



HAL
open science

PtiClic: jeu pour déterminer la lisibilité d'un terme et pour l'acquisition lexical automatique

Mahmoud Charfeddine

► **To cite this version:**

Mahmoud Charfeddine. PtiClic: jeu pour déterminer la lisibilité d'un terme et pour l'acquisition lexical automatique. Sciences de l'Homme et Société. 2013. dumas-00828541

HAL Id: dumas-00828541

<https://dumas.ccsd.cnrs.fr/dumas-00828541>

Submitted on 31 May 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Pticlic : **Jeu pour déterminer la** **lisibilité d'un terme et** **pour l'acquisition lexical** **automatique**

Nom : Charfeddine
Prénom : Mahmoud

Communication

Mémoire de master 2 recherche - 30 crédits – Mention Sciences du Langage

Spécialité ou Parcours : Industrie de la langue

Sous la direction de Virginie Zampa et Mathieu Lafourcade

Année universitaire 2012-2013



Pticlic :

Jeu pour déterminer la lisibilité d'un terme et pour l'acquisition lexical automatique

Nom : Charfeddine
Prénom : Mahmoud

UFR Langage, Lettres et Arts du spectacle, Information et
Communication

Mémoire de master 2 recherche - 30 crédits – Mention Sciences du Langage

Spécialité ou Parcours : Industrie de la langue

Sous la direction de Virginie Zampa et Mathieu Lafourcade

Année universitaire 2012-2013

Remerciements

Je sais que d'ordinaire les remerciements se font par ordre alphabétique, mais j'ai décidé de ne pas respecter cette tradition. Je tiens donc à remercier,

Virginie Zampa et Mathieu Lafourcade pour l'encadrement, les relectures, les conseils et m'avoir fait confiance en me proposant ce stage.

Ma famille pour l'encouragement,

Lulu et Milie, pour leurs patiences et pour leurs présences, leurs conseils et leurs corrections indispensables à la fin.

Régis et Marie pour les idées et les conseils,

Olivier Kraif pour ses remarques et ses conseils indispensables,

Issam, Nader, cecile et elo pour les idées et l'encouragement,

Riadh pour ses corrections en anglais,

Le DIP et le bureau des doctorants, pour l'accueil, le travail, etc.

Pour tous les gens qui ont participé aux questionnaire (Tristan, Manel, Gilles, etc.),

Et évidemment, toutes celles et ceux que j'oublie malgré leur aide.

Sommaire

| | |
|--|-----------|
| PARTIE 1 ETAT DE L'ART | 9 |
| CHAPITRE 1 – INTRODUCTION..... | 10 |
| CHAPITRE 2 – RESSOURCES LEXICALES ET RESEAUX LEXICAUX SEMANTIQUES | 12 |
| <i>Les ressources lexicales</i> | 12 |
| <i>Les réseaux lexico-sémantiques</i> | 12 |
| <i>Les relations sémantiques</i> | 14 |
| CHAPITRE 3 – CONSTRUCTION DE LA RESSOURCE..... | 19 |
| <i>Les outils d'acquisition automatique du lexique existants</i> | 19 |
| <i>Les outils d'acquisition avec intervention humaine par des spécialistes</i> | 20 |
| <i>Les outils d'acquisition lexicale contributifs</i> | 21 |
| CHAPITRE 4 – CONCLUSION | 27 |
| PARTIE 2 CONCEPTION DE LA PLATEFORME | 28 |
| CHAPITRE 5 – LES OBJECTIFS INITIAUX..... | 29 |
| <i>Pticlic</i> | 29 |
| <i>PticlicKids</i> | 30 |
| CHAPITRE 6 – REFONTE DE PTICLIC | 32 |
| <i>Nos choix</i> | 32 |
| <i>Etude conceptuelle</i> | 33 |
| <i>Aspects ergonomiques et sécurité</i> | 36 |
| <i>Public</i> | 37 |
| CHAPITRE 8 – PERSPECTIVES | 53 |
| <i>Questionnaire</i> | 53 |
| <i>Expérimentation à venir</i> | 54 |
| <i>Perspectives d'évolution pour PtiClic</i> | 54 |

Introduction

Dans le cadre de mon mémoire de master 2, j'ai eu l'occasion de faire un stage de recherche. Mon stage se déroule grâce à la collaboration de deux laboratoires. Le laboratoire du LIDILEM et celui du LIRMM.

Le LIDILEM (Linguistique et Didactique des Langues Etrangères et Maternelles) se trouve à Grenoble et se rattache à l'université Stendhal Grenoble 3. Il compte actuellement une cinquantaine de personnels permanents pour une centaine de doctorants. Leurs travaux se scindent en trois axes : le premier dédié aux descriptions linguistiques, TAL, corpus, le deuxième s'intéresse à la sociolinguistique et acquisition du langage et le dernier spécialisé dans la didactique des langues et les recherches en ingénierie éducative.

Virginie Zampa qui est maître de conférences m'encadre dans ce laboratoire. Elle fait partie de l'équipe TAL et est spécialisée dans les EIAH (Environnements Informatiques pour les Apprentissages Humains)

Le LIRMM, quant à lui, dépend de l'université de Montpellier 2. Il est l'acronyme du Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier. Les principaux axes de recherche de ce laboratoire sont : la conception et la vérification de systèmes intégrés, mobiles, communicants, la modélisation de systèmes complexes à base d'agents et les études en algorithmique, bioinformatique, interactions homme-machine, robotique, etc. Il se compose d'environ 350 personnes dont 160 permanents pour 150 doctorants et une quarantaine d'ingénieurs et techniciens.

Mon encadrant au sein de ce laboratoire est Mathieu Lafourcade. Maître de conférences habilité à diriger les recherches, il s'intègre dans l'équipe TAL et a notamment créé un jeu intitulé JeuxDeMots.

Mon thème de recherche se situe au niveau de plusieurs domaines. Le TAL, d'une part, et plus spécifiquement l'extraction des termes de jeux de mots, l'acquisition lexicale automatique dans le but d'assembler des données pour des réseaux lexicaux. Les EIAH (Environnements Informatiques pour les Apprentissages Humain) d'autre part, notamment la dimension liée à l'âge d'acquisition des termes chez les enfants. Mon projet est donc de recréer une plateforme baptisée PtiClic et d'y intégrer une autre plateforme conçue : PtiClicKids.

Partie 1

Etat de l'art

Chapitre 1 – Introduction

La performance des systèmes de traitement des langues est souvent corrélée avec la quantité et la qualité des ressources lexicales utilisées. Différentes méthodes d'acquisition lexicale s'opposent mais quoiqu'il en soit la construction de réseaux lexicaux reste souvent coûteuse et difficile. De plus, il existe un problème essentiel au niveau de l'extraction des termes. En effet, il est rare de parvenir à obtenir des données non bruitées et réellement représentatives des usages des locuteurs. Cependant, à l'heure actuelle, il existe plusieurs méthodes permettant d'acquérir des données lexicales par le biais de jeux qui permettent de résoudre certaines de ces difficultés.

Connaître l'âge d'acquisition des termes et des relations entre termes constituerait une donnée intéressante pour de nombreuses applications du TALN. Ceci permettrait par exemple d'évaluer la difficulté des textes proposés aux élèves. En effet, il n'est pas évident d'évaluer la difficulté lexicale d'un texte. [MESNAGER, 2011] estime que les calculs de lisibilités ne sont pas toujours pertinents car ils ne s'appuient pas forcément sur les bons critères. Plusieurs facteurs rentrent en compte le système anaphorique, les points de vue (les sens qui dépassent la phrase) et la difficulté de la langue (lexique).

Ainsi [MESNAGER, 2002] résume les aspects à prendre en compte pour mesurer la lisibilité d'un texte comme étant la typographie, l'aspect lexical, la syntaxe de la phrase et la cohérence. Dans un autre article [MESNAGER, 1989] prend en exemple les indices de Flesch. Ces derniers se fondent sur le principe statistique : plus un mot est rare, plus il est long, moins il sera compris par l'enfant. De même, plus une phrase est longue, plus sa syntaxe est complexe, moins l'enfant accèdera au contenu sémantique du texte. Ces indices ne sont pas tout à fait exacts. En effet, des mots tels que *tyrannosaure* ou *diplodocus* sont connus par un enfant de dix ans. Ce n'est donc pas la longueur des mots ou la rareté qui va déterminer la complexité. De la même façon, une phrase longue peut être aussi lisible qu'une phrase courte. Nous avons donc choisi de recourir à un jeu afin de connaître l'âge d'acquisition d'un terme par la détermination des relations sémantiques qu'entretiennent les termes.

Dans le cadre de mon projet de mémoire, j'ai choisi de travailler sur l'acquisition lexicale automatique dans le but de collecter des données pertinentes pour des réseaux lexicaux. En effet, les relations entre les mots constituent un enjeu primordial en Traitement Automatique des Langues Naturelles (TALN). Durant cette période, je suis

chargé de l'amélioration d'une plateforme conçue qui comporte deux jeux : PtiClic, pour tous les utilisateurs et PtiClicKids s'adressant à un public d'enfants et préadolescent. Tous deux s'appuyaient sur deux méthodes d'acquisitions lexicales JeuxDeMots¹ (JDM) [LAFOURCADE, 2007] et l'Analyse Sémantique Latente (LSA) [LANDAUER & DUMAIS 1997].

Le présent état de l'art se propose de dresser un rapide tour d'horizon des différents domaines nécessaires à la refonte du logiciel PtiClic. Il se compose au total de quatre chapitres. Dans un premier temps nous aborderons les ressources lexicales et plus particulièrement les réseaux lexicaux sémantiques. Nous verrons ensuite dans le chapitre 3 comment se construisent les différents outils s'appuyant sur les réseaux lexicaux. Enfin nous conclurons cet état de l'art en résumant les choix que nous faisons pour la nouvelle version de PtiClic.

¹ <http://www.jeuxdemots.org/jdm-accueil.php>

Chapitre 2 – Ressources lexicales et réseaux lexicaux sémantiques

Afin de pouvoir déterminer l'âge d'acquisition des mots et des différents types de relations entre termes, nous avons besoin de nous fonder sur des ressources lexicales. Nous allons donc ici présenter ce que sont les ressources lexicales ainsi que les réseaux lexicaux sémantiques.

Les ressources lexicales

[SAADANI & BERTRAND-GASTALDY 2000] s'appuient sur la norme internationale ISO 2788 pour définir les thésaurus comme étant un « *vocabulaire d'un langage d'indexation contrôlé organisé formellement de façon à expliciter les relations a priori entre les notions (par exemple relation générique-spécifique)* ». De nouveaux types de thésaurus ont été mis en évidence après l'apparition des traitements statistiques et linguistiques des langues naturelles et la mise au point de systèmes d'information intelligents. C'est le cas notamment pour les thésaurus de recherche.

L'ontologie quant à elle pourrait être considérée comme l'expression d'une vision du monde décrite sous la forme d'un réseau de concepts. Selon [ZACKLAD, 2007] « *le terme d'ontologie est aujourd'hui utilisé de manière bien plus large et imprécise pour désigner toute classification aisément partageable sur le web grâce à l'usage des standards du W3C que sont le langage XML et le protocole http, sans qu'il soit nécessaire de faire référence au web sémantique formel.* » (p. 8)

La méthode LSA [LANDAUER et DUMAIS 1997] fait également partie des méthodes permettant de mettre en place des réseaux lexicaux avec juste un seul type de relation « a un rapport avec ». Cette dernière méthode sera davantage présentée dans les prochains chapitres. Il est important de noter qu'elle constitue une des deux méthodes d'acquisition lexicale employées dans la première version du jeu PtiClic.

Les réseaux lexico-sémantiques

La construction des réseaux lexicaux constitue la première phase de création d'un outil d'acquisition lexicale. Nous allons donc présenter dans un premier temps, ce qu'est un réseau lexical, en nous focalisant sur deux exemples, avant de s'intéresser aux différentes utilisations qui en sont faites.

Définition

Il s'agit d'un réseau (au sens de graphe) liant des termes entre eux via des relations de différents types. Ces relations peuvent être de nature lexicale, sémantique, ontologiques, etc.

Dans certains réseaux lexicaux (par exemple celui de JDM) les relations en plus d'être typées sont pondérées.. La figure 1 permet de d'illustrer la définition que nous venons d'établir.

