Received: April 2020 / Accepted: July 2020
Published on line: December 2020

JeuxDeMots : Un réseau lexico-sémantique pour le français, issu de jeux et d'inférences

Mathieu Lafourcade
LIRMM - Université Montpellier
mathieu.lafourcade@lirmm.fr

Nathalie Le Brun
Imaginat, Lunel
imaginat@imaginat.name

Abstract

The JeuxDeMots project aims to build a broad knowledge base in French, both common sense and specialised, using games, contributory approaches, and inference mechanisms. A dozen games have been designed as part of this project, each one allowing collecting specific information, or to consolidate the information acquired through the other games. This article describes, both qualitatively and quantitatively, 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), annotation of relations (meta-information), semantic refinements (management of polysemy), the creation of clusters allowing the representation of richer knowledge (n-argument relations). Finally, we will describe some complementary acquisition methods such as a bot for endogenous contributions as well as a chatbot making inferences.

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

Résumé

Le projet JeuxDeMots a pour but de construire une base de connaissances de sens commun et de spécialité, en français, à l'aide de jeux, d'approches contributives, et de mécanismes d'inférence. Une dizaine de jeux ont été conçus dans le cadre de ce projet, chacun permettant de collecter des informations spécifiques, ou de consolider les informations acquises via les autres jeux. Dans cet article nous décrirons quantitativement et qualitativement les données 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.), 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). Enfin, nous décrirons des méthodes d'acquisition complémentaires, à savoir ce que peuvent faire des robots effectuant des contributions endogènes sur les données, ainsi qu'un chatbot réalisant des inférences.

Mots-clefs : réseau lexico-sémantique, base de connaissances, GWAPs, relations lexicales, relations sémantiques, représentation des connaissances, connaissances incertaines et floues, acquisition lexicale.

1. Introduction

Les problématiques du TALN[1] nécessitant l'exploitation de ressources lexicales et sémantiques, optimiser la construction de ces dernières s'avère capital. L'automatisation de cette tâche, quelle qu'en soit la méthode, est susceptible d'introduire divers biais dans la ressource construite. Fabre et Bourigault (1998) mettent ainsi en évidence le bruit important que génère l'ambiguïté lexicale dans l'extraction automatique de relations typées entre termes à l'aide de patrons syntaxiques, et lorsqu'on tente de construire une ressource à partir de textes, la nature du corpus choisi peut également biaiser le résultat. Les ressources lexico-sémantiques peuvent aussi être constituées manuellement de diverses manières, et les approches contributives sont intéressantes dans la mesure où elles permettent l'acquisition de données lexico-sémantiques à un coût bien moindre et beaucoup plus rapidement qu'en sollicitant des experts. On parle de « crowdsourcing » (ou peuplonomie) : il s'agit de solliciter des locuteurs, bénévoles ou rémunérés, sans perdre de vue que la ressource obtenue sera de meilleure qualité avec des contributeurs bénévoles, dans la mesure où leur seule motivation est leur intérêt pour la langue, (d'une façon plus générale, nous parlerons d'intelligence collective, voir Fisher, 2009 ; Steyvers, Miller, Hemmer & Lee, 2009). Les GWAPs, *Games With A Purpose* ('jeux avec finalité d'acquisition de ressources' ; cf. (von) Ahn & Dabbish, 2004 ; (von) Ahn, 2006a ; (von) Ahn, Liu & Blum, 2006b ; (von) Ahn & Dabbish, 2008.), illustrent un autre aspect de l'approche contributive, et c'est celle-ci que nous allons développer dans cet article : la ressource se construit par consensus entre les joueurs, qui sont des contributeurs bénévoles, puisque leurs parties alimentent un réseau lexico-sémantique, lequel s'étoffe et se densifie via des mécanismes d'inférences.

Dans cet article, nous évoquerons d'abord les approches ludo-contributives au service de la constitution de ressources lexicales, en nous focalisant sur les réseaux lexicaux et leur intérêt dans les problématiques relatives au TALN. Nous détaillerons ensuite les conditions que doit remplir un GWAP pour permettre la constitution d'une ressource de qualité. Puis nous illustrerons ces considérations générales par l'exemple du projet JeuxDeMots, en donnant une description détaillée de quelques jeux du projet. Enfin, nous présenterons la ressource obtenue, le réseau lexico-sémantique JeuxDeMots, en donnant des caractéristiques quantitatives et en développant certains aspects de sa construction :

[1] TALN : Traitement Automatique des Langues Naturelles.

certaines difficultés liées au jeu seront évoquées, ainsi que quelques particularités de ce réseau, sans oublier une analyse des erreurs liées à la modalité ludique de sa construction. Une dernière partie sera consacrée aux méthodes complémentaires d'alimentation du réseau, via des procédures automatisées d'acquisitions de relations lexico-sémantiques.

2. Les GWAPs : des jeux pour construire des ressources lexico-sémantiques

Bien que la science dite « participative » soit en plein essor (voir comme exemple emblématique Kawrykow, Roumanis, Kam, Kwak, Leung … Waldispühl, 2012), on recense encore relativement peu de GWAPs, et la plupart sont dévolus au traitement de grandes quantités de données (micro-tâches), ou à la résolution de problèmes complexes (macro-tâches) (voir Lafourcade, Le Brun & Joubert, 2015a), plutôt qu'à la production de ressources (à ce sujet, voir par exemple Mihalcea, 2003 ; Chklovski & Gil, 2005). Notons que dans beaucoup de ces « jeux », notamment ceux concernant le traitement de données existantes, l'aspect ludique est réduit, la motivation des joueurs-contributeurs procédant plutôt de l'idée valorisante d'aider la science. Les questions liées à la motivation des contributeurs de ressources et des écueils potentiels du crowdsourcing ont largement été étudiées (Chandler & Kapelner, 2012 ; Mekler, Brühlmann, Opwis & Tuch, 2013 ; Deci, Koestner & Ryan, 1999 ; Fort, Adda & Bretonnel Cohen, 2011 ; Goh & Lee, 2011 ; Kunkel, Loepp & Ziegler, 2018). Quant à la production de ressources lexico-sémantiques ou de connaissances (Siorpaes & Hepp, 2008 ; Kuo, Lee, Chiang, Wang, Shen … Yung-jen Hsu, 2009), peu de véritables jeux y sont consacrés, et les données produites n'étant généralement pas disponibles, on peut difficilement se faire une idée de l'efficacité de l'approche mise en œuvre. De plus, beaucoup de GWAPs en général, et notamment dans le domaine de la collecte de données lexicales ou de connaissances du monde, ont été mis en œuvre ponctuellement, dans le cadre d'un travail de recherche, d'une thèse, et n'existent plus aujourd'hui. C'est le cas de *Common Consensus* (Lieberman, Smith & Teeters, 2007), ou encore *Verbosity* ((von) Ahn, Kedia & Blum, 2006c). *Wordrobe* (Venhuizen, Basile, Evang & Bos, 2013), un ensemble de jeux d'annotations lexicales semble être toujours accessible. D'autres plateformes perdurent comme *Phrase Detective* (Chamberlain, Poesio & Kruschwitz, 2008, 2009) dont l'objectif est de capturer des informations de coréférence, tout comme *PlayCoref* (Hladka, Mırovsky & Schlesinger, 2009). On peut aussi citer *M-Gwap*, fondé sur l'annotation d'états mentaux (Paolizzo, 2019). Le projet JeuxDeMots est à notre connaissance un des rares à s'être maintenu depuis son initiation en 2007, à proposer un portail de GWAPs pérennes, et dont les données sont l'objet de recherche ou d'applications (Joubert & Lafourcade, 2012 ; Lafourcade & Le Brun, 2017 ; Chapuis & Lafourcade, 2017 ; Lafourcade, Joubert & Le Brun, 2017 ; Lafourcade & Joubert, 2018).

C'est WordNet (Miller, Beckwith, Fellbaum, Gross & Miller, 1990 ; Fellbaum & Miller, 1998 ; ainsi que Vossen, 1998 pour EuroWordNet) qui a fait du réseau lexico-sémantique, généraliste ou spécialisé, une ressource incontournable pour les problématiques du TALN, notamment la

désambiguïsation lexicale (Véronis & Ide, 1990). Un réseau sémantique est un graphe permettant de représenter des connaissances. Un réseau lexico-sémantique est un réseau qui permet de relier des connaissances du monde à des informations lexicales sur le vocabulaire qui les véhicule. En général, les relations lexicales ou sémantiques ne sont pas pondérées au sein des réseaux : on ne sait rien de leur « force » ou de leur fréquence. Mais ces caractéristiques, qui s'avèrent fort utiles pour les algorithmes de propagation en désambiguïsation, sont difficiles à déterminer : un réseau peut être pondéré manuellement et/ou en croisant des données avec plus ou moins d'efficacité, ce qui suppose une vérification, manuelle également (Sagot & Fišer, 2008).

Les réseaux lexico-sémantiques comprennent différents types de relations ontologiques, comme les méronymes ou les hyperonymes, ou lexicales, comme les lemmes, ou les locutions. On peut citer HowNet (Dong & Dong, 2006), une ressource lexicale chinois/anglais construite manuellement, avec de nombreux types de relations. CYC (Lenat, 1995) est une base de connaissances générales également construite manuellement et fondée sur la logique, que l'on peut considérer comme un réseau de relations sémantiques. D'autres approches construisent des réseaux de relations associatives, mono ou multilingues, comme BabelNet (Navigli & Ponzetto, 2010), un réseau multilingue de grande taille construit par extraction/compilation automatique des informations de Wikipédia. Certaines ressources constituent de façon implicite un réseau prenant la forme d'espaces sémantiques (Ploux & Victorri, 1998). Des approches fondées sur le crowdsourcing ont été expérimentées afin de vérifier des données comme la similarité sémantique entre termes (Zesch & Gurevych, 2009).

Dans cet article, nous nous focalisons sur l'approche qui consiste à construire une ressource lexico-sémantique à travers la mise en œuvre d'un GWAP : c'est l'approche développée dans le projet JeuxDeMots (Lafourcade, 2007), qui s'est donné pour finalité la constitution d'un réseau lexico-sémantique de grande taille pour le français. Une portion de ce réseau est représentée Figure 1, montrant qu'il est constitué de relations, que nous définirons comme suit :

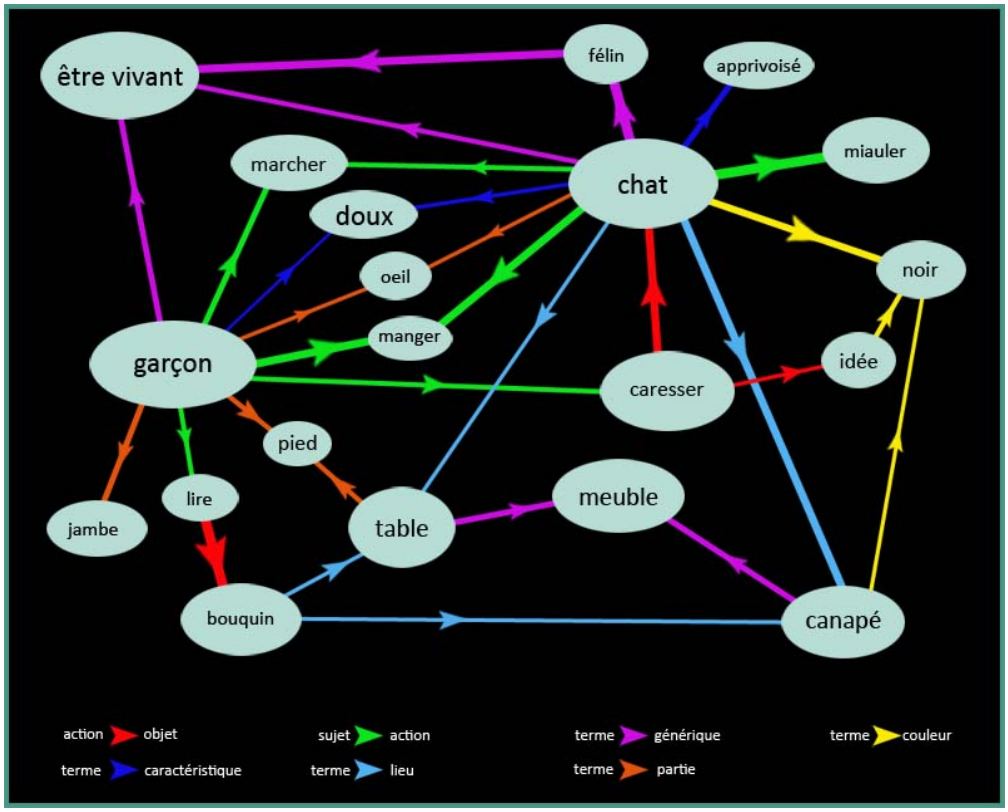


Figure 1. Vue stylisée d'une portion simplifiée du réseau lexico-sémantique JeuxDeMots. Les relations sont orientées, typées et pondérées.

Formellement, une relation est un quadruplet ＜mot1, mot2, type, poids＞ :

- mot1 et mot2, sont des termes du lexique, et en toute généralité, n'importe quelle information symbolique (chaîne de caractères majuscules ne faisant pas partie du vocabulaire et correspondant à une information non ambiguë). Par exemple : *chat, manger, voir,* Nom:Masculin, _INFO-SEM-LIVING-BEING (c'est-à-dire 'être vivant'), _INFO-SEM-PERS ('personne'), etc. ;
- Le type est une information indiquant la nature de la relation, il s'agit d'un symbole : *r-isa* ('est un type de'), *r_has_part* ('a comme partie'), *r_lemma* (''a pour lemme'), *r_agent-1* ('est agent de'), etc. ;
- Le poids est une valeur numérique réelle (positive ou négative).