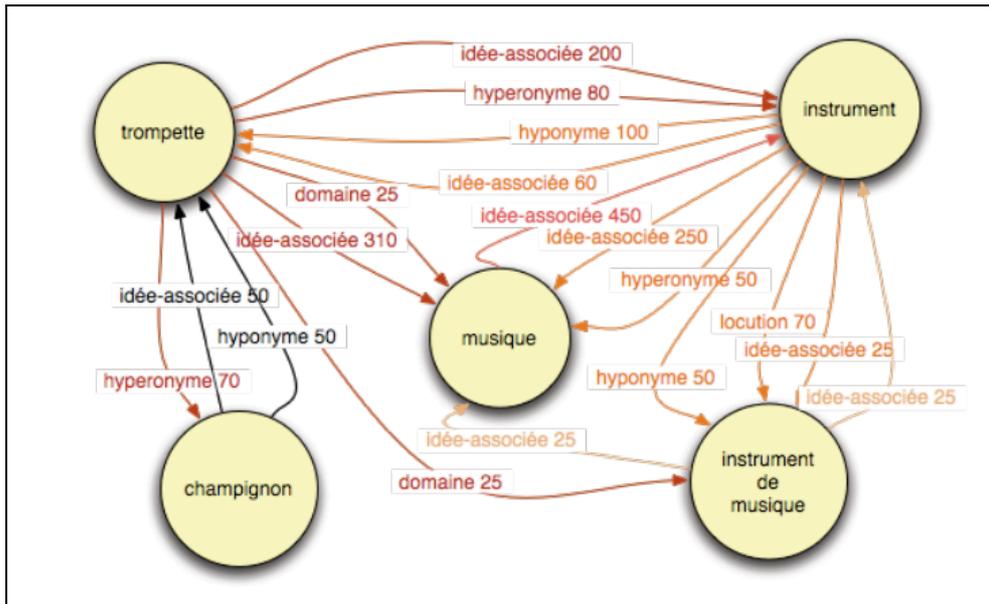


Figure 1 Schéma de description d'un réseau lexical [Zampa & Lafourcade 2011, p.7]

Exemples de réseaux lexicaux

Il existe différentes manières de construire un réseau lexical. Le réseau de JDM, est défini par [LAFURCADE, JOUBERT & RIOU 2009] comme suit « Ces relations sont typées par la consigne imposée aux joueurs ; elles sont pondérées en fonction du nombre de paires de joueurs qui les ont proposées. La structure du réseau lexical que nous cherchons ainsi à obtenir s'appuie sur les notions de nœuds et de relations entre nœuds...Chaque nœud du réseau est constitué d'une unité lexicale (terme ou expression) regroupant toutes ses lexies et les relations entre nœuds traduisent des fonctions lexicales » (p 3). L'ensemble des relations est présenté dans la partie sur les relations sémantiques à la page 9.

[MILLER, BECKWITH & FELLBAUM 1990] proposent grâce à Wordnet de diviser le lexique en cinq catégories : noms, verbes, adjectifs, adverbes, et mots grammaticaux. Ils définissent ensuite pour presque chaque mot un synset (sens du mot).

Ceci leur permet ainsi de savoir quelles sont les relations entre les différents termes à travers les synsets précédemment définis. Les synsets peuvent être considérés comme des concepts et ne sont pas ambigus, dans le cas de wordnet. Les synsets peuvent être considérés comme étant les nœuds. Les relations entre les synsets sont les arcs, Wordnet est structuré par des relations entre synsets et entre les mots. Ceci permet ainsi de retrouver des relations utilisées de type Synonymie, Antonymie, Hyponymie, Méronymie, etc. WordNet est un système d'héritage lexical. La structure de sa base de données prend la forme d'un arbre. Chaque entrée dans le système est constituée d'une référence ou d'un pointeur. Le pointeur étiquette l'hyperonyme par l'arbitraire symbole « @ » et les hyponymes par « ~ » afin de les connecter les uns aux autres. Toutefois, le réseau lexical élaboré par WordNet connaît quelques limites. Ainsi, il est impossible d'identifier un vocabulaire spécifique comme la classification des termes se fait selon des hiérarchies multiples correspondant à des champs sémantiques relativement distincts. Chacun possède son propre vocabulaire. D'autres limites se retrouvent également au niveau des informations sur les savoirs syntagmatiques ; c'est-à-dire qu'il n'existe pas d'informations sur le contexte des mots. Enfin, WordNet ne contient pas d'informations supplémentaires sur les exceptions.

Ainsi, nous avons pu voir qu'il existe à l'heure actuelle différentes méthodes permettant de visualiser les réseaux lexicaux. L'ensemble de ces réseaux nous permet maintenant d'aborder la notion de relation sémantique. Nous verrons tout d'abord définir cette notion qui est importante pour nous avant de nous intéresser aux différentes théories liées à cette notion.

Les relations sémantiques

Définition

Les relations sémantiques constituent une notion centrale de l'acquisition lexicale. Il est important de noter qu'il existe différents types de relation. Elles s'appuient toutes sur les liens de signification entre les mots ou les catégories d'un langage. Toutefois, elles peuvent se diviser en trois catégories distinctes. Les relations sémantiques les plus notables sont les relations hiérarchiques (de l'hyponyme à l'hyperonyme), les relations d'opposition (antonymie) et les relations d'équivalence (synonymie).

Les relations sémantiques de Wordnet

[MILLER, BECKWITH & FELLBAUM 1990] proposent par l'intermédiaire de WordNet un réseau lexical organisé par des relations sémantiques. Sa caractéristique la plus ambitieuse est d'organiser l'information lexicale en termes de signification des mots (les synsets).

Les relations entre les synsets sous WordNet se font grâce aux relations sémantiques. Nous pouvons noter à titre d'exemple « *un mot W1 'IS-A' mot W2* », il y a toujours des relations inverses comme félin hyperonyme de chat et chat l'hyponyme de félin. En effet, le sens sous WordNet est représenté par des synsets, qui, comme nous l'avons abordé dans le chapitre 2 représentent un ensemble de mots quasi-synonymes considéré comme un sens particulier. Pour identifier les noms ils définissent la règle suivante : la relation est symétrique, c'est-à-dire si un terme M1 est proche de terme M2, alors M2 est également similaire à M1.

Wordnet considère comme synonyme la possibilité pour deux expressions de se substituer l'une l'autre dans un même contexte sans en changer le sens. De plus, WordNet est un système d'héritage lexical se fondant sur le principe de connecter tous les hyponymes avec leurs hyperonymes et inversement. Il suppose également que la distinction entre la synonymie et hyponymie peut toujours être faite. Toutefois, dans la pratique, bien sûr, cette distinction n'est pas toujours claire. Lorsque les trois types de relations sémantiques hyponymie, méronymie, et l'antonymie, sont incluses, cela permet d'obtenir un réseau fortement interconnecté des noms. En d'autres termes, cela signifie que l'hyponymie permet d'obtenir une classification ou une hiérarchie ; la méronymie permet de faire le lien entre les termes de même niveau ; et l'antonymie permet de faire les liens de contraire entre les termes. Ils ajoutent que les mots de différentes catégories syntaxiques ne peuvent pas être synonymes, et ne peuvent donc pas former des synsets ; en particulier parce qu'ils ne sont pas interchangeables. Ils en ont donc déduit qu'il est nécessaire de partitionner WordNet en catégories de noms, verbes, adjectifs, et adverbes.

Ainsi, ce réseau lexical construit la séparation des noms grâce à la création de « *unique beginners* » correspondant à un nombre relativement limité de concepts génériques. Un « *unique beginner* » peut donc être considéré comme étant le composant sémantique primitif de tous les mots dépendant de sa hiérarchie. Les arbres ainsi constitués se font grâce à l'héritage sémantique de chacun des « *uniques beginners* ».

L'organisation sémantique des adjectifs est quant à elle organisée en groupes bipolaires. Les auteurs prennent à ce sujet l'exemple suivant : « *Le cluster pour humide / sec, qui définit l'attribut HUMIDITÉ, illustre les appareils de base de codage utilisés, et montre la variété et la diversité des sens qui peuvent être représentés au sein d'un cluster.* » (p.37)².

Ils ajoutent également que les relations sémantiques utilisées pour construire des réseaux de noms et des adjectifs ne peuvent pas être appliquées sans modification, mais doivent être adaptés à la sémantique des verbes.

Concernant les verbes, les auteurs estiment qu'ils sont la catégorie lexicale la plus difficile à étudier. En effet, la sémantique du verbe est beaucoup plus complexe. De plus, les verbes représentent la catégorie lexicale et syntaxique la plus importante de la langue. Les relations entre les verbes diffèrent des noms et des adjectifs.

Ainsi, l'organisation des verbes se fait selon le type des verbes, la citation suivante nous éclaire à ce sujet : « *les verbes d'état et les verbes de changement présentent une structure entièrement différente : ils ont tendance à être organisé en termes d'opposition et les relations de synonymie et ils peuvent être réparties entre les groupes bipolaires qui caractérisent les adjectifs.* » (p.57)

Les relations sémantiques de JeuxDeMots

JDM construit ses relations sémantiques en se reposant sur le fait que les joueurs peuvent proposer des termes qui seront validés ou non selon les concordances entre paires de joueurs. Les relations entre mots sont typées, orientées et pondérées. Les relations proposées aux joueurs sont les suivantes :

² Traduit de l'anglais : « *The cluster for wet/dry, which define the attribute WETNESS or MOISTNESS, illustrates the basic coding devices used, and shows the variety and range of senses that can be represented within a cluster.* »

- idées associées
- domain : un domaine auquel peut appartenir le terme
- domain-1 : donner des termes du domaine
- synonyme
- hyperonyme
- antonyme
- hyponyme
- partie de
- le tout
- locution (hallebarde -> pleuvoir des hallebardes)
- raffinement : ex : blaireau/animal /barbier..
- agent : qui peut faire telle action ?
- agent-1 : que peut faire cet agent ?
- patient : qui peut subir cette action ?
- patient-1 : que peut subir ce patient
- instrument (on peut effectuer l'action du verbe avec quel instrument)
- instrument-1 : que peut on faire avec cet instrument
- lieu : où peut se trouver le terme / où peut se situer l'action
- lieu-1 : ce qui se trouve dans ce lieu
- caractéristique type du mot
- caractéristique-1 : relation inverse de caractéristique (la caractéristique est donnée il faut trouver les termes ayant cette caractéristique)
- magne (plus intense que)
- anti-magn (moins intense que)
- famille (ex : chat chatte, chaterie, chatière, etc.)
- lieu_action : dans quel lieu peut on faire l'action X
- action_lieu : quelles actions peuvent être fait dans le lieu X
- sentiment : quel sentiment est évoqué par ce terme
- manière : de quelle manière (pour un verbe)
- sens : quel est le sens de ...

Cette méthode s'appuie sur des relations de type ontologique (hyperonyme, hyponyme, partie/tout) et sur quelques relations lexicales (synonymes, contraires, termes de la même famille, etc.). Il prend ainsi en compte les mots composés et les expressions dans leur intégralité c'est-à-dire qu'il ne traite pas chaque élément de l'expression comme des mots isolés. Le principe étant de se reposer sur les joueurs qui sont libres de choisir les mots. Lors d'une partie, les mots très récurrents dans le réseau ne permettent pas de gagner de nombreux points. Ceci incite donc les joueurs à trouver d'autres termes. Cela a pour conséquence d'avoir un avantage double pour le réseau de JDM. Il permet d'une part d'enrichir le réseau et d'autre part d'obtenir davantage de relations sémantiques entre les mots. Le filtrage des réponses se fait par la comparaison des résultats de deux joueurs.

Cette technique permet quant à elle de récupérer des relations entre les termes avec le moins de bruit possible ; cela permet donc d'obtenir sur le long terme des relations jugées pertinentes. Toutefois, JDM n'obtient pas de très bons résultats dans des domaines précis. Ceci est en partie dû au fait que les réponses des joueurs sont spontanées. De plus, comme ils ne sont pas experts du domaine et que le temps de jeu pour chaque partie est limité les réponses sont d'autant moins précises.

Maintenant que nous avons présenté les différents types de ressources lexicales dont nous pouvons avoir besoin dans PtiClic, nous allons nous intéresser à la manière dont ces ressources sont construites. Là aussi il existe différentes approches.

Chapitre 3 – Construction de la ressource

Comme nous l'avons déjà indiqué, la performance des systèmes de traitement des langues est souvent corrélée avec la quantité de ressources lexicales utilisées, mais ce n'est pas le seul critère. La manière dont la ressource est construite joue aussi sur sa qualité. Ici aussi, différentes méthodes d'acquisition lexicale s'opposent. En effet, il peut s'agir de construction automatique versus manuelle, par des spécialistes ou non, etc. Nous allons donc les étudier en regardant sur quelles ressources se fondent les différentes méthodes et quelles sont les méthodes qui nous semblent envisageables pour PtiClic.

Les outils d'acquisition automatique du lexique existants

Les premières méthodes d'acquisition auxquelles nous nous intéresserons se fondent sur les corpus.

La méthode polysémique

Cette méthode d'acquisition permet d'extraire les relations à partir d'un corpus existant. Toutefois, plusieurs procédés existent [LEON & MILLON 2005] par exemple construisent leur acquisition de relations lexicales via un corpus de pages web lemmatisé et étiqueté morpho-syntaxiquement. Pour ce faire ils décident de créer un corpus de pages web à partir de la combinatoire lexicale de dix noms français très polysémiques. Le fait d'utiliser des mots très polysémiques permet un accroissement important des relations lexicales contenues dans leur base de données. Ce corpus leur permet de définir les relations qui les intéressent selon des filtres linguistiques (expressions régulières) de leur choix. Afin de pouvoir extraire les relations lexicales qu'ils jugent pertinentes de type NOM ADJECTIF, NOM1 DE NOM2 et VERBE NOM, les auteurs ont défini un seuil de fréquence au moins égal à dix. Cette méthode ne leur permet toutefois pas de récupérer certaines relations caractéristiques. Il est important de noter également que les relations ne sont ni typées ni orientées ni pondérées.

LexSchem

De leur côté, [MESSIANT, NAKAMURA & VOYATZI 2009] construisent leur acquisition à partir de « gros corpus » pour le français par exemple le corpus LM10 qui correspond à dix années du journal *Le Monde*. Le système prend en entrée un corpus traité par un analyseur syntaxique, il filtre ensuite les erreurs par des méthodes statistiques et

fournit un lexique de sous-catégorisation verbale. Ils expliquent enfin que pour chacune des entrées, le lexique fournit son nombre d'occurrences en corpus, sa fréquence relative ainsi que cinq exemples tirés du corpus.

LSA

L'Analyse Sémantique Latente (LSA) [LANDAUER & DUMAIS 1997] prend en entrée des textes à partir desquels elle compose une matrice d'occurrences. Cette méthode construit ses relations sémantiques en s'appuyant sur des méthodes et des analyses statistiques. [LAFOURCADE, JOUBERT & RIOU 2009] définissent cette méthode comme permettant de « *calculer une proximité sémantique entre mots et ainsi produire des nuages de termes appartenant à un même champ sémantique* » (p.2). Elle représente les mots de la langue dans un espace multidimensionnel où chaque terme est défini par un vecteur. Chaque mot est représenté par un vecteur dans un espace à trois cents dimensions et généralement la proximité entre deux mots est fournie par le cosinus de l'angle que forment les deux vecteurs.

Cette méthode nous permet d'obtenir les relations sémantiques entre les termes voire même entre les textes. Toutefois, même si la LSA peut donner des résultats intéressants, reste trop fortement liées aux textes utilisés. De plus, la « proximité sémantique » fournie n'est ni typée, ni orientée.

Les outils d'acquisition avec intervention humaine par des spécialistes

WordNet

[MILLER, BECKWITH & FELLBAUM 1990] présentent leurs travaux sur un outil d'acquisition manuelle, Wordnet. Ce réseau lexical est une des bases de données lexicales les plus renommées pour la langue anglaise. Cette approche manuelle s'intéresse à la structure du contenu sémantique et lexical de la langue. Leurs travaux de recherche, débutés en 1985, leur ont permis de travailler à leur actuelle sur la troisième version de l'outil. Cette version 3.0 de WordNet contient 155 287 mots avec 117 798 noms, 11 529 verbes, 21 479 adjectifs et 4 481 adverbes organisés en 117 659 synsets³. Le principal inconvénient de ce type d'approche est qu'elle reste manuelle. Cela signifie donc qu'elle est coûteuse en temps et en main d'œuvre.

³ La notion de synset a déjà été présentée dans la section intitulée « relations sémantiques » du chapitre 3

Autres ressources lexicales

De leur côté [MESSIANT, NAKAMURA & VOYATZI 2009] pensent qu'il existe plusieurs ressources lexicales syntaxiques, comme les thésaurus et les ontologies qui sont réalisées manuellement. Ces lexiques visent à déterminer les emplois de chaque lemme verbale, ainsi que leurs cadres de sous catégorisation, afin de définir le nombre et le type de ces arguments. Nous pouvons citer à titre d'exemple le dictionnaire syntaxique des verbes français créé par deux linguistes, Dubois & Dubois-Charlier en 1997. Ils s'appuient sur une classification sémantico syntaxique des verbes, ces entrées sont représentées sous la forme d'un couple verbe et schéma de sous catégorisation. [MESSIANT, NAKAMURA & VOYATZI 2009] parlent également de DicoValence qui est un dictionnaire syntaxique manuel construit dans le cadre méthodologique de l'approche pronominale réalisée par [van den Eynde & Blanche-Benveniste, 1978], son principe se fonde sur la relation entre les dépendants du verbe et leurs pronoms pour trouver les valeurs des prédicats.

Les outils d'acquisition lexicale contributifs

Un autre type de méthode d'acquisition lexicale existe. Contrairement aux méthodes que nous venons de présenter, elles prennent pour fondement les contributions volontaires ou issues de jeux. Nous présenterons dans un premier temps, l'Amazon Mechanical Turk (AMT) qui est un outil d'acquisition fondé sur une rémunération à la tâche. Nous nous intéresserons ensuite aux contributions grâce aux jeux. Le premier jeu que nous allons présenter est *The ESP game*. Nous nous intéresserons ensuite à Verbosity avant de finir avec JeuxDeMots.

Volontaire

Amazon Mechanical Turk est une plateforme en ligne conçue par l'entreprise Amazon en 2005 pour mettre en relation des travailleurs (Turkers) avec des gens qui proposent du travail parcellisé c'est-à-dire la segmentation d'un travail en petites tâches. Les fournisseurs de tâches (requesters) divisent un projet en petites tâches simples et proposent un prix pour chaque élément fait. Les Turkers complètent les tâches qu'ils veulent faire et gagnent donc de l'argent. [SAGOT et al 2011] proposent une analyse critique de ce type de plateforme que nous allons reprendre. Ils notent d'une part qu'il y a très peu de financement dans ce domaine de recherche. Les auteurs prétendent que ce type de travail induit quatre problèmes principaux :

- le non-respect du droit du travail avec le système du travail à la tâche qui ne permet pas d'avoir une idée précise du temps à passer sur la tâche
- les coûts cachés dans l'utilisation de la plateforme (développement de l'interface, garde-fous, compensation de la mauvaise qualité des résultats, etc.)
- la qualité du travail fourni nécessite souvent une vérification ultérieure
- la propriété intellectuelle du travail fourni n'est pas bien définie.

Comme le suggèrent les auteurs, il serait intéressant d'utiliser des ressources existantes fournies par la communauté d'experts en TAL, de favoriser le travail en collaboration des experts, mais également le développement de jeux ayant un but pour former les non-experts.

C'est ce dernier point qui nous intéresse plus particulièrement. Les derniers types d'outils que nous allons présenter maintenant sont les outils d'acquisition de type contributif par le jeu. Nous verrons donc trois jeux emblématiques du domaine.

Jeux

The ESP game

Cette première méthode d'acquisition a été mise en place par [VON AHN & DABBISH 2004] ; elle permet d'extraire des relations à partir des réponses communes à deux joueurs. Le principe du jeu est le suivant ; une image est commune aux deux joueurs ; dès qu'ils trouvent le même mot en lien avec cette image, ils passent à l'image suivante. Les deux joueurs doivent se mettre d'accord sur une image. Ce jeu a été conçu afin d'être joué en ligne par un grand nombre de paires à la fois. C'est un système interactif permettant l'étiquetage d'image dans le but d'acquérir des connaissances distribuées en utilisant des personnes connectées.

Le jeu se déroule de façon synchrone entre deux joueurs. Ils ne sont toutefois pas en mesure de communiquer. Des points sont gagnés lorsque les termes des deux joueurs correspondent. L'idée est que les termes fournis par les joueurs constituent des éléments d'indexation pertinents pour décrire l'image. Ne retenir que les mots clés communs fournit la base d'un filtrage minimal permettant d'éviter la majeure partie du bruit. A notre connaissance, ESP a été le premier jeu en ligne dont l'objet était de récupérer de telles données via l'activité des joueurs. D'autres jeux ont vu le jour, notamment concernant la

localisation d'objets dans une image [VON AHN et al., 2006a], ou encore la collecte d'informations de sens commun [LIEBERMAN et al., 2007] et [VON AHN et al., 2006b].

Avant d'attacher une étiquette à l'image, il faut la comparer à un seuil (nombre de paires en accord sur une étiquette). Les images marquées sont quant à elles réintroduites dans le jeu après que plusieurs mois se soient écoulés. En effet, le sens des images peut avoir changé. Les auteurs prennent comme exemple une image de Michael Jackson il y a vingt ans peut être étiquetés comme «superstar», alors qu'aujourd'hui il pourrait être étiquetés comme «criminel» (p. 3). La figure 2 présente une copie d'écran du jeu ESP.



Figure 2 : copie d'écran du jeu *The ESP Game* (Von Ahn & Dabbish, 2004, p.2)

Sur une période de quatre mois ils ont réussi à collecter 1 271 451 étiquettes pour 293 760 images différentes. Au total 13 630 personnes ont participé au jeu durant cette même période.

Il est important de noter que cette méthode ne peut pas être appliquée dans notre cas, notamment à cause des nombreux biais qui existent (triche possible, etc.).

Verbosity

Cette méthode d'acquisition a été réalisée par [VON AHN, KEDIA & BLUM, 2006] Le jeu a été conçu pour être joué en ligne à deux. Les deux joueurs sont choisis au hasard, l'un est défini comme le narrateur et l'autre comme le devineur. Le narrateur obtient un mot secret et doit donner des indices au devineur afin que ce dernier puisse le trouver. Les indices sont sous forme de modèles de phrase à trous accessibles via des cartes où chaque carte contient un modèle de phrase. Le narrateur peut remplir les champs avec tous les mots qu'il souhaite, mis à part le mot secret ou une chaîne contenant ce mot. Le

devineur peut donner son idée par rapport à l'indice du narrateur et celui-ci répondra à ses réponses par « hot » ou « cold ». L'acquisition des connaissances se fait à travers les indices donnés par le narrateur. Par exemple, pour le mot téléphone, lorsque le narrateur dit « il contient un écran » le système acquiert dans sa base de connaissance le fait que le téléphone possède un écran. Ainsi les relations acquises sont de type « l'objet d'un mot », « la hiérarchie de catégorisation », « fournir des données spatial », « homonymie » etc. Le type de modèle va déterminer quelle relation sera utilisée. Une partie dure six minutes pendant lesquelles les joueurs peuvent jouer sur autant de mots qu'ils le souhaitent. Ils ont le choix de passer un mot s'ils se mettent d'accord pour dire qu'il est trop difficile.

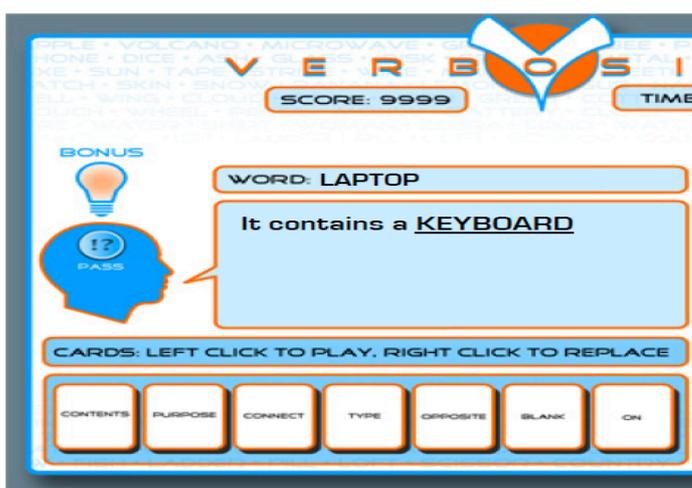


Figure 3 : copie d'écran du jeu VerboSai (Von Ahn, Kedia & Blum, p.2)

Au total 267 personnes ont joué dans une durée de 1 semaine, générant 7.871 faits. Cela signifie que, en moyenne, chaque joueur a contribué 29,47 faits. En termes de temps, chaque personne a joué pour une moyenne de 23,58 minutes en une seule séance, et certains ont joué pendant plus de 3 heures. Nous croyons que ces chiffres montrent la façon dont le jeu est agréable.

A priori cette méthode comporte des failles. Il y a en effet la possibilité de tricher. Par exemple dans le cas du narrateur, il est possible d'utiliser des langages SMS dans les champs des textes à trous pour donner le mot à deviner, nous ne pouvons donc pas utiliser cet outil ici.

JeuxdeMots

JeuxdeMots (JDM) est un jeu développé en 2007 par Lafourcade [LAFOURCADE, 2007], [LAFOURCADE & ZAMPA 2011]. Comme nous l'avons vu dans le chapitre 3, il permet la construction d'un réseau lexical évolutif à l'aide des joueurs ; soit une

construction progressive. Les joueurs peuvent proposer des termes qui seront validés ou non grâce à la concordance entre paires de joueurs. La figure 4 montre un exemple de partie sur JDM.



Figure 4 : Exemple d'une partie sur JDM

Les relations entre mots sont donc typées, orientées et pondérées. Nous avons ainsi pu voir qu'elles reposaient sur des relations de type ontologique (hyperonyme, hyponyme, partie/tout) et sur quelques relations lexicales (synonymes, contraires, termes de la même famille, etc.) ; ce dernier type de relation nous intéresse plus particulièrement car il constitue un des fondements sur lesquels s'appuie le jeu que nous développons, PtiClic. La figure 4 nous montre les résultats de chaque joueur ainsi que les réponses communes qu'ils ont pu faire sur une partie donnée.