Sauf quand c'est nécessaire, on n'indiquera pas le poids des relations dans ce qui suit. Éventuellement on indiquera si le poids est positif ou négatif. Une relation de poids nul doit être interprétée comme une absence d'information (soit comme si la relation n'existait pas), et une relation de poids négatif est une relation fausse. Dans la suite de cet article, on écrira une relation de la façon la plus intuitive possible, par exemple :

chat r_isa félin	(chat est un félin)	(poids non précisé, supposé positif)
**chat r_isa ＜0 oiseau*	(chat n'est pas un oiseau)	(poids négatif non précisé)
chat r_agent-1＝150 chasser	(chat peut chasser)	(poids de 150)

Le projet JeuxDeMots vise donc à collecter des relations sémantiques et lexicales au moyen de jeux. Outre l'analyse des données ainsi collectées depuis l'initiation du projet en 2007, nous évoquerons quelques aspects de la conception de ce type de jeux, à la lumière du recul que nous donnent 13 ans d'expérience.

3. Comment concilier l'intérêt ludique et la qualité des ressources produites ?

Pour remplir son objectif, un jeu dont la finalité est l'acquisition d'une ressource lexicale doit concilier deux grands impératifs : plaire aux joueurs pour en faire des contributeurs efficaces, et permettre l'acquisition de données qualitativement et quantitativement satisfaisantes. Bien que ces impératifs soient communs à tous les jeux à finalité d'acquisition de ressources, qu'elles soient lexicales ou pas, très peu d'analyses concernent les modalités de conception d'un GWAP ; certaines les évoquent de façon très générale ((von) Ahn & Dabbish, 2008), d'autres se limitent à une simple description du jeu (Thaler, Siorpaes, Simperl & Hofer, 2011).

3.1. Comment plaire aux joueurs ?

Pour devenir un bon contributeur, un joueur doit être séduit par le jeu au point d'avoir envie de jouer régulièrement, ce qui suppose :

- *Un intérêt ludique* : la condition *sine qua non* du succès tient à la capacité de l'activité proposée à séduire le joueur et à le fidéliser. Il s'agit de faire la synthèse entre renouvellement et répétitivité, de façon à éviter de lasser le joueur tout en lui fournissant un environnement suffisamment familier pour qu'il puisse progresser de façon tangible. Dans le contexte d'un jeu lexical, le renouvellement et la diversité sont assurés par l'infinie variété non seulement du vocabulaire, mais des thématiques et des niveaux de langue, que vient pimenter l'ambiguïté générée par la polysémie ;
- *Des règles faciles à comprendre, intuitives* : les joueurs étant généralement peu enclins à lire de longues explications, les consignes affichées doivent être à la fois brèves, simples, et précises, afin que le joueur sache exactement et immédiatement ce qu'on attend de lui. Ceci n'exclut pas de proposer dans l'interface du jeu un espace dédié aux explications et conseils pour mieux jouer. Dans un jeu lexical, les consignes doivent être comprises par tous les locuteurs, et donc exclure les termes ou concepts dont la compréhension nécessite une formation linguistique poussée ;
- *Une valorisation et de la compétitivité* : un joueur reviendra d'autant plus volontiers qu'il est en compétition avec ses pairs et qu'un mode de classement dynamique et réactif reflète fidèlement son activité. Et plus celle-là est régulière et fréquente, meilleures sont les contributions ; fidéliser le joueur et encourager la compétitivité au moyen d'un classement et d'un système de récompenses et de bonus améliorent donc qualitativement et quantitativement

les contributions. Par ailleurs, certains des jeux du projet JeuxDeMots, comme Totaki, confrontent le joueur au système et le mettent en situation de le « piéger » ce qui non seulement est valorisant, mais donne le sentiment légitime d'avoir « appris quelque chose » à la machine ;

- *Un jeu évolutif, une progression de la difficulté* : plus le joueur joue, meilleures sont ses contributions et plus il doit pouvoir accéder à des parties plus difficiles, avec des consignes de jeu plus précises et plus exigeantes. Pour éviter la lassitude et entretenir l'addiction, il est fondamental que la difficulté croisse avec l'expérience du joueur. Dans un jeu visant à construire un réseau lexical, la difficulté peut être fonction de divers paramètres, comme la nature de la relation entre les termes (les parties sur la relation cause ou conséquence sont plus difficiles à alimenter que celles demandant des idées associées) ou la nature du vocabulaire proposé, qui va de basique et usuel à plus spécialisé/recherché ;
- *Des parties courtes* : un joueur se connectera au jeu d'autant plus facilement et souvent qu'il sait qu'il peut ne faire qu'une ou deux parties et qu'il aura un résultat immédiat, sous forme d'un score qui viendra influer sur son classement. Un jeu dévolu à l'acquisition de ressources lexicales doit ainsi privilégier les micro-tâches à granularité fine et temps de construction faible. Concernant le jeu principal du projet JeuxDeMots, les parties ont une durée limitée par défaut à 1 minute (mais on peut « acheter du temps »), et la durée moyenne d'une session de jeu, estimée par enquête auprès des joueurs est de l'ordre de 10 minutes, mais peut être nettement plus longue chez les joueurs « accro ».

3.2. Comment acquérir facilement une ressource lexicale de qualité ?

Dans le réseau lexico-sémantique JeuxDeMots, la qualité des données tient au fait qu'elles sont le fruit de consensus entre joueurs : une relation entre deux termes n'est, en effet, validée que si elle a été produite par au moins deux joueurs, dans le cadre d'une partie asynchrone qui préserve l'anonymat des joueurs jusqu'au résultat. Ainsi se constitue depuis 2007 une base de connaissances générales, de sens commun (par exemple, un lave-vaisselle peut tomber en panne), auxquelles viennent s'ajouter des connaissances de spécialités (par exemple, la chloroquine agit contre le paludisme). Toujours en ce qui concerne la qualité de la ressource, notons que contrairement à ce qui se passe dans des tâches de crowdsourcing rémunérées (Sagot, Fort, Adda, Mariani & Lang, 2011), tout laisse penser que la motivation des participants et leur désir de « bien jouer » est un facteur de qualité des données produites. De plus, d'un point de vue éthique, il semble bien plus acceptable de solliciter des personnes au moyen d'une proposition ludique que de s'inscrire dans un système à rémunération faible attirant des personnes en situation économique difficile.

Le prix de revient de la ressource est également un critère à considérer, et dans le cadre d'un GWAP, il est faible, puisque limité au coût du développement, puis de la maintenance. Quant à l'évaluation et la valorisation d'une telle ressource, elles font partie intégrante du travail de l'équipe de recherche impliquée dans le projet. Et lorsque le GWAP a du succès, une partie de l'évaluation est

externalisée et gratuite, puisqu'assurée par la communauté de joueurs. Dans le cadre du projet JeuxDeMots, certains joueurs se sont passionnés au point de devenir contributeurs directs et vérificateurs bénévoles.

Enfin, pour acquérir les données d'une base de connaissances générales, le public visé doit représenter la partie la plus large possible d'une population de locuteurs. Or, force est de constater que la nature même du jeu, assimilé à un « jeu de lettres » cible un certain public, puisque dans le cas du projet JeuxDeMots, les joueurs sont majoritairement des joueuses entre 30 et 60 ans, avec un niveau d'études supérieur. Ce biais s'explique en partie par le fait que le jeu a fait l'objet d'une diffusion via des flyers distribués lors de colloques, soit auprès d'un public plutôt plus diplômé que la moyenne. Mais bien que le projet JeuxDeMots ait également été présenté sur les réseaux sociaux et sur des forums de jeux généralistes, il semble que ce type de jeu ait peu de succès auprès d'un public peu diplômé. La nature du jeu est donc déterminante quant au public recruté.

Plus de 6 000 inscriptions de joueurs ont été comptabilisées depuis 2007, et il est également possible de jouer en « invité », sans avoir à être identifié (« logué ») sur la plateforme. Parmi ces joueurs inscrits, une frange non négligeable est constituée de « fidèles », ainsi qualifiés soit parce qu'ils jouent régulièrement et sans interruption depuis leur inscription (une dizaine de personnes), soit parce qu'ils ont joué très intensément pendant une période donnée : plus de 200 joueurs ont joué plus de 1000 parties, et plus de 3 800 joueurs ont fait entre 100 et 1 000 parties. Le nombre total de parties jouées par les joueurs inscrits s'élève à 1,54 million pour le jeu principal du projet. Après avoir joué intensément pendant une période donnée, certains joueurs reviennent sporadiquement. La fidélisation des joueurs a deux grands avantages :

- Jouer souvent et régulièrement permet de découvrir et d'exploiter toutes les subtilités du jeu, ce qui améliore considérablement la pertinence des contributions, donc la qualité de la ressource produite ;
- Il arrive fréquemment que des joueurs fidèles « s'emparent » d'une thématique qui les intéresse particulièrement (médecine, art, mythologie gréco-latine, bande dessinée, littérature, cinéma, botanique…) et pour laquelle ils sont suffisamment compétents pour produire des connaissances de spécialistes.

Corrélativement, on ne peut qu'en déduire les inconvénients d'un turn-over trop rapide : en ne jouant que quelques parties, on passe à côté de toute la progressivité qui permet de s'améliorer en jouant sur des termes et des relations de difficulté croissante, et les contributions restent à un niveau sommaire.

4. Le projet JeuxDeMots: des jeux d'acquisition et des contre-jeux

Le projet JeuxDeMots, décrit dans Lafourcade (2007) et Lafourcade et Joubert (2010), s'est concrétisé à travers le développement d'un jeu principal, JeuxDeMots[2], de jeux d'acquisition annexes et de « contre-jeux », accessibles à travers un portail[3]. Chacun d'entre eux alimente le réseau lexical suivant des modalités particulières et les contre-jeux ont été conçus pour vérifier et préciser les données introduites dans le réseau via le jeu principal, et compenser au moins partiellement les biais induits par les jeux d'acquisition. Le Tableau 1 récapitule l'ensemble des jeux et contre-jeux du projet, avec leurs finalités.

Jeu	Type	Modalité	Finalité
JeuxDeMots	Jeu d'acquisition	Contribution	Construire un réseau lexical pour le français
Totaki	Contre-jeu/jeu d'acquisition	Devinettes et contribution	Vérifier les informations du réseau, rechercher des lacunes et créer de nouvelles relations
AskYou	Contre-jeu	Votes	Vérifier les relations du réseau
AskIt	Contre-jeu	Votes	Vérifier les relations du réseau, en particulier les relations des raffinements sémantiques
Emot	Jeu d'acquisition	Votes et contribution	Associer aux termes du réseau une émotion ou un sentiment
ColorIt	Jeu d'acquisition	Votes et contribution	Associer aux termes du réseau une information de couleur/texture/aspect
SexIt	Jeu d'acquisition	Votes	Associer aux termes du réseau une connotation sexuelle (ou pas)

[2] http://www.jeuxdemots.org/jdm-accueil.php

[3] http://imaginat.name/JDM/Page_Liens_JDMv8.html

PolitIt	Jeu d'acquisition	Votes	Associer aux termes du réseau une connotation politique
LikeIt	Jeu d'acquisition	Votes	Associer aux termes du réseau une polarité positive, négative, ou neutre
Tierxical	Contre-jeu	Votes	Hiérarchiser les poids des relations au sein du réseau
Top 10	Contre-jeu	Votes et contribution	Hiérarchiser les poids des relations au sein du réseau
Yakadirou	Jeu d'acquisition	Votes	Caractériser les relations de type « a pour lieu » par la préposition pertinente
Selemo	Contre-jeu	Votes	Vérifier les relations du réseau

Tableau 1. Liste des jeux du projet, avec leur type, leur(s) modalité(s) et leur finalité. Les jeux dont le nom est en gras sont ceux qui font l'objet d'une description dans cet article.

À cet ensemble de jeux s'ajoute Diko[4], un outil en ligne de visualisation/édition du réseau lexico-sémantique JeuxDeMots. Les utilisateurs peuvent librement contribuer sur les termes et les types de relations qui les intéressent. Ces contributions doivent être validées par un expert avant de rentrer dans le réseau. L'intérêt principal de cet outil est la vérification manuelle des données ainsi que leur correction, mais également la complétion du réseau pour des relations non jouables ou difficiles. Notons que certains anciens joueurs qui ont fini par se lasser du jeu, mais qui restent intéressés par le projet, préfèrent contribuer directement via cet outil.

La présentation d'une entrée dans Diko ressemble un peu à celle d'un dictionnaire ou un thésaurus classique. Les différentes valeurs pour chaque type de relation peuvent être affichées avec des méta informations (indication de termes polysémiques, statistiques, etc.).

[4] http://www.jeuxdemots.org/diko.php



Figure 2. Vue (partielle) de l'outil Diko en mode « visualisation ». Les termes associés polysémiques sont suivis d'un triangle blanc. Les annotations sont en orange.

Nous allons décrire ci-après certains des jeux listés dans le Tableau 1, pour donner un aperçu des différents types (jeu d'acquisition/contre-jeux) et modalités de jeu (contribution/vote).

4.1. JeuxDeMots, le principal des jeux d'acquisition

C'est avec le lancement de JeuxDeMots (JDM) qu'a été initiée la construction du réseau lexical en juillet 2007, amorcée par un noyau de 150 000 termes sans relation entre eux.

JDM est un jeu « ouvert » dont le but est de collectionner des mots, et dont la règle est la suivante : à partir d'un terme-cible et d'une consigne qui apparaissent à l'écran après qu'il a cliqué sur le bouton « jouer », le joueur doit entrer des propositions dans un champ de texte, en temps limité (Figure 2). La consigne concerne un type de relation parmi la cinquantaine qui peuvent faire l'objet de parties. Nous donnons un échantillon non exhaustif de ces types de relations « jouables » dans le Tableau 2.