The screenshot shows the 'Jeux de mots' website interface. At the top, there is a navigation menu with links: ACCUEIL, CLASSEMENT, EVÉNEMENTS, SOUK, PORTAIL, CONTACT, FORUM, OPTIONS, AIDE, DIKO. On the right, a user profile for 'niniefaitlesmots' is displayed, showing a level of 64, 172923 credits, 165534 honor, 97 blue gems, and 1000 purple gems. A small gift icon and a 'Niveau : 48' badge are also visible. The main content area features the word 'sécurisation' in large purple letters. Below it, two lists of related terms are shown:

- Réponses données par niniefaitlesmots :** sécurisation de l'emploi • e-sécurisation • gestion des risques • loi martiale • économie • population • approche essentialiste • théorie critique • opinion publique • culture • théorie des relations internationales • relations internationales • sécurité • périmètre de sécurité
- Réponses données par ratonlaveur :** sécu • protection • police • dispositif • canot de survie • sauvetage • badge • mot de passe • caméra de vidéosurveillance • vidéoprotection • vidéo • flicage • gilet de sauvetage • surveillance • accès • vidéosurveillance • accès sécurisé • dispositif de sécurité • sécurité • danger

At the bottom of the lists, the word 'sécurité' is highlighted in green. On the left side, there is a vertical logo for 'ES ASSOCIATION' and a black ink splat graphic with a white play button icon.

Figure 5 : Exemple des résultats communs à deux joueurs pour la même partie sur JDM

Depuis son lancement, les joueurs qui ont participé à JDM permettant ainsi de jouer 1 214 511 parties et par conséquent de collecter plus de 2 251 635 relations lexicales pour environ 252 956 termes.

L'ensemble de ce chapitre nous a permis d'observer quelles étaient les différentes méthodes d'acquisition lexicales utilisées à l'heure actuelle dans le domaine du TAL.

Chapitre 4 – Conclusion

Nous venons de voir qu'à l'heure actuelle nous disposons de différentes méthodes d'acquisition lexicale. Cependant, la construction des réseaux reste souvent coûteuse et difficile. Un autre point important est qu'il n'existe pas vraiment de données permettant de simuler l'acquisition et surtout de savoir ce qui est acquis tant en terme de réseau lexical qu'au niveau de la moyenne d'âge d'acquisition.

Si dans [ZAMPA & LAFOURCADE, 2011] l'utilisation conjointe des deux méthodes (LSA et JDM) semblait donner de meilleurs résultats plutôt que de choisir isolément une des deux approches. Il nous semble maintenant que l'accroissement qu'a connu la base de données de JDM est suffisant pour ne plus avoir recours à la méthode LSA. Il est important de noter que PtiClic s'enrichira également grâce aux contributions de ses propres joueurs et enseignants.

Le choix des méthodes d'acquisition lexicale a été effectué dans le but d'arriver à obtenir un jeu contenant un réseau lexical évolutif s'adaptant à la population qui y jouerait. En effet, s'appuyer à la fois sur JeuxDeMots et sur le jeu que nous développons, PtiClic, permet d'avoir une base constamment mise à jour.

Initialement, PtiClicKids était un jeu conçu pour les enfants âgés de six à quinze ans et visait à évaluer l'âge à partir duquel les mots et les relations entre termes, étaient acquises. L'ajout des données concernant l'âge permettrait ainsi de bien sélectionner les mots pour un âge donné (notamment pour les données à caractère pornographique). Le fait que l'enseignant ait la possibilité de créer les différentes parties permet de filtrer les mots du vocabulaire pouvant heurter les plus jeunes publics.

De plus, l'utilisation des données, pour le TALN, concernant l'âge dans les applications de génération des textes permet de générer des versions multiples d'un même texte avec un vocabulaire adapté à différentes tranches d'âge.

Ainsi, PtiClic constitue un jeu éducatif visant à permettre à un public d'enfant ou de préadolescent d'acquérir du vocabulaire ou d'évaluer ses performances lexicales. Le système permet notamment la collecte des différents âges moyens de maîtrise du vocabulaire. Ce type d'information est crucial dans de nombreuses applications du Traitement Automatique des Langues, en particulier la génération de texte en fonction de l'âge de lecteurs ciblés.

Partie 2

Conception de la plateforme

Chapitre 5 – Les objectifs initiaux

A l'origine, il y avait deux jeux distincts : PtiClic et PtiClicKids qui s'appuyaient sur deux méthodes d'acquisition lexicale automatique (JDM et LSA). Les deux méthodes, lors de la conception initiale étaient complémentaires. L'avantage de LSA étant les aspects automatique mais surtout, pour des termes scientifiques, LSA permettait d'introduire du vocabulaire de spécialité, plus précis ou moins usité et qui, de ce fait, était absent du réseau JDM. Quand à JDM, les joueurs ne sont pas forcément des experts des domaines traitées et de ce fait, fournissent des réponses « de base » qui n'apparaîtraient pas dans des livres mais sont importantes, de plus, et c'est l'avantage principal les relations sont typées, orientées et pondérées ce qui n'est pas le cas avec LSA. Mais avec le temps, le réseau JDM s'est suffisamment développé pour que l'utilisation de LSA ne soit plus nécessaire.

Le principal point commun aux deux jeux est qu'ils fonctionnent tous les deux en système fermé. En effet, contrairement à JDM, les joueurs ne peuvent pas proposer de nouveaux termes. Ils ont seulement la possibilité de choisir parmi la liste des termes affichés et qui est générée grâce au réseau lexico-sémantique de JDM et LSA. Ces deux méthodes d'acquisition lexicale sont ainsi présentées.

Les deux sections suivantes vont nous permettre de présenter les deux jeux, PtiClic et PtiClicKids tels qu'ils avaient été initialement conçus.

Pticlic

Selon [LAFOURCADE & ZAMPA 2011], ce jeu permet de construire un réseau lexical. Les connaissances sur le réseau lexical sont acquises par la machine. Il permet donc de compléter le réseau de JDM. La création d'une partie ne dépendant pas d'un utilisateur, tout se faisait automatiquement et de façon aléatoire ; le jeu s'adressant à tous (enfants comme adultes).

Concernant le déroulement du jeu, une partie se déroule en asynchrone entre deux joueurs. Ils ont les mêmes termes ainsi que les mêmes consignes. Il s'agit donc de prendre en compte les propositions communes des deux joueurs.

[LAFOURCADE, JOUBERT & RIOU 2009] décrivent le but initial de PtiClic comme étant la collecte d'une base de données grâce à la création de réseaux lexicaux qui permettent d'enrichir le réseau de JDM tout en aidant les utilisateurs à consolider les associations acquises grâce à JeuxDeMots. Il s'agit donc de tester les relations de JeuxDeMots en les proposant aux utilisateurs afin de renforcer les relations du réseau

lexical. Ce choix de conception permet le renforcement des relations produites par JeuxDeMots ainsi de rendre le réseau lexical plus dense et de diminuer le taux de bruit. Afin de diminuer le risque de silence [LAFOURCADE & ZAMPA 2009] ont ajouté au jeu l’outil LSA afin de trouver d’autres termes n’ayant pas été proposés par les joueurs de JDM.

Interface initiale

La figure 6 représente une partie type de l’ancienne version de PtiClic , un mot cible est proposé au joueur (ici « professeur »), ainsi qu’un nuage de mots (« prof », « vacances scolaires », etc.) et qu’un certain nombre de relations (« ... est une partie de professeur », « une caractéristique de professeur est ... », etc.). Le joueur glisse les mots dans les relations qu’il veut, chaque mot ne devant pas forcément être dans une relation.

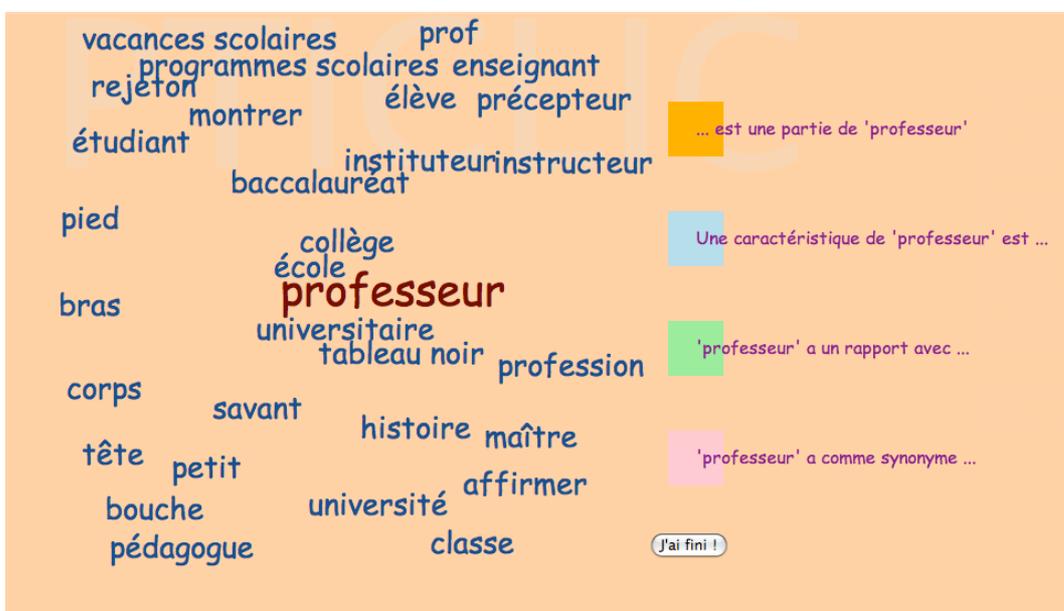


Figure 6 : Exemple de jeu avec l’ancien PtiClic

PticlicKids

Ce logiciel s’adressait essentiellement aux enfants, le jeu pouvant être fait avec un enseignant ou en autonomie [ZAMPA & LAFOURCADE, 2011]. Il s’agit de proposer un outil pédagogique qui permette de récupérer pour un terme donné d’un lexique, l’âge d’acquisition du terme. L’enseignant a tout de même la possibilité d’ajouter des relations s’il le souhaite. L’enseignant peut aussi indiquer l’âge des élèves auxquels la partie est adressée.

Le jeu fait une sélection de vingt mots à partir de chaque méthode JDM et LSA. C'est ensuite à l'enseignant de sélectionner les mots qu'il désire garder pour le jeu. Cela permet donc de calculer la maîtrise d'une relation entre deux mots pour un âge donné. Les auteurs donnent à ce sujet l'exemple suivant :

« Msyn,8 (chat, félin) = 0,5. exemple la maîtrise de la relation de type synonyme entre chat et félin pour un enfant de 8 ans est égale à 0.5 aussi on peut savoir la maîtrise d'une phrase exemple d'auteur dans le texte T : « les oiseaux gazouillent », si M8 (oiseau) = x et M8 (gazouiller) = y alors L8(T) = moyenne(x,y) »

Ainsi, comme nous l'avons déjà vu, contrairement à JDM il s'agit d'un jeu fermé où les joueurs ne peuvent pas proposer de nouveaux termes mais seulement choisir parmi la liste des termes affichés qui est généré par JDM. PtiClicKids un jeu conçu pour les enfants âgés de six à quinze ans et vise à évaluer l'âge à partir duquel les mots et les relations entre les termes étaient acquis. Le but est donc à terme de permettre la génération des textes avec des versions multiples d'un même texte dans un vocabulaire adapté à différentes tranches d'âge.

Dysfonctionnement

D'un point de vue ergonomique le nuage des mots n'était pas très adapté aux besoins des utilisateurs. Les mots se superposaient les uns sur les autres c'était donc à l'utilisateur de les séparer. Ce choix de conception avait été fait dans le but de divertir les joueurs. Il nous faudra donc trouver une autre technique plus adaptée.

Il faut bien viser la case de couleur correspondant à la relation sémantique pour mettre les mots et il n'est pas possible de ce qui a été mis dans chaque « case ».

Le jeu PtiClicKids n'ayant pas été implanté et PtiClic n'ayant pas beaucoup de joueurs, nous avons donc décidé de les fusionner. En effet, le nouveau jeu connaît quelques modifications fondamentales à plusieurs niveaux. D'un côté au niveau du système, mais également au niveau de l'interface et plus particulièrement des créations de parties, de l'ergonomie et de la façon de jouer.

Chapitre 6 – Refonte de PtiClic

Nos choix

Notre travail porte sur la refonte de PtiClic, logiciel conçu pour permettre des acquisitions lexicales par l'apprenant mais aussi par le système. Outre ces connaissances PtiClic, vise à obtenir l'âge moyen d'acquisition du lexique chez les enfants qu'il s'agisse du sens particulier d'un mot ou d'une expression ou des relations qu'ils entretiennent. Ces informations, inexistantes pour l'instant, sont importantes que ce soit pour les recherches en acquisition, en TAL en particulier la génération de texte en fonction de l'âge de lecteurs ciblés, en linguistique, etc.