Type de relation	Nom de la relation	Consigne de jeu associée
Idées associées	r_associated	Donner des idées associées au terme-cible
Sens/signification	r_meaning/glose	Donner des mots évoquant les différents sens possibles du terme-cible
Thèmes/domaines	r_domain	Donner des noms de domaines relatifs au terme-cible (par exemple, *rougeole* relève du domaine de la médecine)
Synonymes	r_syn	Donner des synonymes du terme-cible
Contraires	r_anto	Donner des contraires (antonymes) du terme-cible
Génériques	r_isa	Donner des génériques (hyperonymes) du terme-cible (par exemple *animal* et *mammifère* sont des génériques de *chat*)
Spécifiques	r_hypo	Donner des spécifiques (hyponymes) du terme-cible (par exemple, *mouche, abeille, moustique*, sont des hyponymes d'insecte)
Parties/constituants/éléments	r_has_part	Donner des parties (méronymes) du terme-cible (par exemple, *roues, volant, siège, moteur, pour automobile*)
Tout	r_holo	Donner des "touts" (holonymes) du terme-cible (par exemple *squelette, corps, animal, être humain pour membre supérieur*)
Agent	r_agent	Donner des agents (ou sujets) du verbe-cible (par exemple *marathonien* est un agent de *courir*)

Agent-1 (relation inverse de Agent)	r_agent-1	Donner sous forme de verbes des actions typiques pour le terme-cible (exemple, *chat* et *miauler*, *ronronner*, *griffer*, *manger*…)
Patient	r_patient	Donner des patients (ou objets) du verbe-cible (par exemple *pomme* est un patient de *croquer*)
Patient-1 (relation inverse de Patient)	r_patient-1	Donner des actions typiques dont le terme-cible peut être l'objet (par exemple, *pomme* peut être l'objet de *manger, croquer, éplucher*…)
Lieu	r_lieu	Donner des lieux possibles pour le terme-cible (par exemple, *nid* est un lieu typique pour *œuf*)
Caractéristique	r_carac	Donner des caractéristiques typiques du terme-cible (par exemple, *froide, chaude, liquide, limpide*, pour *eau*)
Caractéristique-1	r_carac-1	Donner des termes qui peuvent avoir la caractéristique-cible (par exemple, *original* peut caractériser un *dessert*, un *projet*, un *homme*…)
Cause	r_causatif	Donner des causes possibles pour le terme-cible (par exemple *négligence* est une cause possible pour *accident*)
Conséquence	r_conseq	Donner des conséquences possibles pour le terme-cible (par exemple, *accident* est une conséquence possible de *négligence*)
But	r_but	Donner des buts possibles pour le terme-cible (par exemple *s'entraîner* ou *entraînement* peuvent avoir pour but *réussir* ou *remporter la victoire*)

Tableau 2. Échantillon non exhaustif des types de relations jouables, notation dans le réseau lexical, et consignes correspondantes. Cette liste est volontairement limitée aux types de relations les plus courantes et exclut les types les plus délicats à jouer, qui seront évoqués plus loin (cf. Section 5.2.1.).



Figure 3. Exemple d'une partie de JDM où le joueur doit indiquer à quoi lui fait penser l'expression « ne pas chercher à savoir ». Il lui reste 22 secondes pour faire des propositions, et les réponses déjà fournies sont *ignorer*, *ne pas vouloir savoir*, *fermer les yeux*, *savoir*, *ne pas s'en mêler*, *ne pas chercher*, et *fermer les yeux*. Les termes en orange (*ne pas chercher* et *ne pas s'en mêler*), ne sont pas (encore) connus du système. Le dernier terme proposé (*fermer les yeux*) a plusieurs sens, *fermer les yeux* 'fermer les paupières' et *fermer les yeux* 'indulgence', et le joueur est invité à en choisir un, s'il le souhaite.

La Figure 3 montre une partie de JeuxDeMots où il est demandé d'énumérer ce que peut évoquer le terme ne pas chercher à savoir (consigne du type de relation « idées associées »). Le joueur a fait 7 propositions en tout (à droite de l'image). À l'issue de la partie, ses propositions sont confrontées à celles d'un autre joueur, stockées dans la base, sur le même terme et la même consigne. Ainsi une partie de JDM se déroule de façon asynchrone, le joueur actif ne connaissant pas l'identité du joueur qui a fourni les réponses auxquelles vont être confrontées ses propres propositions, ce qui rend impossible de tricher en « choisissant » son adversaire. Le pseudo de ce dernier n'est dévoilé qu'avec le résultat, qui est le nombre de propositions communes aux deux joueurs (Figure 3). Ce sont ces réponses communes qui sont mémorisées par le système pour alimenter le réseau lexical, et les relations ainsi acquises sont pondérées en fonction du nombre de paires de joueurs qui les ont proposées. Au sein du réseau, une relation est un quadruplet, constitué par le mot-cible, le mot qui lui a été associé par au moins une paire de joueurs, le type de la relation correspondant à la consigne, et son poids.

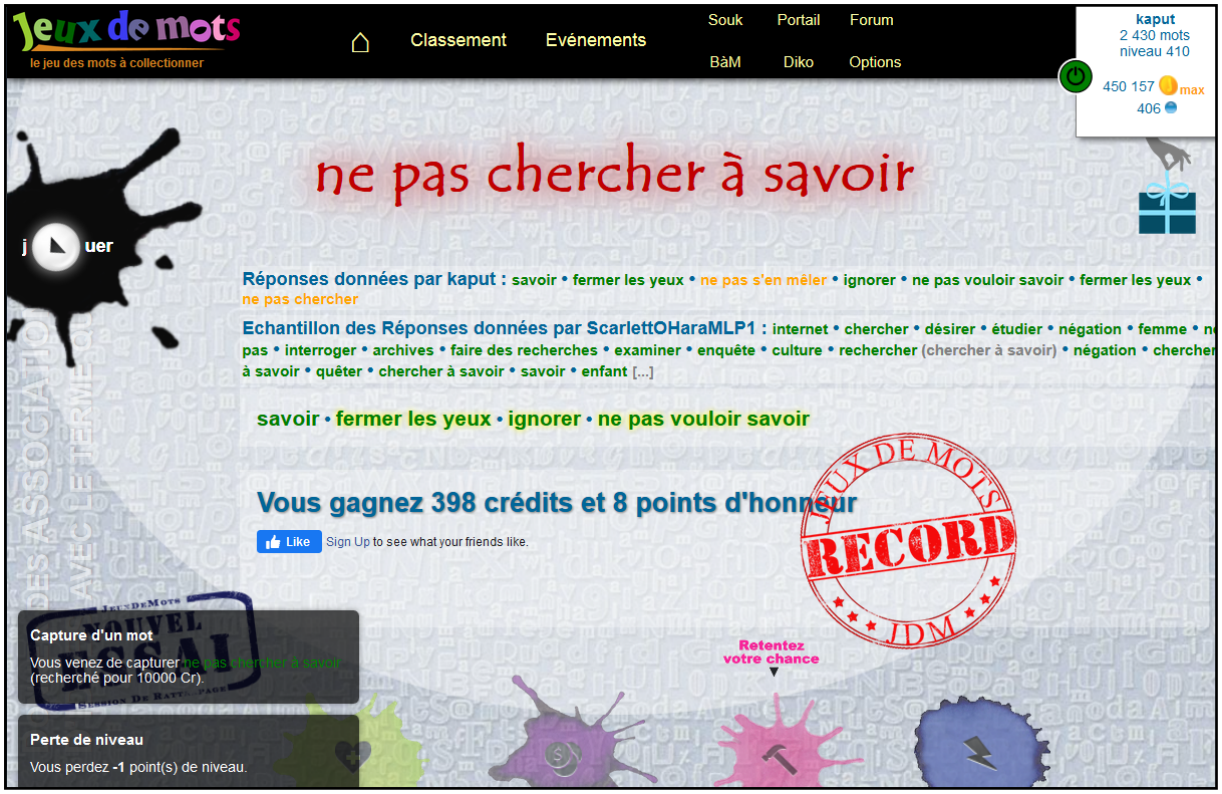


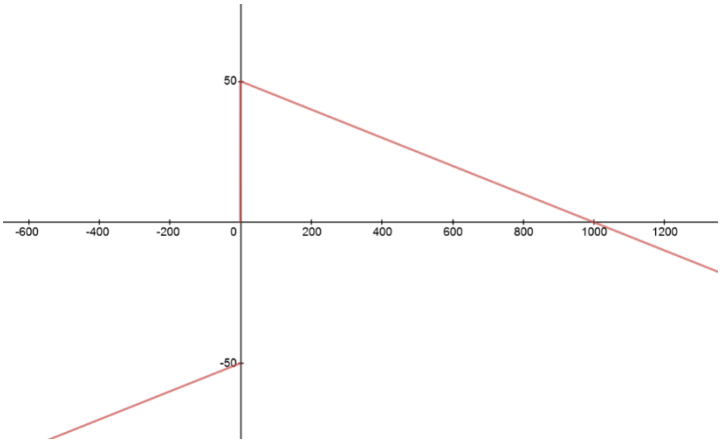
Figure 4. Résultat de partie de JDM sur le terme « ne pas chercher à savoir » avec la relation idées associées. Les deux joueurs ont fourni 4 réponses communes (*savoir*, *fermer les yeux*, *ignorer* et *ne pas vouloir savoir*), et gagnent 398 crédits et 8 points d'honneur.

Le résultat de la partie liste les propositions communes aux deux joueurs, et donne le score sous forme d'un gain en « crédits », qui représentent non seulement la « monnaie » du jeu, mais déterminent le classement principal, et en « points d'honneur », un autre paramètre d'évaluation du joueur.

Les gains en crédits, sont calculés en fonction de l'état du réseau avant la fin de la partie. L'état du réseau est apprécié par le poids des relations entre le mot-cible et les propositions du joueur, et par le type de la relation. La fonction de poids est la suivante :

$$\frac{(1000 - x)}{20} \cdot \mathrm{sign}(x + 1)$$

et a le tracé suivant :



Une relation nouvelle (inconnue dans le réseau) rapporte donc 50 crédits. Une fois devenue « taboue » (c'est-à-dire à partir d'un poids de 1000), une relation commence à faire perdre des crédits. Proposer un terme déjà connu via une relation ayant un poids négatif fait également perdre des crédits. Par exemple, pour la consigne consistant à donner des termes qui sont agents de *voler>déplacement aérien*, le candidat *autruche* fera perdre 54 crédits. En effet, la relation *voler>déplacement aérien r_agent autruche* est pondérée à -75.

Quel que soit le score, le joueur a le choix entre diverses actions[5]: s'il est déçu par le résultat et considère que l'autre joueur a fourni des réponses inadéquates, il peut lui intenter un procès. Il peut aussi l'attaquer, lui lancer une « patate chaude », l'envoûter, investir dans cette partie s'il estime que ses propositions méritent d'être proposées à d'autres joueurs plus performants, retenter la partie. S'il a fait un « jackpot » (un score supérieur ou égal à 200 crédits), il a droit à des actions assimilables à des « récompenses » : ajouter un ami, proposer un nouveau terme, changer de pseudo, associer une image au terme sur lequel il vient de jouer, ou « lancer une patate douce » à un joueur de son choix. Et s'il a fait un score supérieur à la valeur du mot, il remporte le mot, qui s'ajoute à sa collection personnelle, listée dans sa « boîte à mots ». Un mot qui n'appartient à personne a par défaut une valeur de 200 crédits, et un mot possédé par un joueur a la valeur du score qu'il a réalisé pour le remporter.

On remarquera que beaucoup de ces actions favorisent les interactions entre joueurs, de façon à titiller l'esprit de compétition, susciter l'émulation et pousser les joueurs à réagir aux « agressions », donc à jouer : un joueur « attaqué », prévenu par mail, aura à cœur de se « venger », une attaque réussie se traduisant pour la victime par le vol d'un de ses mots. Mais les interactions suscitées peuvent aussi avoir des effets pédagogiques : intenter un procès conduit à rendre publics le terme, la consigne, les réponses des deux joueurs et les griefs du plaignant afin que la communauté des joueurs « rende la justice » sous la forme d'un verdict (coupable, innocent, relaxé au bénéfice du doute), à la suite de

[5] Toutes ces actions et toutes les stratégies pour « bien jouer » font l'objet d'un document d'aide détaillée disponible via un lien dans l'interface du jeu : http://imaginat.name/JDM/AideJDMv1web.pdf

débats et de votes. Cette procédure « juridique », qui se déroule sur un espace dédié du site, améliore, en favorisant les échanges entre joueurs, la compréhension des consignes délicates à interpréter. Enfin, les interactions prennent aussi la forme de « cadeaux » que peuvent se faire les joueurs entre eux, un cadeau étant une partie dont le terme et la consigne sont choisis par le donateur, et donc une incitation à jouer et à rendre la pareille. L'échange de cadeaux est l'un des aspects « coopératifs » du jeu, sous-tendu par une motivation d'ordre social : il s'agit de se montrer sous un jour favorable en offrant des parties intéressantes. Du point de vue du joueur une partie « intéressante » est une partie permettant de donner de nombreuses réponses pertinentes, et qui par voie de conséquence, va enrichir le réseau. Ainsi, JDM bénéficie d'une tension entre coopération et compétition : mieux on joue, plus on augmente son score, mais en favorisant aussi celui de l'autre puisque le score profite aux deux protagonistes ; et la compétition induite par un classement multicritères pousse à se démarquer, ce qui suppose de jouer le plus et le mieux possible.

Parmi les jeux autres jeux d'acquisition, tous consistent à poser une question ouverte au joueur, mais pour certains, la modalité de réponse est mixte : soit le joueur clique sur une des réponses proposées, soit il indique sa propre réponse via un champ de texte, lorsqu'aucune des propositions ne lui convient. Cette modalité mixte concerne Emot et ColorIt, des jeux destinés à associer à des termes-cibles des émotions/sentiments pour le premier, et des informations de couleur et d'aspect pour le second. Les autres jeux d'acquisition sont tous des jeux à votes (également appelés « jeux à clics »). Signalons que cette modalité, du fait qu'elle ne nécessite pas de périphérique de saisie, se prête tout particulièrement au jeu sur smartphone/tablette dans les situations d'attente ou les transports en commun.

4.2. LikeIt, un jeu d'acquisition à votes

Le principe d'un jeu à votes est de répondre à une question en cliquant sur une réponse parmi un choix fermé de propositions. Il est toujours possible de « passer » la question sans répondre. Ce type de jeu permet d'acquérir des informations sur des relations spécifiques via une modalité d'interaction rapide, bien adaptée aux smartphones.



Figure 5. 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.