Afin de mener à bien la conception et le développement de cet outil, nous avons besoin de nous fonder sur des ressources lexicales. C'est pourquoi nous avons analysé les différents types de ressources lexicales et leur méthode de construction. Le choix des méthodes d'acquisition lexicale a été effectué dans le but d'arriver à obtenir un jeu contenant un réseau lexical évolutif s'adaptant aux joueurs. Si dans [ZAMPA & LAFOURCADE, 2011] l'utilisation conjointe de deux méthodes (LSA et JDM) semblait donner de meilleurs résultats plutôt que de choisir isolément une des deux approches. Il nous semble maintenant que l'accroissement qu'a connu la base de données de JDM est suffisant pour ne plus avoir recours à la méthode LSA. De plus, le principal problème rencontré avec la LSA était que les connaissances restaient liées au corpus de base et que les relations ne sont ni typées ni orientées. En effet, s'appuyer à la fois sur JeuxDeMots et sur le jeu que nous développons, PtiClic, permet d'avoir une base constamment mise à jour. Il est important de noter que PtiClic s'enrichira également grâce aux contributions de ses propres joueurs et enseignants.

Maintenant que le choix du réseau lexical a été fait, il nous reste à implémenter la nouvelle version du jeu. Nous présentons par la suite les choix au niveau de la conception informatique, avant de nous intéresser aux choix ergonomiques et sécuritaires. Le fonctionnement du jeu, côté joueur, côté enseignant et côté système représente la dernière section que nous aborderons.

Etude conceptuelle

La conception représente une étape importante du processus de réalisation d'un projet informatique. Cette partie s'intéresse essentiellement à l'interaction du système avec son environnement, elle permet de faciliter le travail de certaines tâches et ce en décomposant le système en des sous-systèmes afin de cerner les difficultés que nous pourrions rencontrer.

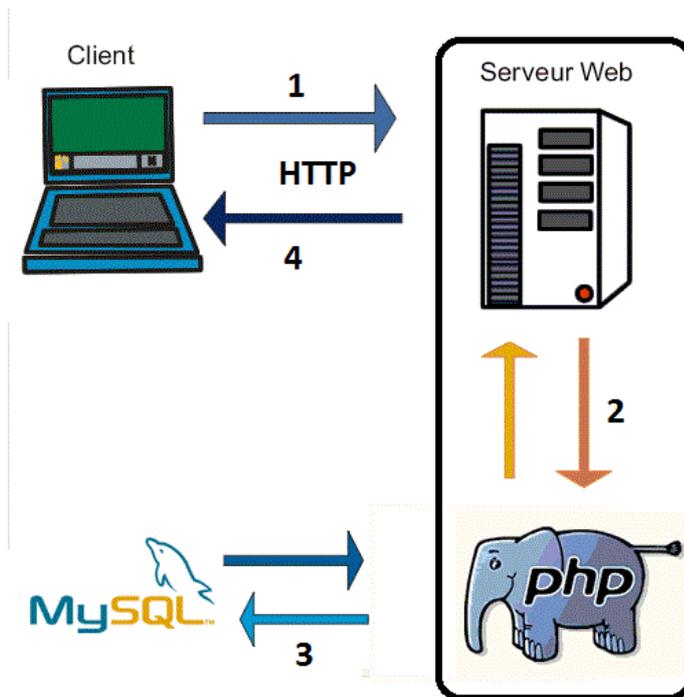


Figure 7 : Structure de l'architecture trois tiers⁴

Notre application se fonde sur une architecture de type 3-tiers (architecture à trois niveaux) qui est un modèle qui permet de séparer les trois niveaux (ou couches) en les interrogeant. Elle se compose de la partie client (utilisateur), serveur d'application et serveur de base de données.

Le client consulte l'interface du site pour demander un service. Des requêtes sont envoyées au serveur d'application qui les traite. Ensuite, il interroge le serveur de la base de données. Ce dernier extrait les données demandées qui seront renvoyées par la suite au client passant par le serveur d'application.

⁴ L'image est extraite à partir du site Php Débutant disponible à l'adresse suivante <http://www.phpdebutant.org/article118.php> dernière consultation le 18 avril 2013

Outils de conception

Présentation

UML (Unified Modeling Language) est un langage de modélisation unifiée permettant de représenter les composants d'un système informatique. Ce type de modélisation se compose habituellement de treize diagrammes différents. Nous nous intéresserons plus particulièrement à trois d'entre eux qui rentrent en compte dans le processus de conception d'une plateforme telle que PtiClic : les diagrammes de cas d'utilisation, de classe et de séquence. Les deux premiers permettent d'avoir une vue statique de PtiClic tandis que le troisième offre une vue dynamique de l'application.

Vue statique

Diagrammes de cas d'utilisation

Ce type de diagramme nous permet d'obtenir un aperçu général du fonctionnement du système que nous mettons en place. Il comporte deux entités les acteurs et les cas d'utilisation. Est considéré comme acteur l'utilisateur qui va interagir sur la plateforme. Dans notre cas ils sont représentés par les inscrits et les personnes particulières (visiteurs). Les fonctionnalités changent d'un type d'utilisateur à l'autre. La figure 8 illustre notamment le diagramme pour le visiteur. Les autres diagrammes de ce type sont présentés dans l'annexe 1.

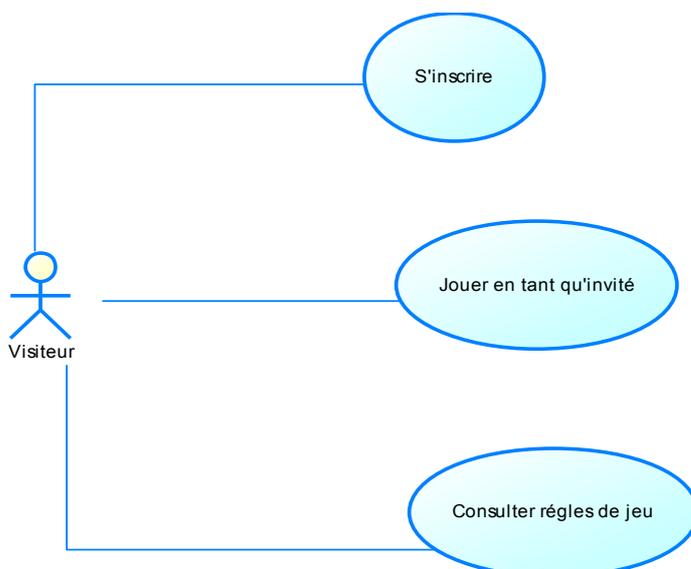


Figure 8 : Diagramme de cas d'utilisation pour le visiteur

Diagrammes de classes

Ces diagrammes comportent des classes et des associations. Une classe représente une liste d'attributs et de méthodes tandis que les associations relient les classes les unes aux autres. Le tout nous donnant une vue objet de l'application. La figure 9 présente le visiteur. Les diagrammes décrivant les autres utilisateurs sont présentés dans l'annexe 2 (figure 38 et 39).

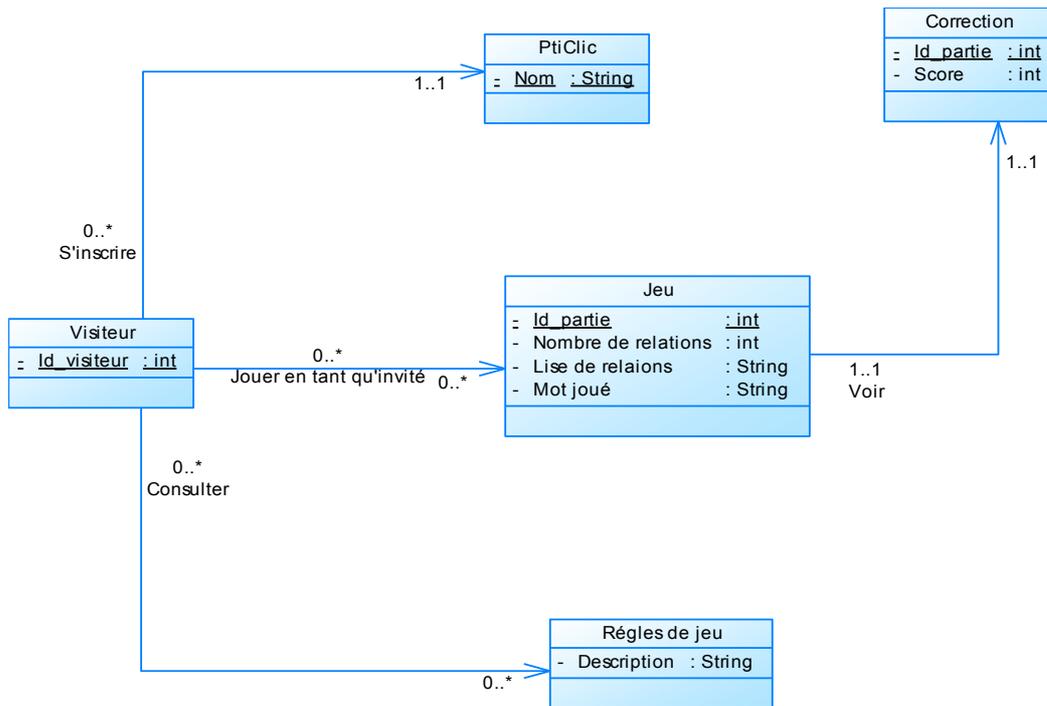


Figure 9 : Diagramme de classes pour le visiteur

Vue dynamique

Diagrammes de séquence

Le diagramme de séquence comporte des acteurs et des objets. Il a pour but d'expliquer les interactions des acteurs avec le système tout en fournissant le déroulement du scénario. Si le visiteur souhaite s'inscrire sur la plateforme, le diagramme de séquence qui représente cette démarche est présenté dans la figure 10. Un autre diagramme de séquence présentant le déroulement de la création de partie pour l'enseignant est contenu dans l'annexe 3.

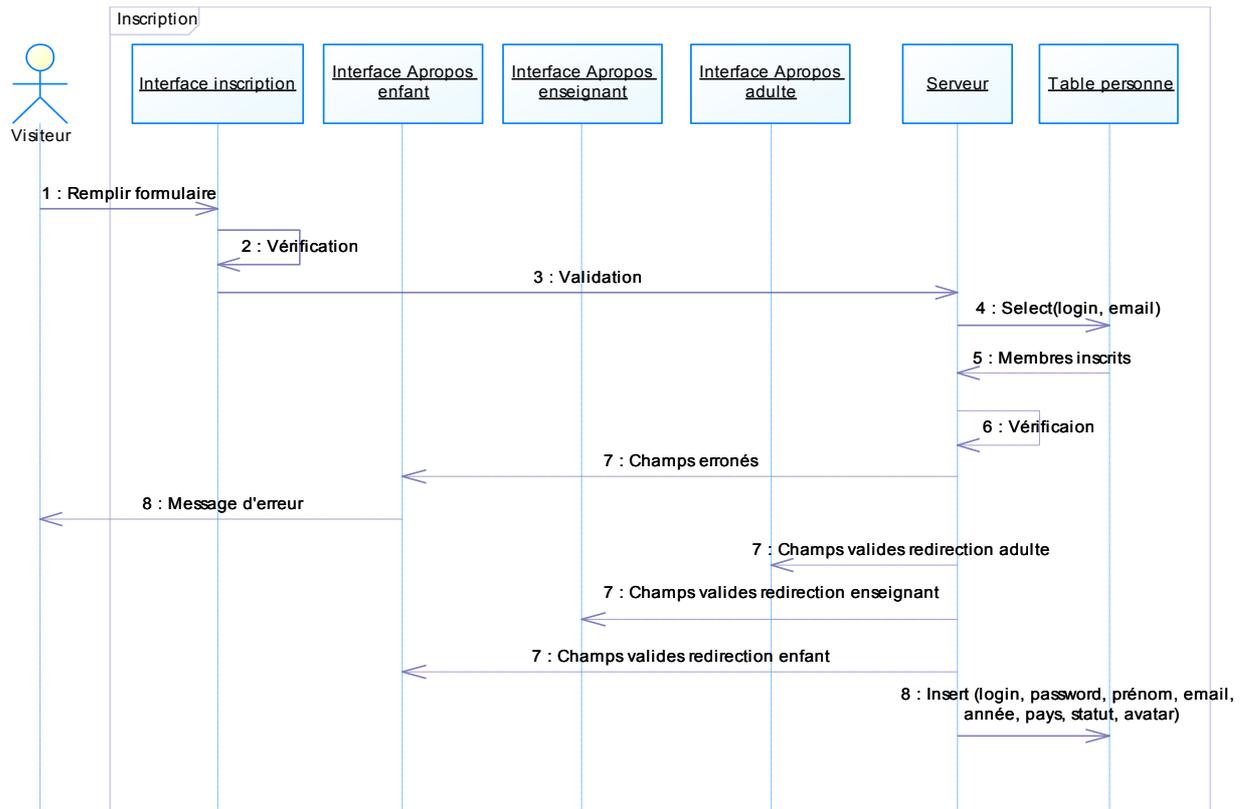


Figure 10 : Diagramme de séquence pour l'inscription

Aspects ergonomiques et sécurité

Cette partie présente tout d'abord les choix graphiques que nous avons faits, puis le détail des différents types d'utilisateurs et enfin le mode d'inscription à la plateforme.

Charte graphique

Police

Par souci de lisibilité et de clarté, des points d'autant plus importants qu'il s'agit d'un jeu, nous avons choisi d'utiliser la police Trebuchet MS pour le contenu et une police de type Arial pour les titres dans l'application.