LikeIt, décrit et analysé dans (Lafourcade et al., 2015), a été conçu pour affecter une polarité aux termes du réseau, ce qui peut être un paramètre intéressant en analyse d'opinion, ou encore pour filtrer des contenus textuels. À la question « est-ce que vous aimez l'idée de X », le joueur peut cliquer sur « oui, j'aime l'idée », « non je n'aime pas l'idée » ou « ça ne me fait ni chaud ni froid », dotant ainsi le terme X d'une polarité positive, négative, ou neutre. Une étude préliminaire a montré que les données de polarité pouvaient être exploitées avec profit dans le cadre de la désambiguïsation lexicale : prises isolément, elles permettent de sélectionner le sens correct d'un terme polysémique en contexte dans 50% des cas. Une approche simple consiste à calculer la polarité moyenne des termes au voisinage du terme-cible polysémique, et ensuite de sélectionner le raffinement du terme-cible se rapprochant le plus de cette polarité moyenne. Vouloir effectuer une désambiguïsation lexicale uniquement avec cette approche n'est pas efficace (puisque limitée à 50% de bons résultats), mais elle constitue néanmoins une heuristique intéressante qui peut être combinée avec profit à d'autres exploitant d'autres informations que la polarité.

Comme c'est le cas pour tous les jeux à votes, LikeIt est un jeu à parties courtes et relance immédiate, et la « récompense » pour le joueur est de voir comment sa réponse se situe par rapport à celles de l'ensemble des joueurs (Figure 5). Le pourcentage de personnes qui ont voté comme lui procure au joueur soit la satisfaction d'avoir « bien » répondu (c'est-à-dire comme la majorité), soit, dans le cas contraire, d'être original, de se démarquer du « commun des mortels ».

Notons que si la polarité est évidente pour des termes comme *maladie* ou *sourire*, c'est un paramètre très relatif dans de nombreux cas : pour certaines personnes, le terme *vitesse* va être associé à l'ivresse et à la liberté et perçu positivement, tandis que pour d'autres, il est connoté négativement en raison de l'idée de danger qui peut lui être associée. Il en est de même pour *voiture* : neutre dans l'absolu, le terme peut être associé soit à l'idée de pollution, d'accident, d'embouteillage, et jugé négativement, soit représenter un espace et un moyen de liberté et recevoir un vote positif.

Pour proposer au joueur des termes intéressants à « polariser », la sélection ne se fait pas au hasard dans le réseau, ce qui donnerait une proportion trop importante de termes neutres, avec le risque de lasser le joueur. La sélection est pseudo-aléatoire : à partir de quelques termes polarisés manuellement soit positivement, soit négativement, pour amorcer le processus, l'algorithme en sélectionne un aléatoirement, et propose au joueur, soit ce terme, soit un terme pris au hasard parmi ceux qui lui sont reliés. C'est une façon simple de sélectionner des candidats pertinents pour la polarisation. LikeIt a ainsi permis de polariser plus de 24 000 termes en 3 mois, et à ce jour, au 1er septembre 2020, plus de 1 million de termes sont dotés d'une polarité.

Les contre-jeux sont pour la plupart des jeux à votes. Ce sont des outils de consolidation et de vérification de la ressource, puisque le joueur doit donner son avis sur un énoncé issu des données du réseau, en cliquant sur une des propositions affichées (c'est pourquoi on parle aussi de « jeux à clics »). On entend par « consolidation » la confirmation d'une relation dont le poids est faible, via une question fermée sur sa validité. Ces jeux, qui proposent des parties rapides à relance immédiate aux joueurs, sont un moyen de modifier la ressource avec une granularité très fine : les réponses modifient par exemple le poids d'une relation avec un incrément très faible. L'intérêt majeur des jeux à votes est de compenser les erreurs liées au forçage[6] de réponses, qui a souvent lieu dans les jeux dont la modalité est la contribution directe via un champ de texte.

Depuis l'introduction de ce jeu, et comme pour tous les autres, nous avons mené des enquêtes régulières auprès des joueurs volontaires afin de savoir comment ils percevaient le jeu, quels étaient leurs comportements de joueurs, et de collecter d'autres informations relevant de la notion de profil psychologique. Dans le cas de LikeIt, le public interrogé était majoritairement constitué de jeunes adultes, étudiants, également répartis quant au genre. Il est apparu que le choix de la polarité d'un terme dépendait fortement de l'expérience de vie du joueur, ou du contexte immédiat. Plus le joueur est fatigué, plus ses réponses reflètent fidèlement ce qu'il pense. Certains joueurs ont avoué ne pas répondre sincèrement, mais orienter leurs réponses pour tenter d'être dans la majorité, ou au contraire de se démarquer. On a observé statistiquement, avec les données de joueurs identifiés, que ces comportements ne durent pas et ne biaisent pas les données recueillies.

Concernant LikeIt, 135 personnes ont été interrogées entre novembre 2015 et janvier 2020. Le temps moyen de jeu est d'environ 15 minutes, ce qui correspond à une cinquantaine de réponses environ.

La modalité de jeu à votes concerne aussi les contre-jeux, comme Askit, que nous présentons ci-après.

---

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

4.3. AskIt, un contre-jeu à votes



Figure 6. Exemples de deux écrans consécutifs de AskIt, le deuxième montrant au joueur comment se situe sa réponse par rapport à celles des autres joueurs et la progression d'une barre de score. Cet écran est remplacé par une nouvelle question sans intervention du joueur (relance immédiate).

Askit est un jeu de questions portant globalement sur les relations entre les termes, en particulier en ce qui concerne les termes polysémiques, avec trois réponses possibles : « oui », « non », et « possible, mais bof » quand la question n'est pas pertinente. On peut aussi « passer » la question. Les réponses permettent la vérification, et le cas échéant, la consolidation de la relation sur laquelle porte la question : dans l'exemple de la Figure 6, la réponse négative permet de valider la polysémie de *voilier* et de confirmer la différence entre les raffinements *voilier>bateau* et *voilier>oiseau*. Les parties sont ultrarapides, une nouvelle question arrive immédiatement après l'écran de résultat, sans intervention du joueur, qui est poussé à continuer par une double motivation (Figure 6, écran de droite) :

- Curiosité de savoir comment se situe sa réponse par rapport à celle des autres joueurs,
- Désir de voir progresser la barre d'état et de la remplir pour passer au niveau supérieur[7].

Le phénomène de « trollage » (fait de donner sciemment une réponse erronée) est négligeable (inférieur à 1%), et d'une façon générale, le jeu est apprécié, notamment parce que les questions liées à la polysémie s'avèrent souvent insolites et amusantes : « Est-ce qu'un coucou (avion) a des plumes ? »

[7] Les « niveaux » sont fonction du nombre de questions auxquelles le joueur a répondu depuis le début du jeu.

4.4. Totaki, un jeu de devinettes pour évaluer et apprendre

Totaki est un jeu à part dans la mesure où c'est un contre-jeu, mais c'est aussi un jeu d'acquisition, puisqu'il permet de pallier les lacunes du réseau et de l'enrichir d'informations inédites émanant des joueurs.

Lancé en 2010, il a été conçu comme un outil susceptible d'évaluer la qualité du réseau : le joueur tente de faire deviner un mot au système en lui fournissant des indices, qui sont soit des relations typées (X a comme conséquence *frissonner*), soit des idées associées (X est en rapport avec *fièvre*). L'identification du mot par le système indique que le réseau lexical est correctement renseigné pour ce terme. Si ce n'est pas le cas, le joueur est invité à entrer la solution et le système mémorise les indices fournis et crée les relations correspondantes. Signalons que le fait que certains joueurs tentent de « piéger » le système, ce qui est fréquent, conduit à enrichir le réseau en créant des relations inédites. L'outil de vérification se double donc d'un mécanisme d'apprentissage et de détection d'incohérences, lorsque les informations fournies par le joueur sont en contradiction avec les données du réseau. Une incohérence typique est d'avoir une relation *A r B* négative[8] et une relation *B r-1 A* positive (*r-1* étant la relation inverse de *r*). Cependant, on ne peut savoir automatiquement laquelle des deux relations est correcte.

Cet outil de vérification du réseau remplit parfaitement sa fonction initiale, puisque le mot pour lequel le joueur fournit des indices est deviné dans 82% des cas (Lafourcade, Joubert, Schwab & Zock, 2011), alors qu'un humain soumis à la même procédure identifiera le terme dans 48% des cas. Depuis son lancement, plus de 296 000 parties de Totaki ont été jouées.

L'algorithme de résolution de Totaki est le suivant : Totaki démarre avec L, une liste vide, destinée à accueillir les « candidats », c'est-à-dire les réponses potentielles, avec leurs poids. Rappelons que le joueur donne comme indice soit un mot, qui va alors être traité par défaut comme une idée associée au terme à trouver, soit une relation spécifique, par exemple « est un animal » (*r_isa animal*) ou « fait partie d'une voiture » (*r_holo voiture*). Au premier indice I1 fourni, sont ajoutés à la liste L les termes qui lui sont reliés dans le réseau (assoc(I_1)), soit par le type de relation idées associées lorsque le joueur a fourni un simple terme comme indice, soit par le type de la relation qu'il a utilisé pour donner l'indice. Si parmi ces termes figure l'indice I_1, il est retiré de la liste L : L ← assoc(I_1) – {I_1}. La liste L est alors triée par poids décroissants, et le terme en première position est proposé comme réponse et retiré de la liste. Si le joueur indique que c'est la bonne réponse, le processus est terminé. Dans le cas contraire, le joueur est invité à fournir un deuxième indice I_2. Dans la liste L, ne sont alors retenus que les termes en commun avec ceux associés à I_2 (assoc(I_2)) moins I_1 et I_2 : L ← L ∩ assoc(I_2) – {I_1, I_2}. Après tri de cette nouvelle liste par poids décroissants, le terme en première position est proposé et retiré de la liste. Si ce n'est pas la bonne réponse, le processus se poursuit de la sorte jusqu'à

[8] Nous rappelons qu'une relation pondérée négativement est une relation considérée comme fausse.

l'identification du mot à deviner, ou jusqu'à ce que la liste L soit vide. Totaki tente alors une heuristique tolérante : il reprend les termes associés par les différents indices fournis et considère leur somme plutôt que leur intersection, en retirant les indices eux-mêmes (I_1 à I_n) et les termes déjà proposés. Après tri de cette liste, le premier terme est proposé. Cette heuristique est tentée deux fois de suite avant que Totaki ne « jette l'éponge ».

Augmenter le nombre d'étapes de l'heuristique tolérante dégrade les résultats. Le compromis idéal entre « trouver la bonne réponse » et « minimiser le nombre d'étapes » est effectivement d'utiliser l'heuristique, mais pas plus de deux fois. Signalons l'intérêt d'exploiter dans cet algorithme les poids des relations au sein du réseau (plutôt que de considérer les associations comme booléennes). En effet, des associations « fortes » arrivent rapidement (par exemple, *outil* est lié à *marteau* avec un poids élevé).

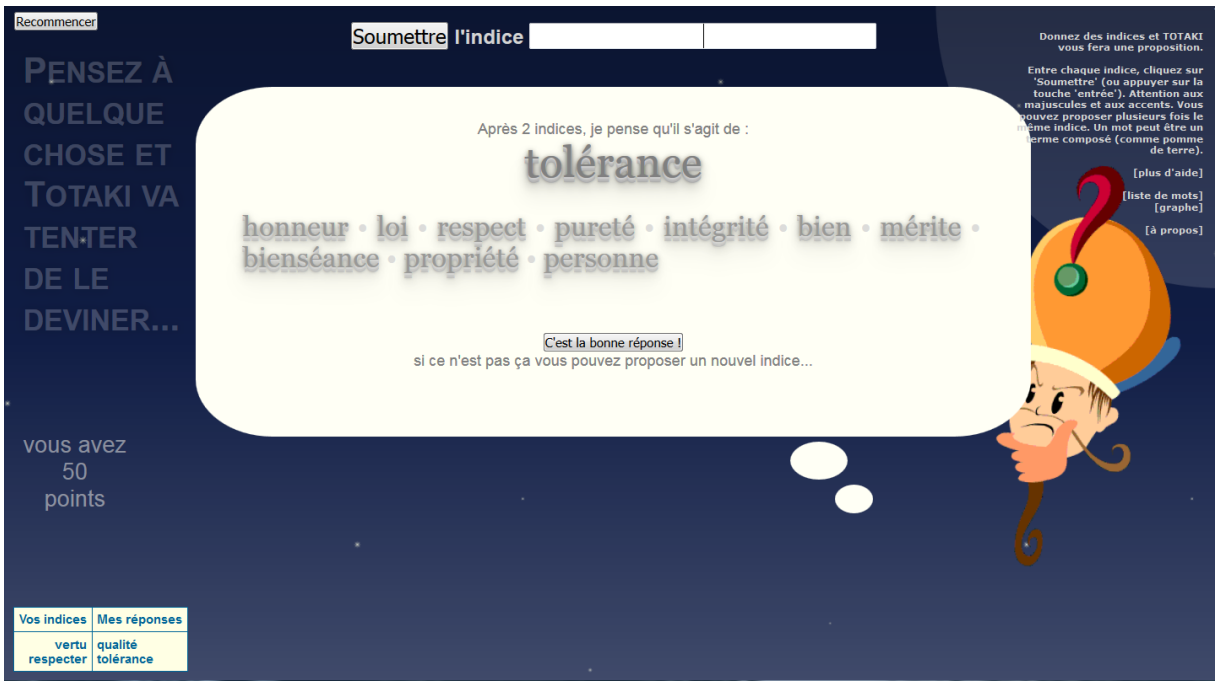


Figure 7. Partie de Totaki. Les indices donnés par le joueur sont affichés en bas à gauche. Le joueur a fourni les indices *vertu* et *respecter*, et Totaki a proposé *qualité* suite au premier indice, et *tolérance* suite au deuxième. Si *tolérance* est effectivement le terme à trouver, le joueur va cliquer sur la mention « c'est la bonne réponse », et il gagnera des points. Si ce n'est pas la bonne réponse, il sera invité à fournir un 3e indice.

5. La ressource obtenue, et les modalités ludiques de sa constitution

5.1. Présentation quantitative et qualitative du réseau lexico-sémantique JeuxDeMots

Depuis son initiation, en 2007, à partir d'une base de 150 000 termes, le réseau lexical s'est considérablement étoffé, puisqu'il compte actuellement, au 1er septembre 2020, plus de 4,2 millions de termes reliés par plus de 328 millions de relations lexico-sémantiques, qui se répartissent en environ 120 types. Parmi ces types de relations, une bonne moitié peuvent faire l'objet d'une partie dans JDM,

le jeu principal d'acquisition du projet JeuxDeMots. Initialement, les concepteurs avaient proposé un ensemble d'une vingtaine de types de relations « jouables ». Sous l'impulsion des joueurs assidus, et en concertation avec eux, de nombreux types de relations ont été ajoutés depuis.

145 272 435 associations libres	1 488 526 domaines (champ lexical)	2 035 767 synonymes
5 815 contraires (antonymes)	8 9 686 594 génériques (hyperonymes)	9 700 377 spécifiques (hyponymes)
8 529 456 parties (méronymes)	8 805 433 tout (holonymes)	1 036 524 lieux typiques
3 113 226 agents (verbes)	113 084 patients (verbes)	116 482 causes
41 421 instruments (verbes)	1 034 686 caractéristiques typiques	16 004 plus intense (magn)

Tableau 3. Nombre d'occurrences des relations les plus représentatives du réseau lexico-sémantique JeuxDeMots obtenues au 1er septembre 2020.

Certains types de relations comme les lemmes (1 192 316 relations dans le réseau) et les catégories morphosyntaxiques (19 099 610 relations dans le réseau) sont trop peu variables pour être intéressantes à jouer. Ces relations sont donc établies soit manuellement par des contributeurs bénévoles, via Diko, le dictionnaire contributif associé au projet JeuxDeMots (cf. Section 4), soit de façon automatique via des mécanismes d'inférence sur le réseau (approches endogènes) ou des mécanismes d'extraction d'information depuis des contenus textuels (approches exogènes).

5.2. Quelques aspects relatifs à la modalité ludique

5.2.1. Jeu principal : quelques consignes délicates

Le mode de jeu par défaut de JDM invite le joueur à jouer sur le type de relation d'idées associées avec du vocabulaire courant, mais comme nous l'avons évoqué précédemment dans la présentation du jeu (cf. Section 4.1.), les joueurs inscrits ont accès à un très grand nombre de types de relations et de thématiques. Nous évoquons ci-après quelques types de relations plus ou moins délicats à jouer, d'après les résultats des parties, et les débats entre joueurs sur les divers espaces de discussion.

- *Le type de relation sens/significations* : pour les termes polysémiques, ce ne sont pas les joueurs qui fournissent les raffinements sémantiques, mais ils sont invités, via des parties sur le type de relation *sens/significations*, à fournir des gloses, c'est-à-dire des mots représentatifs de chacun des sens du terme. Les gloses les plus pertinentes seront ensuite validées par un administrateur en tant que raffinements sémantiques. Les parties *sens/significations* sont jugées difficiles, mais intéressantes : en temps limité, il faut à la fois répertorier toutes les acceptions possibles d'un terme, et trouver les mots les plus adéquats pour les désigner. Par exemple, il est plutôt difficile d'identifier en 1 minute (durée d'une partie par défaut) tous les sens possibles d'un verbe très polysémique comme *prendre* (32 raffinements sémantiques dans le réseau lexico-sémantique JeuxDeMots) ou encore *tourner* (19 raffinements), et de trouver des termes pour illustrer chaque sens identifié.

- *Les types de relation cause et conséquence* : les retours des joueurs, et leurs débats lors de procès, par exemple, ou sur le forum associé au jeu, montrent qu'il existe une certaine confusion entre ce qui relève de la cause et ce qui relève de la conséquence : *le désir est-il la cause ou la conséquence de l'amour ?* Il est même parfois impossible de trancher : *la perte de conscience* peut aussi bien être la cause que la conséquence d'un *accident*.
- *Les types de relation instance et hyponyme* : les joueurs confondent fréquemment les instances et les hyponymes, et considèrent que les secondes incluent les premières : ainsi, *Idéfix* est caractérisé aussi bien comme hyponyme de *chien*, que comme instance. En revanche, un hyponyme n'est quasiment jamais donné pour une instance.
- *Les types de relation rôle agentif et implications agentives* : malgré les explications et les exemples, distinguer les *implications agentives*, qui sont les étapes qui président à la construction d'un artefact, du *rôle agentif*, i.e. le ou les verbe(s) transitif(s) qui lui donne(nt) naissance, n'est pas intuitif pour les joueurs : on va par exemple trouver *battre les blancs en neige* comme rôle agentif de *soufflé au fromage*, alors que c'est une implication agentive, le rôle agentif étant plutôt *faire, préparer ou confectionner*.
- *Le type de relation plus intense (magn)* donne lieu à des interprétations très subjectives, *plus intense* étant souvent confondu avec un mélioratif : pour *dessert*, on va trouver *bon dessert* qui est un mélioratif, les termes satisfaisant à la relation *plus intense* étant plutôt *dessert copieux* ou *ronde des desserts*.
- *Le type de relation termes de la même famille*, pour lequel nous attendions des termes de la même famille lexicale, formés par dérivation morphologique, a dû être scindé en deux nouvelles relations : *termes étymologiquement apparentés* et *termes dérivés morphologiquement*. En effet, les joueurs donnaient pêle-mêle les deux catégories de termes : au terme *jour*, étaient associés pour ce type de relation *journalier, journellement,* ou *journal*, mais aussi *diurne, lundi* ou *quotidien*. Cependant, d'autres erreurs subsistent, évaluées par échantillonnage à environ 5% des termes associés par ces types de relations, ce qui est supérieur aux taux d'erreur relevés pour les autres : dans les dérivés morphologiques de *statistiques*, on trouvera par exemple *statique*.

5.2.2. Quelques observations relatives au comportement des joueurs

Il est manifeste que le désir de trouver absolument des réponses pousse de nombreux joueurs à « forcer la consigne », comme nous l'avons évoqué plus haut (cf. Section 4.3.), et ceci bien qu'il soit possible de « passer » la partie, ou encore d'indiquer qu'il n'y a pas de réponse pertinente possible, comme c'est stipulé dans chaque partie, sous la zone de saisie : *si vous ne savez pas répondre, il faut passer la partie. Si vous estimez qu'il n'y a pas de réponse possible, mettez ****.

Certains joueurs ont ainsi tendance à proposer des réponses qu'ils estiment « approchantes », quand ils ne trouvent rien qui réponde bien à la consigne. On trouvera ainsi *stylo, cahier, traitement de texte*, comme conséquences de *faire ses devoirs*, alors qu'il s'agit plutôt d'instruments. Notons que le travers

inverse existe aussi : il consiste à considérer qu'il n'y a pas de réponse possible et à mettre *** quand on ne trouve rien. Une évaluation par échantillonnage montre que le forçage de réponses est à l'origine d'environ 0.1% de relations erronées, mais que ce biais est largement contrebalancé par les contre-jeux. Nous entendons par « échantillonnage » la sélection automatique et aléatoire d'un petit pourcentage de relations à vérifier. La vérification de ces relations est assurée manuellement par des membres de l'équipe de recherche impliquée dans le projet, et par des volontaires bénévoles.

Grâce aux espaces d'échanges que représentent les procès, le « Blabla » (une zone de chat présente dans toutes les pages du jeu à l'exception des parties), et le forum, les joueurs interagissent et s'entraident, les plus anciens ayant manifestement à cœur d'éclaircir vis-à-vis des nouveaux la sémantique de certaines consignes délicates, notamment quand ils « ratent » une partie à cause d'une mauvaise interprétation de la consigne par l'autre joueur. Ceci vient confirmer la dimension « collaborative » du jeu principal, déjà évoquée plus haut, dans sa présentation (cf. Section 4.1.). Il est manifeste que plus on joue sur une longue période, mieux on joue, en raison de l'expérience, mais aussi de l'interaction avec les pairs.

Chaque partie (nous désignons ici par ce terme l'ensemble des réponses données pour un terme et une consigne) est stockée avec le pseudonyme du joueur, ce qui permet de calculer et de fournir au joueur, outre le nombre de parties jouées depuis son inscription, divers paramètres de classement dont le *niveau d'expertise* (capacité à faire plus que son score moyen), et l'*efficacité* (nombre moyen d'intersections[9] par partie). L'efficacité moyenne des joueurs en 2020 est de 2,76, tandis que le meilleur joueur a une efficacité de 11. Le nombre moyen de mots proposés par les joueurs est égal à 9,6, mais les bons joueurs en proposent fréquemment une cinquantaine. Par ailleurs, et toujours pour favoriser la coopération et la convivialité, chaque joueur inscrit peut découvrir la liste de ses *bff* (Best Friends Forever), les joueurs avec lesquels il a réalisé ses meilleurs scores.

L'analyse qui suit est issue de multiples enquêtes menées régulièrement depuis le lancement du projet, sous forme de questions ouvertes portant soit sur un jeu spécifique, soit sur l'ensemble de jeux proposés. Certaines enquêtes ont pris la forme d'un questionnaire par mail envoyé aux joueurs inscrits, d'autres d'interviews de groupes de joueurs, typiquement des groupes d'étudiants. La participation aux enquêtes a toujours été libre et n'a jamais fait l'objet d'une quelconque rétribution.

Pourquoi les joueurs jouent et qui sont-ils ? Comme nous l'avons brièvement évoqué à la Section 3.2, tous jeux confondus, les femmes représentent environ 65% des joueurs, avec une moyenne d'âge supérieure à 40 ans. La durée moyenne de fréquentation de la plateforme JeuxDeMots est d'environ 4 semaines. Un petit nombre de joueurs (environ une vingtaine) jouent de façon plus ou moins régulière depuis plus de 8 ans. Les joueurs interrogés déclarent tous un intérêt, à des degrés divers, pour les choses de la langue : le vocabulaire, la sémantique, la linguistique, etc. À notre

[9] Une « intersection » est un terme en commun avec l'autre joueur.

connaissance, aucun joueur indifférent à ces questions n'a fréquenté la plateforme JeuxDeMots plus que le temps de quelques parties.

Par ailleurs, des jeux sur le langage sont globalement perçus comme une activité intellectuelle, donc « noble », ce qui fournit une forme de justification à ceux qui s'y adonnent de façon intensive : non seulement ils n'ont pas l'impression culpabilisante de perdre leur temps à « jouer », mais en plus, la richesse des thématiques de jeu au sein du portail JeuxDeMots est telle qu'ils affirment se cultiver.

Ces aspects relatifs à la perception des jeux semblent unanimement partagés, quels que soient l'âge et le sexe. En outre, le fait que les jeux proposés soient gratuits (totalement gratuits, pas de *pay to win* ou de *pay to fast*, ou autre incitation mercantile), qu'ils soient les instruments d'un projet de recherche libre et ouvert, que les données recueillies soient totalement accessibles et libres de droit, sont autant de facteurs qui attirent et séduisent des joueurs potentiels, et confèrent au projet JeuxDeMots une bonne réputation et une dimension éthique.

Conformément aux observations de Mekler, Florian, Klaus et Alexandre (2013), qui ont étudié l'effet des points sur la gamification, les joueurs réguliers du jeu principal sont extrêmement stimulés par les différents classements (crédits, points d'honneur, nombre de mots possédés, nombre de parties faites, niveau d'expertise, efficacité, etc.). Ces classements fournissent l'opportunité de se comparer aux autres, ce qui est un facteur motivationnel important, bien que relativement commun à une grande majorité des jeux, et des jeux en ligne en particulier. Dans le jeu principal, découvrir les réponses de l'autre en fin de partie peut se révéler valorisant (quand on a fourni plus de réponses et/ou qu'on les estime de meilleure qualité), mais aussi instructif (quand l'autre a fourni des réponses inédites et pertinentes). La nature du matériau manipulé (la langue) joue un rôle fondamental dans ces perceptions.

La nature des données manipulées et les modalités ludiques participent aussi fortement au côté addictif des jeux de la plateforme. Quel que soit le type de jeu et sa modalité, attendre pour visualiser ce qu'a répondu un autre joueur, ou quel couple mot/consigne, ou quelle question fermée va proposer l'Intelligence Artificielle (le moteur qui sélectionne les données de jeu proposées lors des parties), ou encore ce qu'a répondu la majorité, est accrocheur et pousse à rejouer, notamment dans le cas des jeux à votes. Certains joueurs ont ainsi déclaré avoir joué compulsivement à certains jeux à votes (Askit, LikeIt, SexIt, etc.), pendant des soirées entières, via leur smartphone. Quelques joueurs jouent au jeu principal depuis 2007 de façon régulière, avec occasionnellement des périodes d'activité intense lorsque de nouveaux joueurs se mettent à jouer avec assiduité et menacent de les dépasser dans le classement.

Nous avons découvert des phénomènes de biais (positifs) intéressants dans certains jeux. Par exemple, dans le jeu principal, il arrive que des joueurs fournissent dans un premier temps les réponses qui leur viennent spontanément, puis celles qu'ils supposent pouvoir être données par d'autres. Pour des joueurs assidus et habitués, cela peut se traduire par un nombre conséquent de propositions. Ce type de comportement s'observe notamment pour les parties impliquant des réponses potentiellement

genrées, par exemple le type de relation « Agent, » pour le terme *faire le ménage*, et consiste à fournir les réponses qui pourraient être données par quelqu'un du sexe opposé. La majorité des joueurs étant des femmes, le réseau comporte un nombre important de termes et de relations particulièrement sexistes : certaines joueuses proposent en effet les réponses qu'elles pensent que les hommes donneraient, et quand ces joueuses se rencontrent via une partie, elles font entrer dans le réseau ces « relations sexistes », et ce, sans qu'aucun homme n'ait été impliqué. Mais le but du projet étant de construire un panel d'associations aussi large que possible pour des processus d'analyse automatique de textes, et non d'effectuer des études de psycholinguistique, ce type de biais ne constitue pas un problème, mais sert au contraire le but recherché.

5.3. Particularités du réseau lexical JeuxDeMots

5.3.1. Polysémie et raffinements

Les raffinements sont les représentations des différents sens d'un terme, sous la forme *terme>glose*, la *glose* étant un mot ou une expression qui illustre le sens du terme. Ainsi, le terme *avocat* a deux raffinements : *avocat>fruit* et *avocat>justice*. Mais un raffinement peut lui aussi être polysémique et avoir des raffinements, qui sont des sous-raffinements du terme d'origine.