Couleurs

Nous avons choisi des couleurs claires et non-agressives pour la plateforme. En effet, les utilisateurs peuvent être amenés à passer un certain temps sur celle-ci ; il était donc préférable de choisir des couleurs les plus neutres possible. Nous avons ainsi sélectionné la couleur blanche pour l'arrière plan du contenu par défaut et les utilisateurs

ont également trois autres interfaces au choix. La figure 11 illustre les différentes interfaces possibles.

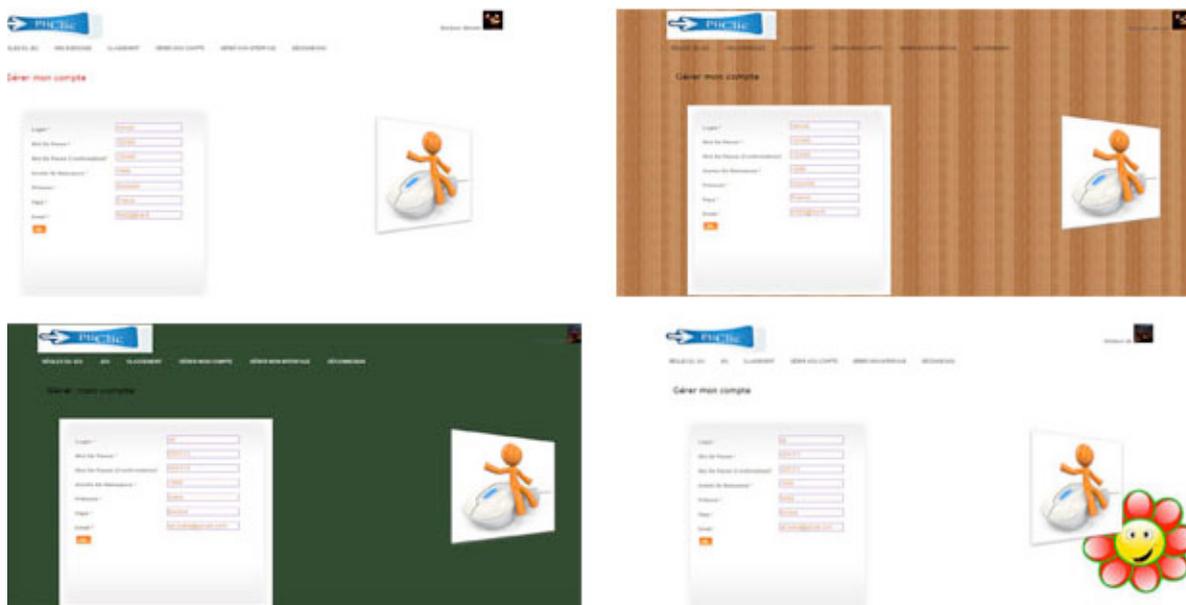


Figure 11 : Présentation des différentes interfaces

Sécurisation des données

L'authentification est nécessaire pour que l'utilisateur puisse accéder à ses informations. Toutefois, les informations contenues dans la plateforme (login, mot de passe, année de naissance, email, etc.) ne nécessitent pas d'avoir un niveau de sécurité très élevé.

Inscription

L'inscription sur la plateforme est libre. Les utilisateurs rentrent leurs informations en tenant compte de différents paramètres tels que l'âge et le statut (enfant, adulte, enseignant). Une fois que l'enseignant est inscrit sur PtiClic, il peut inscrire des élèves dans ses groupes ou les rajouter à partir de la liste existante.

Public

PtiClic est un jeu gratuit accessible en ligne qui s'adresse prioritairement à un public d'enfants ou de préadolescents, que ce soit en autonomie ou au sein d'une séquence didactique, mais qui reste ouvert à tout type de public. Cette application présente un intérêt pour le joueur lambda, qui peut acquérir des connaissances lexicales, et pour un enseignant, qui peut créer des parties sur les points qu'il veut aborder avec ses élèves ou réutiliser des parties créées par des collègues. Le jeu présente également un intérêt pour le

système qui va être complété et mis à jour par les résultats des parties, pour les chercheurs qui travaillent dans le domaine de l'acquisition, etc.

Les utilisateurs ont accès à des fonctionnalités différentes selon s'ils sont inscrits ou non, puis en fonction de leur statut. Il existe quatre grands types d'utilisateur : invité (sans inscription), adulte, enfant et enseignant. Nous verrons dans un premier temps les fonctionnalités disponibles pour les utilisateurs invités, puis celles pour tous les utilisateurs enregistrés. Nous détaillerons ensuite les fonctionnalités spécifiques selon le statut.

Invité

Ce type d'utilisateur a la possibilité de s'inscrire sur la plateforme, de jouer en tant qu'invité et de consulter les règles du jeu.

Pour accéder à l'ensemble des fonctionnalités que nous venons de détailler, le visiteur dispose d'un menu lui permettant de naviguer entre les différentes rubriques. La figure 12 montre les différentes rubriques possibles

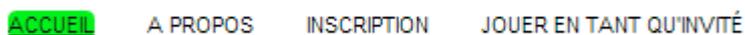


Figure 12 : Menu d'un invité

Accueil

Cette page permet aux utilisateurs de s'authentifier pour pouvoir accéder à toutes les fonctionnalités de la plateforme. Elle est composée d'un formulaire sur la partie droite de la page. Les champs nécessaires à l'authentification sont le login et le mot de passe de l'utilisateur. Il a aussi des liens à sa disposition sur la partie gauche qui ramènent vers des moteurs de recherches et vers le site de jeux de mots sur lequel se fonde notre réseau lexical (figure 13).



Figure 13 : Page d'accueil

A propos

Cette page explique comment jouer et quelles sont les fonctionnalités disponibles pour le type d'utilisateur courant. Elle est accessible à tous, seule la liste des fonctionnalités change en fonction du statut de l'utilisateur. Elle se compose de deux parties, celle de gauche décrivant les fonctionnalités ainsi que les règles de jeu et celle de droite contenant une vidéo montrant comment jouer (figure 14).

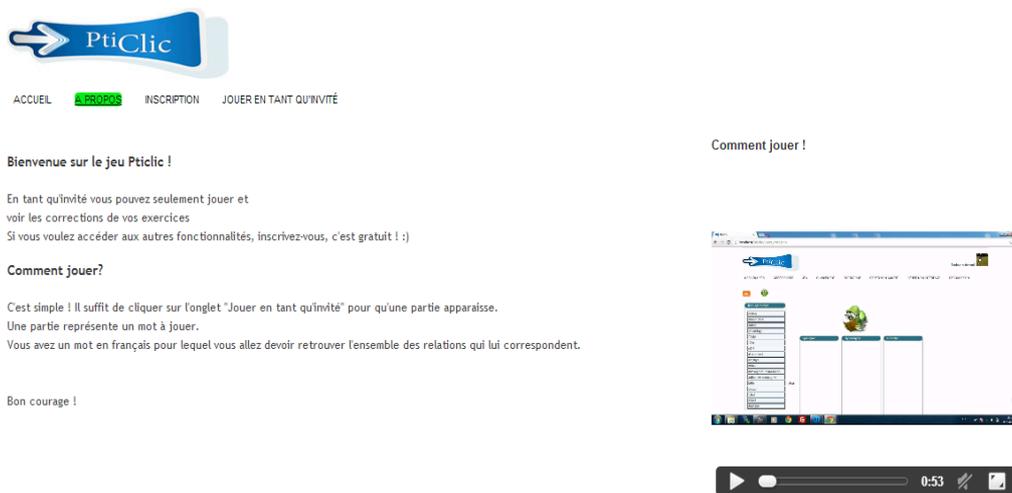


Figure 14 : Page A propos

Page d'inscription

La partie gauche de cette page permet aux utilisateurs de s'inscrire dans la plateforme. Le formulaire d'inscription requiert différentes informations : le login et le mot

de passe qu'il souhaite utiliser, son année de naissance, son prénom, son statut, son pays ainsi que son email (figure 15).

Avatar *

Login *

Mot De Passe *

Mot De Passe (Confirmation) *

Année De Naissance *

Prénom *

Statut * Enseignant

Pays * France

Email *

OK

Figure 15 : Page d'inscription

Page de jeu

Cette page permet aux utilisateurs de jouer. Le concept du jeu ne change pas ; les joueurs sont face à une liste de relations possibles (trois dans la figure 16 : synonymes, hyperonyme et corbeille) avec un terme donné (ici « chat ») et un nuage de mots (« clavardage », « minet », « animal », « félin », « chaton », « être vivant », « messagerie instantanée », etc.). L'attribution des relations se fait grâce à un glisser-déposer jusque dans les colonnes des différentes relations; tous les mots du nuage doivent être placés. Le joueur met dans « corbeille » les mots qui, dans l'exemple suivant, ne sont ni des « synonymes » ni des « hyperonymes » de « chat ».

Les parties proposées à ce type d'utilisateur sont générées automatiquement à partir des parties enregistrées par des enseignants. Ces derniers les ont créées de façon semi-automatique ou manuelle comme nous le verrons par la suite.



ACCUEIL A PROPOS INSCRIPTION **QUEL EST AN ANIMÉ**

OK



| liste des termes |
|------------------------|
| animal de compagnie |
| messagerie instantanée |
| minet |
| vertébré |
| animal |
| être vivant |
| mistigri |
| MSN |
| mammifère |
| clebardage |
| compagnon |
| tchat |
| diacgue |
| bête |
| félin |
| mincu |
| matou |
| féliné |
| chaton |

chat



Synonyme de chat

Hyperonyme de chat

Corbeille

Figure 16 : Page de jeu

A la fin d'une partie, les joueurs ont accès à la « correction ».

Quand il s'agit d'une partie pour les invités, les réponses correspondant à celles données par l'enseignant sont considérées comme justes et celles qui en diffèrent comme fausses. Les réponses fausses apparaissent en rouge et la catégorie à laquelle elles appartiennent est précisée entre parenthèse. Les réponses qui n'étaient placées dans aucune relation sont colorées en orange et affichées dans la bonne catégorie. Enfin, les réponses justes apparaissent en vert. Le joueur obtient un score calculé en pourcentage en fonction du nombre de réponses justes; le score maximum est de 100 points. A titre d'exemple, une partie contenant dix-neuf mots où le joueur a eu quatorze réponses justes obtient un score de 74 points. La figure 17 présente les résultats de cette partie.

Réponse juste
 Réponse mal placée
 Réponse oubliée
 Le mot principal de la partie est **chat**

Vous gagnez 74 points

Synonyme de chat

minet
 minou
 chaton
 dialogue
 MSN
 tchat
 clavardage (Il avait été placé dans Hyperonyme)
 matou
 mistigri
 maccsaria inct antanúa

Hyperonyme de chat

félin
 animal
 mammifère
 félidé
 être vivant
 bête
 vertébré
 compagnon
 animal de compagnie (Il avait été placé dans Synonyme)

Figure 17 : Feed-back et résultats du joueur : la correction

Pages communes

Une fois connectés, les utilisateurs ont accès à d'autres fonctionnalités. Adulte, enfant et enseignant ont accès au classement, à la gestion de leur compte et à la gestion de l'interface.

Classement

Le classement de tous les joueurs est affiché sur cette page avec pour chacun le pseudo, le nombre de parties jouées, le score maximal et le score minimal obtenus toutes parties confondues, le score moyen par partie et le score général. Il est possible de trier le score selon les critères notés ci-dessus (voir figure 18).

Classement

| | Pseudo | Score général | Nombre de partie | Moyenne par partie | Score maximal | Score minimal |
|---|---------|---------------|------------------|--------------------|---------------|---------------|
| 1 | mahmoud | 403 | 1 | 101.0000 | 101 | 101 |
| 2 | denzel | 228 | 1 | 74.0000 | 74 | 74 |

Trier selon le pseudo

Trier selon le score général

Trier selon les nombres des parties

Trier selon la moyenne par partie

Trier selon le score maximale

Trier selon le score minimale

Figure 18 : Page classement

Gérer mon compte

Ici, l'utilisateur peut modifier chaque information le concernant grâce à un formulaire se trouvant sur la partie gauche de la page (voir figure 19).