Ainsi, le terme *train* présente deux niveaux de raffinement : on dénombre en effet 12 sens, soit 12 raffinements, parmi lesquels le raffinement *train>chemin de fer* admet lui-même 3 raffinements : *train>chemin de fer>jouet, train>chemin de fer>convoi ferroviaire*, et *train>chemin de fer>mode de transport*.

5.3.2. Relations inverses

Parmi les types de relation existant dans le réseau lexico-sémantique JeuxDeMots, on trouve un nombre important de paires « relation/relation inverse », comme les relations générique et hyponyme (*r_isa/r_hypo*), holonyme et méronyme (*r_holo/r_has_part*), possesseur et possédé par (*r_own/r_own-1*), a pour agent et agent de (*r_agent/r_agent-1*). On peut se demander pourquoi ne pas garder qu'un seul type de relation pour une paire donnée, ce qui réduirait la taille de la ressource. Plusieurs éléments de réponses peuvent être apportés :

- Les poids entre une relation *A r B* et *B r-1 A* ne sont pas forcément les mêmes. Cette caractéristique est importante pour toutes les applications où lister et ordonner les termes reliés par une relation est nécessaire. Par exemple, les requêtes *$x r_isa animal* (x est un animal) et *animal r_hypo $x* (animal peut avoir comme spécifique x) pourront donner des ensembles de termes semblables, mais ces termes seront listés dans un ordre différent. Quand dans un processus automatique, il s'agit de sélectionner l'objet le plus représentatif d'un type particulier (l'objet qui vient généralement le plus vite à l'esprit des gens) cette propriété est

essentielle. Par exemple, pour un locuteur français à qui on demande de citer un outil, l'outil cité en premier est en général *marteau*, et non pas *tournevis*. Il est important que les données du réseau puissent refléter cela ;

- La redondance inhérente à ces paires de types de relations offre de nombreuses opportunités de détections d'erreurs, comme des contradictions du type *A r B* a un poids positif tandis que *B r-1 A* a un poids négatif. Les incohérences de ce type sont très bien détectées automatiquement. En revanche, elles seront corrigées, soit manuellement, soit par les joueurs eux-mêmes : l'Intelligence Artificielle qui examine le réseau et détecte les problèmes est en effet capable de proposer des parties qui pourront les résoudre.

5.3.3. Pondérations négatives

Les relations peuvent être pondérées négativement, ce qui a un triple intérêt :

- Représenter des exceptions, par exemple que le manchot ne vole pas, bien que ce soit un oiseau, ce qui s'écrit : *voler > déplacement aérien r_agent < 0 manchot* ou *manchot > oiseau de mer r_agent-1 < 0 voler*
- Représenter des informations négatives, par exemple qu'un homme n'a pas d'utérus, ce qui s'écrit : *homme > être humain de sexe masculin r_has_part < 0 utérus*
- Conserver dans le réseau sous forme de relations négatives des erreurs qui présentent un intérêt, parce qu'elles sont fréquentes : beaucoup de gens pensent à tort que les araignées sont des insectes. Cette erreur est mémorisée dans le réseau sous la forme *araignée r_isa < 0 insecte*.

Cette relation erronée est également annotée *souvent cru vrai*, ce qui nous amène à expliquer ce que sont les annotations au sein du réseau lexico-sémantique JeuxDeMots.

5.3.4. Annotations

Il est possible, pour les concepteurs/administrateurs du jeu, ainsi que pour les anciens joueurs qui sont devenus des contributeurs assidus et réguliers, d'ajouter des méta-informations aux relations sous forme d'annotations, les termes d'annotations étant des mots présents dans le réseau. Par exemple les relations *nez rouge r_holo clown* ou *nez rouge r_holo ivrogne* (*nez rouge* est une partie de *clown/d'ivrogne*) sont annotées fréquent, tandis que *nez rouge r_holo individu* est annotée *possible*, tout comme la relation *nez rouge r_carac factice* (*nez rouge* peut avoir factice comme caractéristique) (Figure 8). Comme le montre la figure, les annotations sont des relations qui relient non pas deux termes, mais une relation et un terme. Les termes d'annotations sont variés : *fréquent, possible, rare, improbable, toujours vrai,* etc. L'annotation *non pertinent* concerne des relations qui sont vraies dans l'absolu, mais non informatives, comme, par exemple, *Blanche-Neige r_has_part yeux*, le fait d'avoir des yeux étant vrai pour tout mammifère, voire pour tout animal.

Les annotations peuvent jouer un rôle déterminant dans les mécanismes d'inférence, par exemple en servant de filtres lors de la recherche d'inférences déductives : une annotation *non-pertinent* sur une des prémisses arrête le processus et oblige l'algorithme à en trouver une autre ; par exemple, sans cette annotation *non-pertinent* la relation *animal r_has_part atomes* (animal a comme partie atomes) pourrait conduire par inférence déductive à la création d'une relation *X r_has_part atomes* pour tous les *X* qui sont des hyponymes d'*animal*, ce qui n'a pas d'intérêt. Ainsi, cette annotation bloque le mécanisme lors d'un processus d'apprentissage, ce qui évite de saturer la base de connaissances avec des relations qui sont vraies, mais peu utiles. En revanche, le processus n'est pas bloqué lors d'une analyse de texte, mais son résultat ne sera pas mémorisé dans le réseau.

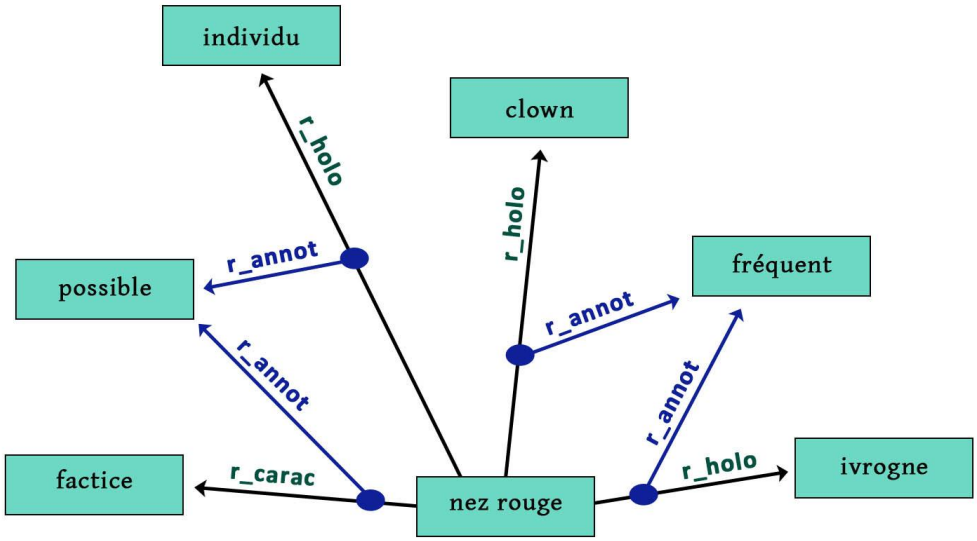


Figure 8. Des annotations (*r_annot*) dans le réseau lexico-sémantique JeuxDeMots : Les clowns et les ivrognes sont fréquemment pourvus d'un nez rouge. Un nez rouge peut être factice.

Les annotations peuvent être introduites par des contributeurs via l'outil Diko (cf. Section 4.), l'outil Hélix (cf. Section 6.2) ou par des joueurs, avec certains jeux de la plateforme (Askit, Yakadirou, etc.).

5.3.5. Représentation de connaissances via des termes agrégés

On peut aussi créer au sein du réseau des agrégés, qui sont des objets lexicaux qui représentent des vérités factuelles ou des hypothèses issues de la connaissance du monde, composés à partir des termes du réseau, ou récursivement à partir d'autres agrégés. Par exemple, il est possible de représenter des connaissances comme « un médecin peut soigner un blessé » sous la forme :

médecin r_agent-1 [soigner r_patient blessé]

Cet agrégé constitue un nœud du réseau. Il est encodé sous une forme compacte non ambiguë :

médecin [peut] soigner [objet] blessé

Et avec la forme suivante dans la base de connaissances :

::>66:59071>24:14722071 (médecin [peut] *id:14722071*)
id:14722071 correspondant à::>14:7781>29:72805 (soigner [objet] blessé)

Dans cette notation, « id » désigne l'identificateur, qui est une valeur numérique associée de façon unique à un nœud ou une relation.

66 est l'id de la relation *r_head* (tête d'un syntagme nominal) avec comme argument *médecin* (id:59071) et 24 est l'id de la relation *r_agent-1* qui demande un syntagme verbal (dont l'id est ici id:14722071).

14 est l'id de la relation *r_patient* avec comme argument *blessé* (id :7781) et 29 indique que l'agrégé a pour prédicat *soigner* (id:72805).

Les caractères : et > servent de séparateurs et facilitent le décodage automatique de l'information contenue dans l'agrégé.

Ces nœuds particuliers que sont les agrégés peuvent être associés à d'autres nœuds via des relations (tout comme les termes « standard »). Par exemple :

médecin [peut] *soigner* [objet] *blessé r_lieu hôpital, clinique, champ de bataille, route, rue*, …

Les agrégés potentiels sont créés automatiquement par le système à partir des informations du réseau : par exemple, comme la relation *lion r_agent-1 dévorer* (un lion peut dévorer) existe, le système va créer (entre autres) l'agrégé verbal *dévorer [sujet=r_agent] lion* et proposer une partie sur cet agrégé avec la consigne correspondant au type de relation « Patient » (*r_patient*) afin de savoir ce que peut dévorer un lion. En pratique, une modalité du jeu donne au joueur la possibilité de choisir une partie parmi 5 propositions de parties avec des agrégés. Le joueur en choisit une, et c'est ainsi que des relations sont ajoutées aux agrégés.

5.4. La question des erreurs

Nous avons évalué statistiquement[10] depuis 2015 à 0.4% le taux de relations fausses, les relations « discutables[11] » représentant quant à elles, moins de 1,7%. Nous sommes en mesure de repérer automatiquement environ 77% des relations erronées. Il s'agit entre autres d'incohérences du type : A est l'hyperonyme de B et B est l'hyperonyme de A alors que A et B sont plutôt des synonymes, mais

[10] L'échantillonnage est effectué automatiquement et l'évaluation est manuelle. Cette évaluation est faite assez régulièrement (tous les 6 mois environ) par des volontaires. Les relations fausses sont ensuite corrigées.

[11] Dans ce contexte, on considère une relation comme discutable si au moins deux personnes (des locuteurs natifs du français) n'arrivent pas à s'accorder sur sa validité lors d'une discussion à ce sujet. Il ne s'agit pas d'un défaut de compétence, les ressources dictionnairiques ou encyclopédiques étant disponibles.

aussi d'incohérences liées à la polysémie. Ainsi, s'il est vrai qu'une *plante* a un *pied*, il n'en est pas moins exact qu'un *pied* a une *plante*. Il n'y a plus d'incohérence lorsqu'on utilise les raffinements respectifs de *pied* et *plante* : *plante>botanique r_has_part pied>botanique* et *pied>anatomie r_has_part plante>anatomie*. S'il est facile de détecter automatiquement ce genre d'incohérence, il est nettement plus difficile de les corriger automatiquement.

Les quelque 23% de relations fausses qu'on ne détecte pas, car il n'y a pas de contradiction ou d'incohérence détectable automatiquement de façon simple, sont en général dues à des défauts d'informations au sein du réseau. Comme nous avons aussi environ 12% de relations détectées automatiquement comme fausses, alors qu'elles sont justes, nous ne procédons pas à une invalidation systématique des relations que le système estime fausses. Les contre-jeux permettent de résorber assez facilement les relations erronées de faible poids, soit environ 3% de l'ensemble des relations détectées comme telles.

6. Méthodes d'acquisition complémentaires

6.1. Robots de contribution endogènes

Dans l'optique de consolider le réseau lexical JeuxDeMots, et éventuellement de découvrir/produire de nouvelles informations, de nombreux robots logiciels ont été mis au point dans le cadre du projet JeuxDeMots. Ces robots (ou bots) ont une activité qualifiée d'endogène dans la mesure où elle ne fait pas appel à des ressources extérieures au réseau (il existe par ailleurs une activité exogène, assurée par d'autres bots). Ces bots peuvent avoir les activités suivantes (liste non exhaustive) :

- Vérifier qu'une relation et sa relation inverse sont cohérentes (il ne doit pas exister une relation à poids positif et son inverse à poids négatif). Par exemple, si *x r_isa y* est positive alors *y r_hypo x* doit aussi être positive ;
- Dans le cadre de la polysémie, vérifier ou faire vérifier aux joueurs si les relations d'un raffinement donné sont valides pour les autres raffinements du terme. Si ce n'est pas le cas, annoter, au niveau du terme général, cette relation comme contrastive (exemple : *avocat r_agent-1 plaider*, cette relation est contrastive, puisque non valide pour le raffinement *avocat>fruit*) ;
- Fournir des contributions (qui devront être validées) sur l'ensemble des types de relations. Les mécanismes mis en jeu sont variés, et relèvent d'inférences déductives, inductives et abductives. Ces mécanismes peuvent être selon les cas, soit codés « en dur » soit le résultat d'apprentissages statistiques/symboliques ;
- Découvrir des règles par apprentissage et, une fois qu'elles ont été validées (soit par un expert soit indirectement par des non-experts (les joueurs), les appliquer sur le réseau. Les règles peuvent être par exemple :

- *$x like %que & $w r_pos Adj: => $x r_pos Adj:Fem+SG & $x r_pos Adj:Mas+SG*
  (Si un adjectif se termine par « que », alors il est aussi bien féminin que masculin)
- *$x r_carac ovipare => $x r_agent-1 pondre>zoologie*
- *$x r_isa être humain => $x r_has_part cœur*
- …

Ces règles sont appliquées à intervalles réguliers sur le réseau lexico-sémantique JeuxDeMots afin de produire de nouvelles relations ;

– Proposer des raffinements si les génériques d'un terme sont mutuellement exclusifs, le terme étant alors nécessairement polysémique. Par exemple :

- *$x r_isa être humain & $x r_isa aliment* (exemple : *avocat*)
- *$x r_isa poisson & $x r_isa oiseau* (exemple : *colin*)
- *$x r_isa film & $x r_isa roman* (exemple : *les Liaisons dangereuses*)
- ...