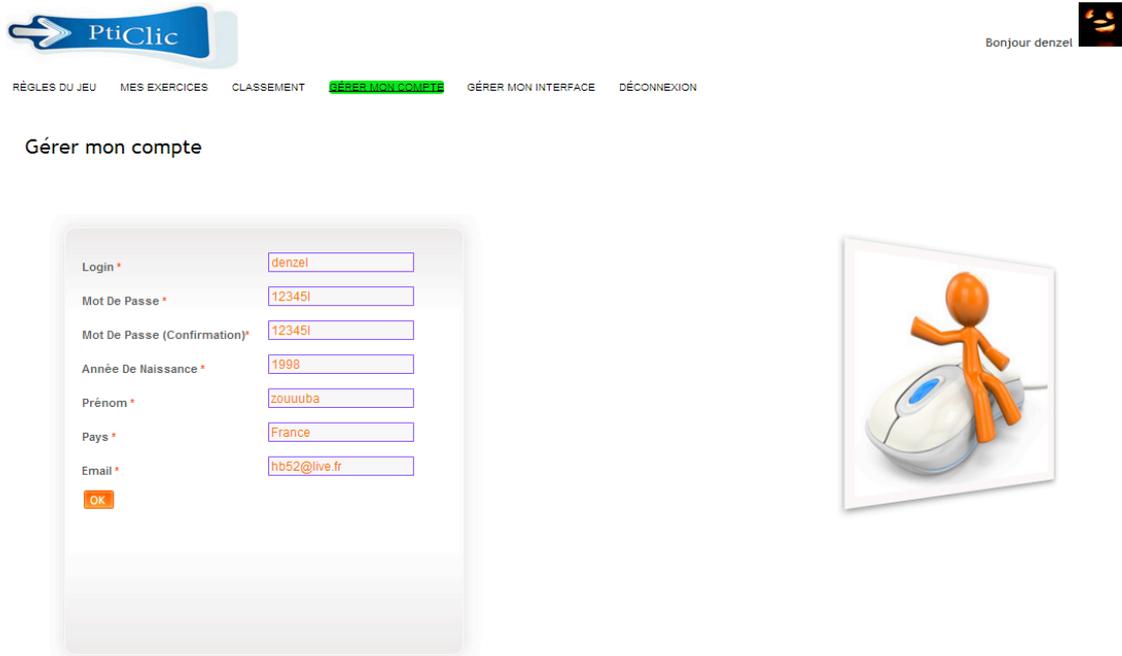
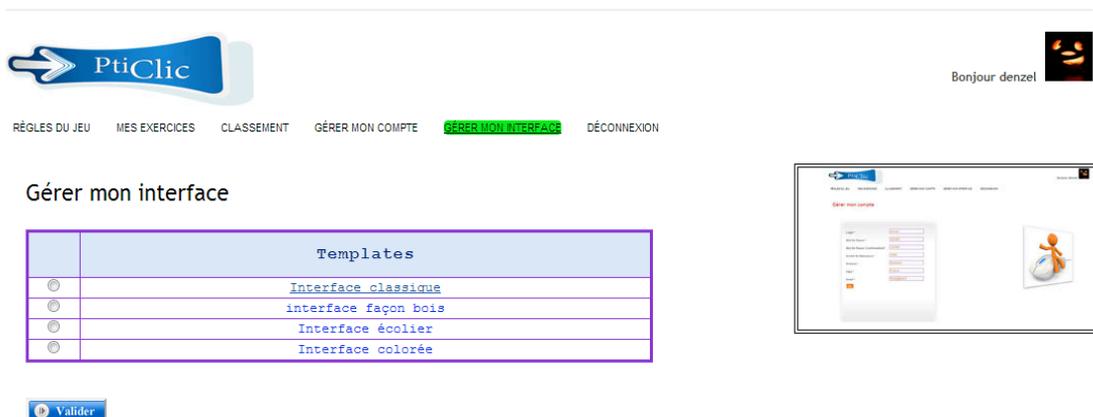


Figure 19 : Page gérer mon compte

Gérer mon interface

Cette page permet à l'utilisateur de choisir l'interface qu'il souhaite avoir dans la plateforme. Nous avons choisi par défaut la couleur blanche, les autres interfaces proposés étant « façon bois », « écolier » et « colorée ».

Au survol du lien, un aperçu de la page avec l'interface choisie s'affiche (voir figure 20).



| Templates | |
|-----------------------|--------------------------------------|
| <input type="radio"/> | Interface classique |
| <input type="radio"/> | interface façon bois |
| <input type="radio"/> | Interface écolier |
| <input type="radio"/> | Interface colorée |

Figure 20 : Page Gérer mon interface

Spécificités propres à l'utilisateur « Adulte »

Ce type d'utilisateur a accès aux mêmes fonctionnalités qu'un invité mais il a la possibilité de gérer son compte et son interface ainsi que de voir sa position dans le classement également.

De la même façon, un menu lui permet de naviguer entre les différentes rubriques. La figure 21 montre les différentes rubriques possibles

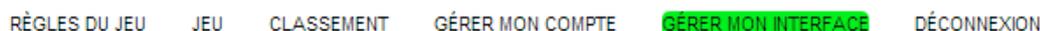
A horizontal menu bar with six items: 'RÈGLES DU JEU', 'JEU', 'CLASSEMENT', 'GÉRER MON COMPTE', 'GÉRER MON INTERFACE', and 'DÉCONNEXION'. The item 'GÉRER MON INTERFACE' is highlighted with a green background.

Figure 21 : Menu adulte

En ce qui concerne le jeu, la génération de partie est également spécifique pour les utilisateurs adultes.

Les parties qui sont proposées pour les adultes sont générées automatiquement à partir des parties enregistrées des cinquante mots les plus fréquents en français (femme, homme, enfant, animal, maison, etc.) et toutes leurs relations. Au moment de la génération d'une partie, tout est choisi de façon aléatoire : le mot principal, les relations et le nombre de relations, ce qui permet d'obtenir la liste des termes à jouer.

Les réponses correspondant à celles données par JDM sont considérées comme justes et celles qui en diffèrent comme fausses. Les systèmes de coloration et de score restent les mêmes.

Spécificités propres à l'utilisateur « Enfant »

Compte tenu de leur statut d'enfant, ces utilisateurs n'ont pas accès aux parties générées automatiquement afin d'éviter de donner des parties contenant des mots d'argot, des mots trop spécifiques qu'ils ne connaissent pas ou des termes sexuels. La figure 22 montre que l'onglet « jeu » est différent entre les deux types d'utilisateurs. En effet, l'enfant a une rubrique baptisée « Mes exercices ».

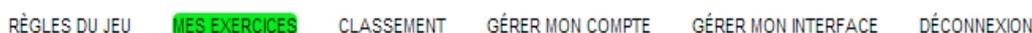
A horizontal menu bar with six items: 'RÈGLES DU JEU', 'MES EXERCICES', 'CLASSEMENT', 'GÉRER MON COMPTE', 'GÉRER MON INTERFACE', and 'DÉCONNEXION'. The item 'MES EXERCICES' is highlighted with a green background.

Figure 22 : Menu enfant

En effet, les parties accessibles aux enfants sont des parties créées par des enseignants, attribuées en fonction de l'âge du joueur qui est une des informations requises lors de l'inscription.

Ce type d'utilisateur est donc rattaché à un ou plusieurs groupes. Pour être inscrit en tant qu'enfant il faut donner son âge au moment de l'inscription. En effet, les parties créées par les enseignants sont attribuées en fonction de l'âge.

Il joue les parties faites par ses enseignants En cliquant sur cette rubrique, il accède à la liste des exercices à faire donnés par ses enseignants, affichés à partir du titre et du mot principal à jouer (figure 23). Il peut également voir le prénom et le nom de son enseignant ainsi que la date de création de la partie. Il peut également consulter la liste des parties déjà faites classée par leur date de réalisation.



Figure 23 : Page exercices à faire et faits pour un enfant

Une partie en cours ressemble à ce que montre la figure 24. Il y a donc la liste des mots à jouer (colonne de gauche), le mot principal et les différentes catégories dans lesquelles il faut faire glisser les mots à jouer. Bien entendu, il peut placer les mots qu'il juge distrayants, c'est-à-dire ce qu'il ne juge pas appartenir à une des relations proposées, dans la corbeille.

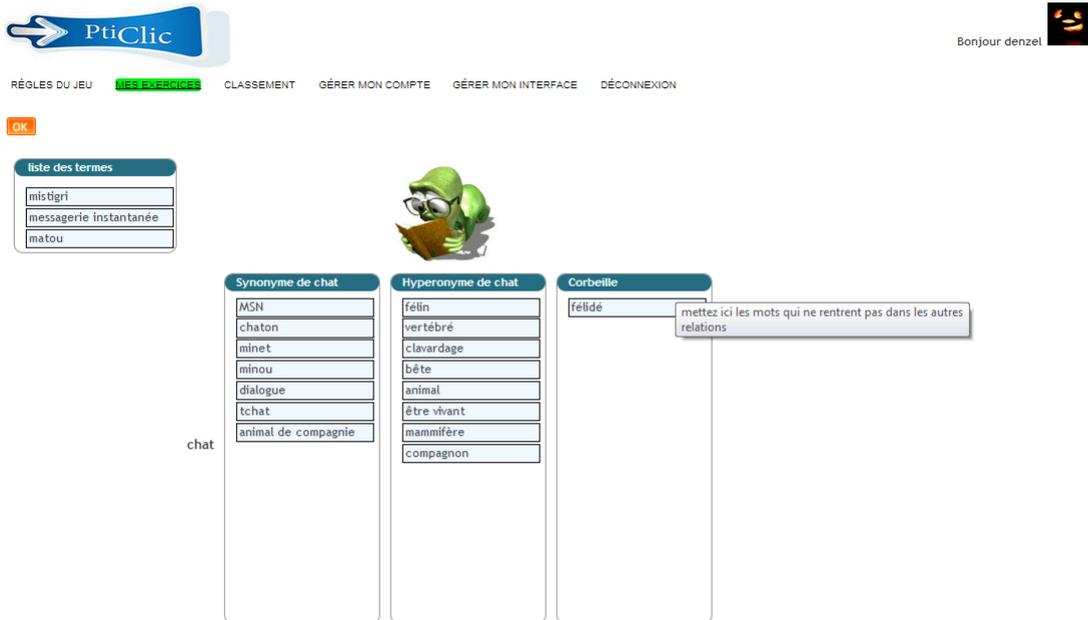


Figure 24 : Partie en cours

Une fois l'exercice fini, l'enfant a la possibilité de consulter les corrections données par ses enseignants, ce qu'illustre la figure 25.

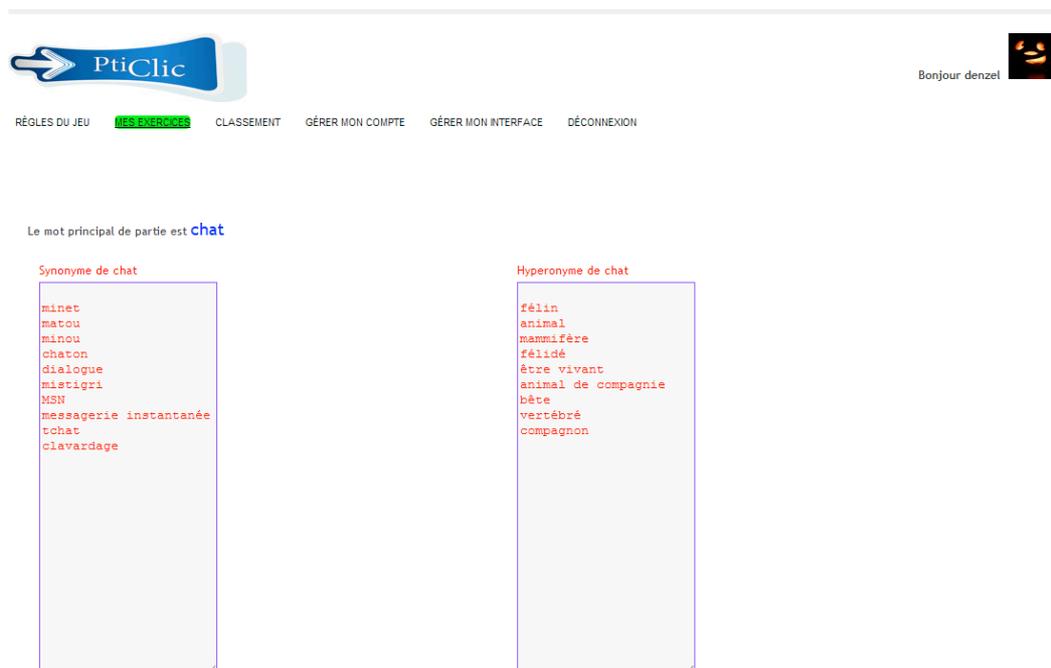


Figure 25 : Correction donnée par l'enseignant

Quand un enfant a fini toutes les parties données par ses enseignants, il peut faire d'autres exercices plus faciles ou plus difficiles (plus ou moins deux ans d'écart d'âge) qui peuvent être donnés par ses enseignants ou par d'autres enseignants (figure 26).



Figure 26 : Page parties supplémentaires

Spécificités propres à l'utilisateur « Enseignant »

L'enseignant est l'utilisateur qui a le plus de fonctionnalités à sa disposition. Il peut non seulement jouer, gérer ses paramètres personnels, consulter son classement mais dispose aussi de ses propres catégories. Son menu contient le plus d'onglets (figure 27).



Figure 27 : Menu enseignant

L'enseignant peut gérer ses groupes ainsi que ses élèves, consulter des statistiques sur l'âge d'acquisition des termes et des relations utilisées et aussi créer des parties.

Mes groupes

Cette page permet de gérer les groupes ; il peut donc ajouter, modifier, supprimer et consulter les parties données pour un groupe. Il peut aussi voir le classement au sein de chaque groupe et gérer la liste des élèves de chaque groupe (ajout, modification ou suppression). Sur la partie gauche de la page se trouve le tableau contenant les groupes et les fonctionnalités possibles et sur la partie droite, une vidéo montrant comment utiliser ces fonctionnalités (figure 29).