En moyenne, une relation introduite par des joueurs permet de produire automatiquement 56,3 relations via des mécanismes de raisonnement. Il a été évalué par les porteurs du projet que moins de 1/10 000 des relations produites de cette manière est fausse (la plupart du temps l'erreur provient d'une polysémie non clairement identifiée).

Est-il désirable de faire augmenter la taille réseau lexico-sémantique à l'aide de mécanismes d'inférence ? Le temps d'accès à une relation du réseau est fonction de log(N), N étant le nombre de relations contenues dans ce réseau. Lors d'une analyse sémantique de texte, les processus de choix de relations sémantiques entre termes du texte exploitent les connaissances du réseau, et quand une information n'est pas directement présente, des mécanismes d'inférences sont invoqués. Ces mécanismes d'inférence effectuent en moyenne une vingtaine d'accès au réseau (au minimum 2 et au maximum une centaine). On comprend bien dans ces conditions l'intérêt de minimiser le nombre d'inférences à réaliser, et donc de maximiser le nombre de cas où la réponse est directement disponible dans le réseau (qu'elle soit négative ou positive).

6.2. Hélix, un chatbot raisonneur

Hélix[12] est un chatbot permettant de tester l'ensemble des connaissances contenues dans le réseau lexico-sémantique JeuxDeMots. Hélix a les propriétés suivantes :

– L'interrogation peut se faire en pseudo-langage naturel ;

[12] http://www.jeuxdemots.org/rezo-ask.php?text=1

- Tout ce qui est dit/demandé à Hélix est mémorisé afin de faire l'objet de vérifications et de raisonnements. En particulier, Hélix proposera à l'Intelligence Artificielle de JeuxDeMots de faire vérifier par les joueurs les relations/questions qui lui auront été soumises ;
- Quand Hélix ne trouve pas directement la réponse à la question dans le réseau JeuxDeMots, il entreprend de calculer la réponse la plus probable à l'aide de mécanismes d'inférences, soit, par ordre de priorité: la déduction (il existe un générique pour lequel la relation existe), l'induction (il existe un spécifique pour lequel la relation existe), la transitivité (pour les relations pour lesquelles c'est pertinent, par exemple A est dans B et B est dans C alors A est dans C), l'abduction (trouver des exemples proches similaires), et l'association d'idées.

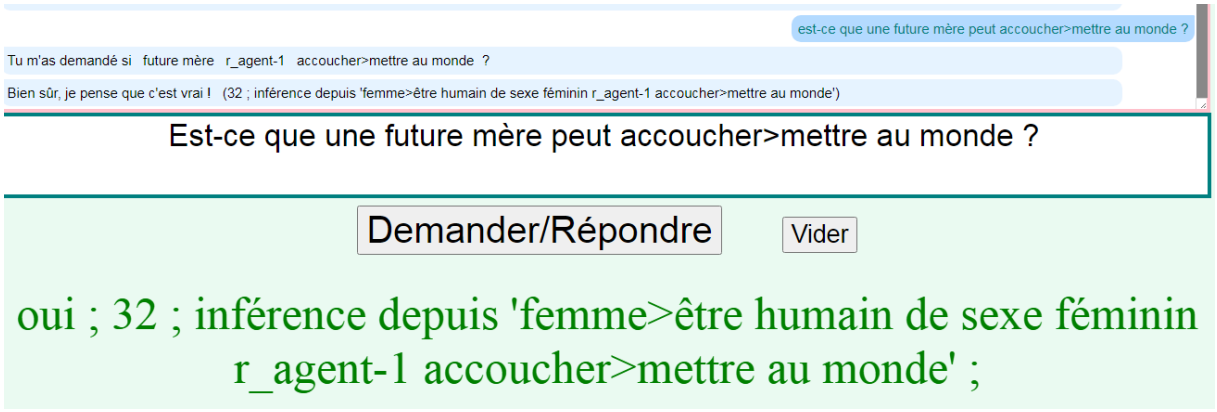


Figure 9. Interface de communication avec Hélix avec un exemple de question impliquant un raffinement. La réponse est issue d'une inférence. De haut en bas, nous avons le rappel de la conversation en cours, la zone de saisie pour la question en cours, la réponse donnée par le système.

Hélix peut gérer les raffinements ainsi que les termes agrégés. Hélix tient compte des relations fausses (à poids négatif) lors des raisonnements visant à produire la réponse.

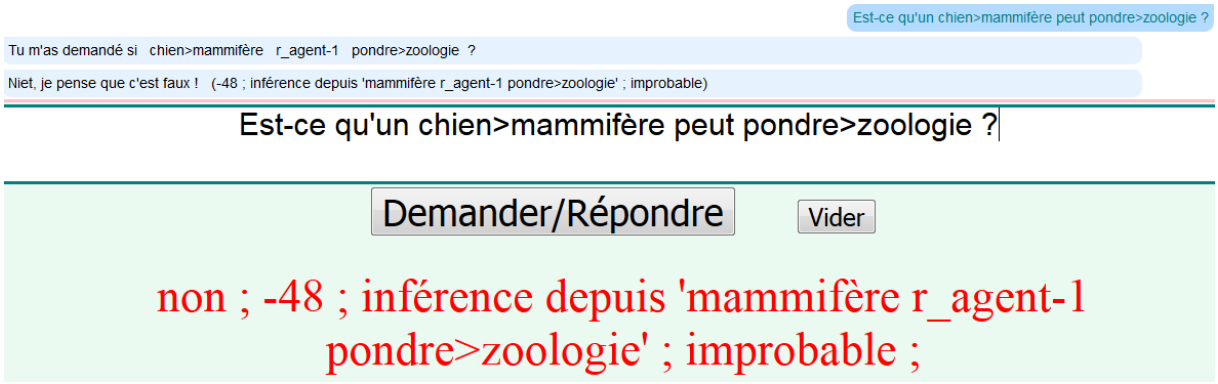


Figure 10. Exemple de question posée à Helix impliquant deux raffinements. La réponse est issue d'une inférence dont le résultat est négatif (avec annotation).

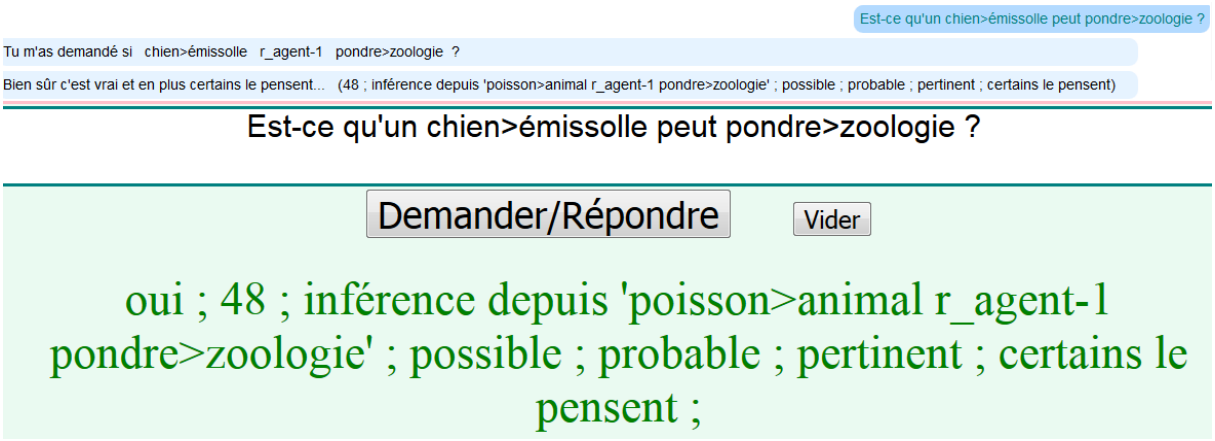


Figure 11. Exemple de question posée à Helix impliquant deux raffinements. La réponse est issue d'une inférence dont le résultat est positif (avec annotation).

6.3. Extraction depuis des textes

Plusieurs applications d'extraction de relations sémantiques depuis des textes ont été mises au point en exploitant le réseau lexico-sémantique JeuxDeMots. Par exemple, à partir du texte suivant :

> Un poisson vivipare est un poisson qui se reproduit par viviparité, c'est-à-dire en mettant au monde des alevins et non en pondant de œufs comme le font les ovipares. Cependant, il est plus juste de parler d'ovoviviparité et de poisson ovovivipare. En effet, si l'embryon vivipare se nourrit dans le sein de sa mère, l'embryon ovovivipare se nourrit dans l'œuf, qui éclot un peu avant la naissance. (Wikipedia, *s.v. poisson vivipare*)

On obtient les relations suivantes, qui sont placées en attente de validation dans le réseau lexico-sémantique JeuxDeMots.

(1) (274682)-(68467) poisson vivipare ISA poisson ---- relation déjà VRAIE dans le réseau C (106) = ' . un poisson_vivipare est un poisson qui se reproduit par viviparité , c'est-à-dire en mettant au_monde'

(2) (16503)-(68467) se reproduire HAS-AGENT poisson inférence oui; 61; inférence depuis 'mimer r_agent_de animal'; certains le pensent; C (125) = ' . un poisson_vivipare est un poisson qui se reproduit par viviparité , c'est-à-dire en mettant au_monde des alevins et_non '

(3) (112239)-(153665) pondre HAS-PATIENT oeufs ---- relation déjà VRAIE dans le réseau C (135) = ' viviparité , c'est-à-dire en mettant au_monde des alevins et_non en pondant de oeufs comme le font les ovipares . Cependant, il est '

(4) (112239)-(265249) pondre HAS-AGENT ovipares ---- relation déjà VRAIE dans le réseau C (135) = ' viviparité , c'est-à-dire en mettant au_monde des alevins et_non en pondant de oeufs comme le font les ovipares . Cependant , il est '

(5) (140318)-(83193) parler CONCERNING ovoviviparité inférence ne sait pas; certains le pensent; 0; C (135) = ' les ovipares . Cependant, il est plus juste de parler d'ovoviviparité et de poisson_ovovivipare. En_effet, si l'embryon vivipare '

(6) (92260)-(105621) embryon CARAC vivipare ---- relation déjà VRAIE dans le réseau C (129) = ' d' ovoviviparité et de poisson_ovovivipare . En_effet , si l' embryon vivipare se nourrit dans le sein de sa_mère , l' embryon '

(7) (47466)-(92260) se nourrir HAS-AGENT embryon inférence C (131) = ' de poisson_ovovivipare . En_effet , si l' embryon vivipare se nourrit dans le sein de sa_mère , l' embryon ovovivipare se nourrit'

(8) (47466)-(0) se nourrir HAS-AGENT embryon vivipare inférence C (131) = ' de poisson_ovovivipare . En_effet , si l' embryon vivipare se nourrit dans le sein de sa_mère , l' embryon ovovivipare se nourrit '

(9) (33882)-(7430) se nourrir LOCATION-OF-ACTION sein inférence oui; 61; inférence depuis 'manger r_action_lieu balcon'; certains le pensent; C (139) = ' de poisson_ovovivipare . En_effet , si l' embryon vivipare se nourrit dans le sein **de sa_mère , l' embryon ovovivipare se nourrit dans l' '**

(10) (7430)-(2459806) sein IS-A-PART-OF mère inférence ne sait pas; certains le pensent; 0; C (121) = ' En_effet , si l' embryon vivipare se nourrit dans le sein de sa_mère , l' embryon ovovivipare se nourrit dans l' oeuf , '

(11) (92260)-(94339) embryon CARAC ovovivipare ---- relation déjà VRAIE dans le réseau C (114) = ' vivipare se nourrit dans le sein de sa_mère , l' embryon ovovivipare se nourrit dans l' oeuf , qui éclot un peu '

(12) (33882)-(116175) se nourrir LOCATION-OF-ACTION oeuf inférence ne sait pas; certains le pensent; 0; C (116) = ' dans le sein de sa_mère , l' embryon ovovivipare se nourrit dans l' oeuf , qui éclot un peu avant_la_naissance . '

Certaines relations extraites existent déjà dans le réseau lexico-sémantique. L'outil Hélix est sollicité afin d'évaluer si une relation qui n'existe pas déjà dans le réseau est crédible ou pas. On remarquera que dans l'exemple ci-dessus le contexte textuel de la relation extraite est fourni. Le terme embryon vivipare est inconnu du réseau (id = 0) (voir ligne *oui* ci-dessus), mais est proposé à l'insertion.

7. Conclusion

Ce tour d'horizon du projet JeuxDeMots met en évidence la pertinence de l'approche GWAPs, ou jeux à finalité, combinés avec des processus automatisés d'inférences, pour l'acquisition de ressources lexico-sémantiques de bonne qualité pour le français. La difficulté d'une telle approche tient non seulement à concevoir un véritable jeu, mais surtout à rendre interdépendants, à travers l'activité ludique, l'intérêt du joueur et la qualité des données générées. Un autre enjeu d'importance, non abordé ici, est le liage d'une telle ressource, construite par peuplonomie ludifiée, à d'autres ressources existantes (Tchechmedjiev, Mandon, Lafourcade, Laurent & Todorov, 2017). Par ailleurs, la ressource construite peut être utilisée dans le cadre de l'apprentissage du vocabulaire (par des enfants ou en seconde langue) en relation avec des jeux dits "sérieux" (voir Lafourcade & Zampa, 2009 et Zampa & Lafourcade, 2010).

Le projet JeuxDeMots établit nettement que des connaissances de sens commun, aussi bien que des données proprement lexicales, si l'on exclut celles qui requièrent une expertise linguistique, peuvent être efficacement collectées via des jeux tous publics. Par ailleurs, l'approche ludique permet aussi de remplacer avantageusement, que ce soit qualitativement ou quantitativement, un travail de caractérisation manuelle confié à des spécialistes : nous avons ainsi conçu un jeu, LikeIt qui permet de caractériser le vocabulaire avec une information de polarité, et les résultats, obtenus très rapidement,

sont parfaitement conformes aux attendus. Dans un même ordre d'idée, citons aussi ColorIt, (Lafourcade et al., 2014), qui permet de collecter des informations de couleurs, ou PolitIt (Tisserant & Lafourcade, 2015) qui explore la connotation politique potentielle du vocabulaire courant.

L'intérêt de cet ensemble de jeux dans la construction d'un réseau lexico-sémantique, mais aussi son originalité, est d'être le siège d'une autoévaluation permanente de la ressource construite via une forme de boucle vertueuse : ce sont les joueurs, qui par l'intermédiaire des contre-jeux, évaluent et rectifient les données qu'ils ont contribué à produire, en répondant à des questions fermées de façon à valider ou invalider les informations recueillies. Et cette approche ludo-contributive résout également la question de l'échantillonnage : il est effectué par les joueurs eux-mêmes, qui choisissent, via les parties terme/relation qu'ils s'offrent en cadeaux, ce qui leur semble intéressant à renseigner, ce qui est la garantie d'une ressource lexico-sémantique reflétant fidèlement la vie quotidienne et l'ensemble des connaissances qui la structurent.

Bibliographie

Ahn, L. von, & Dabbish, L. (2004). Labelling Images with a Computer Game. In *CHI 2004: connect: April 24-29, Vienna, Austria: conference on human factors in computing systems: conference proceedings* (pp. 319-326). New York: Association for Computing Machinery Press. ISBN 978-1-58113-702-6

Ahn, L. von (2006a). Games with a purpose. *Computer 39*(6), 92-94. https://doi.org/10.1109/MC.2006.196

Ahn, L. von, Liu, R., & Blum, M. (2006b). Peekaboom: a game for locating objects in images, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 55-64). New York: Association for Computing Machinery Press. https://doi.org/10.1145/1124772.1124782

Ahn, L. von, Kedia, M., & Blum, M. (2006c). Verbosity: a game for collecting common-sense facts, In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 75-78). New York: Association for Computing Machinery Press. https://doi.org/10.1145/1124772.1124784

Ahn, L. von, & Dabbish, L. (2008). Designing Games With a Purpose. *Communication of the ACM 51*(8), 58-67. https://doi.org/10.1145/1378704.1378719

Chamberlain, J., Poesio, M., & Kruschwitz, U. (2008). Phrase Detectives: A Web-based Collaborative Annotation Game. In S. Auer, S. Schaffert & T. Pellegrini (Eds.), *Proceedings of I-SEMANTICS '08: International conference on semantic systems* (pp. 42-49). Graz: Verlag der Technischen Universität Graz. http://hdl.handle.net/10419/44448

Chamberlain, J., Poesio, M., & Kruschwitz, U. (2009). A demonstration of human computation using the Phrase Detectives annotation game. In P. Bennett (Ed.), *HCOMP '09: Proceedings of the ACM*

SIGKDD Workshop on Human Computation (pp. 23-24). New York: Association for Computing Machinery Press.

Chandler, D., & Kapelner, A. (2012) Breaking monotony with meaning: Motivation in crowdsourcing markets. *Journal of Economic Behavior & Organization, 90*, 123-133. https://doi.org/10.1016/j.jebo.2013.03.003

Chapuis, M., & Lafourcade, M. (2017). Identifying Polysemous Words and Inferring Sense Glosses in a Semantic Network. In C. Gardent & C. Retoré (Eds.) *Proceedings of the 12th International Conference on Computational Semantics (IWCS 2017)*, Computing Natural Language Inference (CONLI) workshop. Montpellier, France.

Chklovski, T., & Gil, Y. (2005). Improving the design of intelligent acquisition interfaces for collecting world knowledge from web contributors. *In K-CAP '05: Proceedings of the 3rd international conference on Knowledge capture* (pp. 35-42). https://doi.org/10.1145/1088622.1088630

Deci, E., Koestner, R., & Ryan, R. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological bulletin 125*(6), 627-668. https://doi.org/10.1037/0033-2909.125.6.627

Dong, Z., & Dong, Q. (2006). *HowNet and the Computation of Meaning.* WorldScientific Publishing Co., NJ, USA. https://doi.org/10.1142/5935

Fabre, C., & Bourigault, D. (2006). Extraction de relations sémantiques entre noms et verbes au-delà des liens morphologiques. In *Proceedings of TALN'06*, Leuven, Belgique, (pp. 121-126).

Fellbaum, C., & Miller, G. A. (1998). *WordNet*. Cambridge MA: MIT Press.

Fisher, L. (2009). *The Perfect Swarm: The Science of Complexity in Everyday Life*. NY: Basic Books.

Fort, K., Adda, G., & Bretonnel Cohen, K. (2011). Amazon mechanical Turk: Gold mine or coal mine? *Computational Linguistics*, *37*(2). https://doi.org/10.1162/COLI_a_00057

Goh, D., & Lee, C. (2011). Perceptions, quality and motivational needs in image tagging human computation games. *Journal of Information Science 37*(5), 515-531. https://doi.org/10.1177/0165551511417786

Hladka, B., Mırovsky, J., & Schlesinger, P. (2009). Play the language: play coreference. In *Proceedings of the ACL-IJCNLP 2009 Conference Short Papers*, (pp. 209-212). Stroudsburg, PA: Association for Computational Linguistics. https://doi.org/10.5555/1667583.1667648

Joubert, A., & Lafourcade, M. (2012). A new dynamic approach for lexical networks evaluation. In *LREC'12: Eight International Conference on Language Resources and Evaluation*. Istanbul, Turkey. European Language Resources Association (ELRA). ISBN 978-2-9517408-7-7

Kawrykow, A., Roumanis, G., Kam, A., Kwak, D., Leung, C., Wu, C., Zarour, E., Sarmenta, L., Blanchette, M., & Waldispühl, J. (2012). Phylo: a citizen science approach for improving multiple sequence alignment. *PloS one*, *7*(3). https://doi.org/10.1371/journal.pone.0031362

Kunkel, J., Loepp, B., & Ziegler, J. (2018). Understanding latent factors using a GWAP. In *Proceedings of the Late-Breaking Results track part of the Twelfth ACM Conference on Recommender Systems (RecSys'18)*. Vancouver, BC, Canada.

Kuo, Y., Lee, J-C., Chiang, K-Y., Wang, R., Shen, E., Chan, C-W., & Yung-jen Hsu, J. (2009). Community-based game design: experiments on social games for commonsense data collection. In *Proceedings of the ACM SIGKDD Workshop on Human Computation, HCOMP'09*, (pp. 15-22). New York: Association for Computing Machinery Press. https://doi.org/10.1145/1600150.1600154

Lafourcade, M. (2007). Making People Play for Lexical Acquisition. In *Proceedings of the SNLP 2007, 7th Symposium on Natural Language Processing*. Pattaya, Thailande. https://hal-lirmm.ccsd.cnrs.fr/lirmm-00200883

Lafourcade, M., & Zampa, V. (2009). PtiClic: a game for vocabulary assessment combining JeuxDeMots and LSA. In *Proceedings of the CICLing (Conference on Intelligent text processing and Comptational Linguistics)*. Mexico.

Lafourcade, M., & Joubert, A. (2010). Computing Trees of Named Word Usages from a Crowdsourced Lexical Network. *Investigationes Linguisticae, XXI*, 39-56. https://doi.org/10.14746/il.2010.21.3

Lafourcade, M., Joubert, A., Schwab, D., & Zock, M. (2011). Évaluation et consolidation d'un réseau lexical grâce à un assistant ludique pour le « mot sur le bout de la langue ». In *Proceedings of the TALN'11* (pp. 295-306). Montpellier, France. https://hal.archives-ouvertes.fr/hal-00661659

Lafourcade, M., Le Brun, N., & Zampa, V. (2014). Les couleurs des gens. In *Proceedings 20e Conférence sur le Traitement Automatique du Langage Naturel – TALN 2014* (pp. 103-113). Marseille, France. https://hal-lirmm.ccsd.cnrs.fr/lirmm-01471671

Lafourcade M., Le Brun, N., & Joubert, A. (2015a). *Jeux et intelligence collective – résolution de problèmes et acquisition de données sur le Web.* ISTE éditions. ISBN: 978-1-78406-052-7.

Lafourcade, M., Le Brun, N., & Joubert, A. (2015b). Vous aimez ?... ou pas ? LikeIt, un jeu pour construire une ressource lexicale de polarité. In *22e Conférence sur le Traitement Automatique des Langues Naturelles – TALN 2015*. Caen, France. https://hal-lirmm.ccsd.cnrs.fr/lirmm-01471662

Lafourcade, M., Joubert, A., & Le Brun, N. (2017). Si les souris étaient des reptiles, alors les reptiles pourraient être des mammifères ou Comment détecter les anomalies dans le réseau JDM ? In *24rd French Conference on Natural Language Processing – TALN 2017* (pp. 158-164), Orléans, France. ATALA https://www.aclweb.org/anthology/2017.jeptalnrecital-court.20

Lafourcade, M., & Le Brun, N. (2017). Extracting semantic relations via the combination of inferences, schemas and cooccurrences. In *Proceedings of the International Conference on Recent Advances*

in Natural Language Processing – RANLP 2017 (pp. 417-423). Varna, Bulgaria. https://doi.org/10.26615/978-954-452-049-6_055

Lafourcade, M., & A. Joubert (2018). Si beaucoup d'oiseaux volent, peut-on en déduire que tous les oiseaux volent ? Production endogène de règles déductives dans un réseau lexico-sémantique. In *Proceedings of the TALIA – PFIA 2018*. Nancy, France. https://hal.archives-ouvertes.fr/hal-02269036/document

Lenat, D. (1995). CYC: A large-scale investment in knowledge infrastructure. *Communications of the ACM*, *38*(11), 33-38. https://doi.org/10.1145/219717.219745

Lieberman, H., Smith, D., & Teeters, A. (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.

Mekler, E. D., Brühlmann, F., Opwis, K,. & Tuch, A. N. (2013). Disassembling gamification: the effects of points and meaning on user motivation and performance. *CHI EA '13: CHI '13 Extended Abstracts on Human Factors in Computing Systems* (pp. 1137-1142). https://doi.org/10.1145/2468356.2468559

Mihalcea, R. (2003). Open MindWord Expert: Creating large annotated data collections with web users help. In *Proceedings of the EACL 2003, Workshop on Linguistically Annotated Corpora (LINC 2003)*.

Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D., & Miller, K. (1990). Introduction to WordNet: an On-Line Lexical Database. *International Journal of Lexicography*, *3*(4), 235-244. https://doi.org/10.1093/ijl/3.4.235

Navigli, R., & Ponzetto, S. (2010). BabelNet: Building a very large multilingual semantic network. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics* (pp. 216-225). Uppsala, Sweden.

Paolizzo, F. (2019) *M-GWAP: An Online and Multimodal Game With A Purpose in WordPress for Mental States Annotation.* https://arxiv.org/abs/1905.12884v1

Ploux, S., & Victorri, B. (1998). Construction d'espaces sémantiques à l'aide de dictionnaires de synonymes. *Traitement Automatique des Langues*, *39*(1), 161-182. https://halshs.archives-ouvertes.fr/halshs-00009433

Sagot, B., & Fišer, D. (2008). Construction d'un wordnet libre du français à partir de ressources multilingues. *In Proceedings of the TALN 2008*. Avignon, France. https://hal.inria.fr/inria-00614707

Sagot, B., Fort, K., Adda, G., Mariani, J., & Lang, B. (2011). Un turc mécanique pour les ressources linguistiques : critique de la myriadisation du travail parcellisé. In *Proceedings of the TALN'2011 – Traitement Automatique des Langues Naturelles*. Montpellier, France. https://hal.inria.fr/inria-00617067

Siorpaes, K., & Hepp, M. (2008). Games with a Purpose for the Semantic Web. *IEEE Intelligent Systems*, *23*(3), 50-60. https://doi.org/10.1109/MIS.2008.45

Steyvers, M., Miller, B., Hemmer, P., & Lee, M. D. (2009). The wisdom of crowds in the recollection of order information. In *Proceedings of the NIPS'09: Proceedings of the 22nd International Conference on Neural Information Processing Systems* (pp. 1785-1793). NY: Curran Associates Inc.

Tchechmedjiev, A., Mandon, T., Lafourcade, M., Laurent, A., & Todorov, K. (2017). Ontolex JeuxDeMots and Its Alignment to the Linguistic Linked Open Data Cloud. *International Semantic Web Conference* (pp. 678-693). Vienna, Austria.

Thaler, S., Siorpaes, K., Simperl, E., & Hofer, C. (2011). A Survey on Games for Knowledge Acquisition. *STI Technical Report*.

Tisserant, G., & Lafourcade, M. (2015). PolitIt, du crowd-sourcing pour politiser le lexique. In *Étudier le Web politique : Regards croisés, Institut des Sciences de l'Homme*. Lyon.

Venhuizen, N., Basile, V., Evang, K., & Bos, J. (2013). Gamification for Word Sense Labeling. Proceedings of the 10th International Conference on Computational Semantics (IWCS 2013), pp. 397-403. https://www.aclweb.org/anthology/W13-0215/

Véronis, J., & Ide, N. (1990). Word Sense Disambiguation with Very Large Neural Networks Extracted from Machine Readable Dictionaries. In *Proceedings of 13th International Conference on Computational Linguistics (COLING'90)* vol. 2, (pp. 389-394). Helsinki.

Vossen, P. (1998). *EuroWordNet: A Multi-lingual Database with Lexical Semantic Networks.* Dordrecht, The Netherlands: Kluwer.

Zampa, V., & Lafourcade, M. (2010). PtiClic et PtiClic-kids : jeux avec les mots permettant une double acquisition. In *Proceedings TICE'2010, 7e colloque TICE*. Nancy.

Zesch, T. & Gurevych, I. (2009). Wisdom of crowds versus wisdom of linguists measuring the semantic relatedness of words. *Natural Language Engineering* (pp. 25-59). Cambridge University Press. https://doi.org/10.1017/S1351324909990167