La colonne « Options » permet de gérer les groupes. La première icône (engrenage) supprime le groupe, la deuxième (crayon) permet d'éditer le groupe⁵, la troisième icône renvoie à la page de consultation des parties attribuées au groupe⁶ tandis que la quatrième

⁵ Voir figure 41 de l'annexe 4

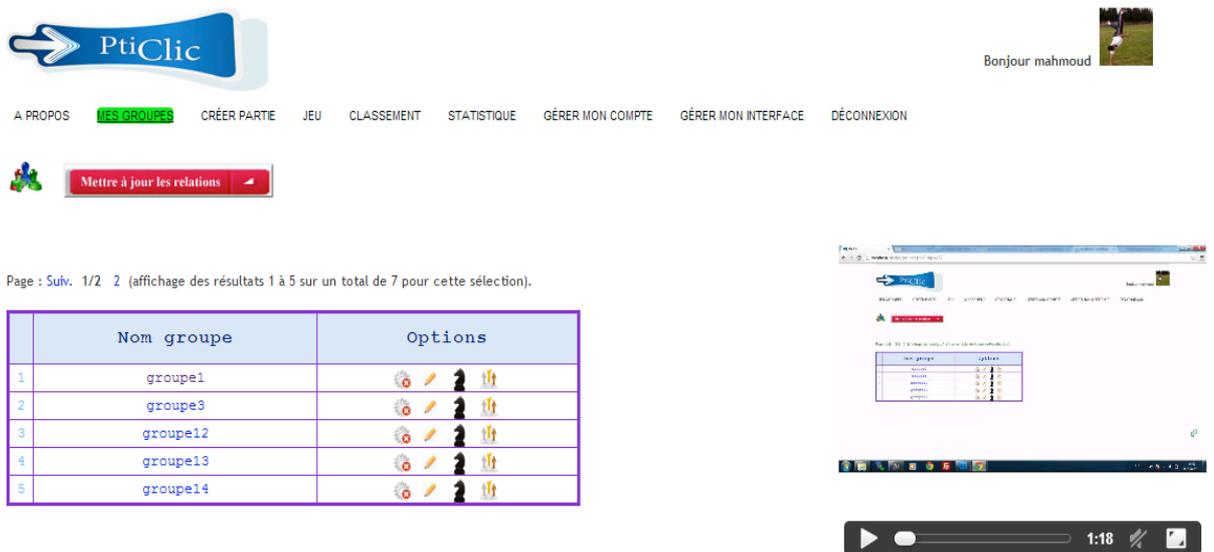
⁶ Voir figure 42 et 43 de l'annexe 4

redirige l'enseignant vers la page de classement des élèves au sein du groupe sélectionné⁷ (voir figure 28).

| | Nom groupe | Options |
|---|------------|---|
| 1 | groupe1 |     |
| 2 | groupe3 |     |
| 3 | groupe12 |     |
| 4 | groupe13 |     |
| 5 | groupe14 |     |

Figure 28 : Opérations possibles sur un groupe

Un enseignant joue également le rôle de modérateur de relations. En effet, il peut mettre à jour les relations en fonction des différentes réponses des joueurs. Cette mise à jour permet la synchronisation des relations par rapport aux termes contenus dans JDM et PtiClic. Pour ce faire, l'enseignant doit cliquer sur le bouton rouge en haut à gauche « Mettre à jour les relations ».



Page : Suiv. 1/2 2 (affichage des résultats 1 à 5 sur un total de 7 pour cette sélection).

| | Nom groupe | Options |
|---|------------|---|
| 1 | groupe1 |     |
| 2 | groupe3 |     |
| 3 | groupe12 |     |
| 4 | groupe13 |     |
| 5 | groupe14 |     |

Figure 29 : Page mes groupes

Pour accéder à la liste des élèves d'un groupe, l'enseignant doit cliquer sur le nom de groupe qu'il souhaite consulter. Il est alors redirigé vers la page présentée par la figure 30.

⁷ Voir figure 44 de l'annexe 4

C'est sur cette page que l'enseignant gère ses élèves. A gauche de la page se trouve la liste des élèves sous forme de tableau. Ce dernier contient différentes informations telles que le login, le prénom, le mot de passe et l'âge. Le mot de passe est visible pour l'enseignant que dans le cas où l'enfant a été ajouté par lui ; dans le cas contraire il apparaît sous forme d'étoiles. L'enseignant peut supprimer ou modifier⁸ un membre de la liste (icône 3 montrée dans la figure 30). Il peut aussi inscrire un nouvel enfant⁹ (icône 2) ou le rajouter à partie d'une liste existante¹⁰ (icône 1) grâce aux icônes situées juste au dessus du tableau. Un clic sur une des ces trois icônes affiche un formulaire dans la zone 4.

The screenshot shows the PtiClic interface. At the top left is the PtiClic logo. To the right, it says 'Bonjour mahmoud' next to a small profile picture. Below the header is a navigation menu with items: A PROPOS, MES GROUPES (highlighted in green), CRÉER PARTIE, JEU, CLASSEMENT, STATISTIQUE, GÉRER MON COMPTE, GÉRER MON INTERFACE, and DÉCONNEXION. Below the menu are three circular icons: a green arrow pointing left (1), a person icon with a plus sign (2), and a person icon (1). Below these icons is the text 'Page : Suiv. 1/2 2 (affichage des résultats 1 à 5 sur un total de 9 pour cette sélection)'. The main content is a table with 5 rows and 5 columns. The columns are 'Login', 'Prénom', 'Mot de passe', 'Âge', and 'Opérations'. The 'Opérations' column contains icons for adding, deleting, and editing. The third icon in the first row is circled in red and labeled '3'. To the right of the table is a large red rectangular box labeled '4'.

| | Login | Prénom | Mot de passe | Âge | Opérations |
|---|---------|---------|--------------|-----|-------------------------|
| 1 | denzel | zouuba | 123451 | 15 | [Add] [Delete] [Edit] 3 |
| 2 | bilel | bolbol | ***** | 14 | [Add] [Delete] [Edit] |
| 3 | kljkljk | lkljlkj | ***** | 14 | [Add] [Delete] [Edit] |
| 4 | nidio | nidhal | ***** | 8 | [Add] [Delete] [Edit] |
| 5 | iheb87 | iheb | ***** | 16 | [Add] [Delete] [Edit] |

Figure 30 : Liste des élèves d'un groupe

Page statistique

Cette page permet de consulter les statistiques sur l'âge d'acquisition d'une relation entre deux termes, le nombre de fois où le premier mot a été joué, le nombre de fois où le deuxième mot a été joué, le nombre de parties contenant le premier mot comme mot principal et le nombre de parties contenant le deuxième mot comme mot principal (voir figure 31).

Les statistiques de l'âge d'acquisition par rapport à l'âge du joueur sont affichées à gauche en deux tableaux, avec également un affichage graphique sous forme de

⁸ Voir figure 45 et 46 de l'annexe 5

⁹ Voir figure 48 de l'annexe 5

¹⁰ Voir figure 47 de l'annexe 5

diagramme en dessous. La partie de droite contient une vidéo montrant comment faire les statistiques.

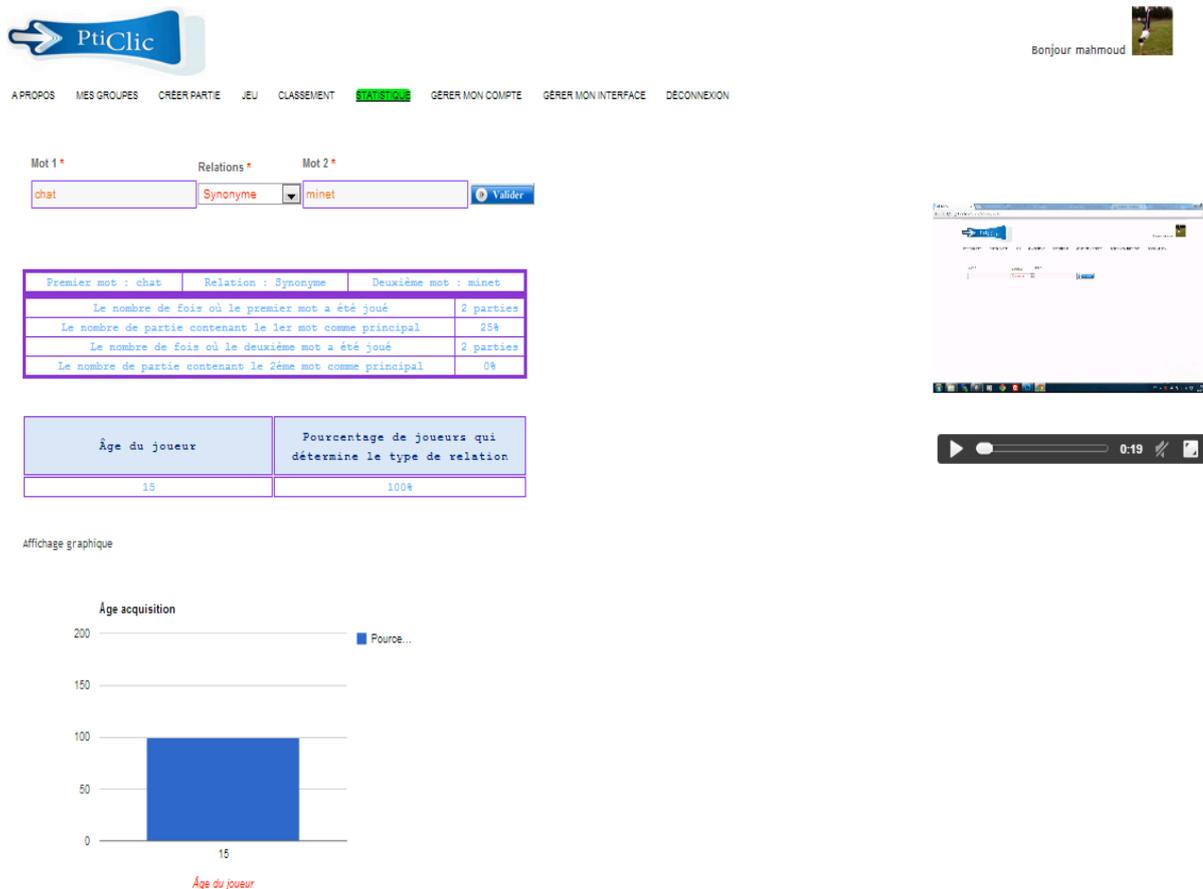


Figure 31 : Page Statistique

Page création de partie

Dans cette nouvelle version, nous avons implémenté une interface pour aider l'enseignant à créer des parties adaptées à une tranche d'âge donnée. La création de partie est une des innovations majeures du nouveau PtiClic. Il y a trois façons de créer une nouvelle partie :

- Création manuelle : l'enseignant choisit lui-même les mots et les relations qu'il veut proposer aux joueurs.
- Création mixte : l'enseignant sélectionne les relations extraites de JDM à partir du terme de son choix. Il peut ensuite modifier les relations et les mots à sa guise.

De façon générale, la création de partie se fait en plusieurs étapes. La première étape est de sélectionner les types de relations qu'il souhaite faire travailler. Pour ce faire,

il doit cocher deux à quatre relations dans la liste de choix possibles. Il doit ensuite saisir le mot qu'il souhaite utiliser pour la partie et valider. La partie inférieure de la page contient alors une sélection de termes dans les colonnes correspondant aux relations choisies. Il peut ensuite modifier les listes proposées en enlevant ou ajoutant des termes. Enfin, l'enseignant doit donner un titre, l'âge du public visé et le groupe auquel la partie est attribuée pour pouvoir valider l'exercice. Comme sur la plupart des pages, l'utilisateur peut s'aider du tutoriel vidéo disponible sur le côté droit.

La figure 32 présente l'interface de création d'une partie mixte accessible à un enseignant. Ici, il a choisi de travailler sur le mot cible « chat » et a sélectionné les relations « synonyme » et « antonyme ». Il a alors accès à ce qui est fourni dans le réseau de JDM et peut choisir les termes qu'il désire garder.

Créer partie semi-automatique Créer partie manuelle

Veuillez sélectionner les relations lexicales

Veuillez choisir entre 2 et 4 relations

Synonyme : Hyperonyme : Antonyme : Hyponyme :
 Idées associés : Domaine : Est une partie de : Caractéristique :
 Raffinement : Position : Locution : Lieu :
 Agent : Patient : Instrument : Sentiment :

Ajouter une relation à partir de jeux de mots

Entrer le mot à chercher

chat

Synonyme de chat

minet
matou
minou
chaton
dialogue
mistigri
MSN
messagerie instantanée
tchat
clavardage

Antonyme de chat

chien

test 15 Groupe : groupe1

Figure 32 : Page de création mixte d'une partie

Les parties créées par les enseignants sont attribuées à la catégorie d'utilisateur de type « enfant » et correspondent à un âge donné. L'enseignant peut choisir de garder

« privée » une partie, c'est-à-dire qu'elle n'est accessible que pour un « groupe » d'élèves, ou « publique », c'est-à-dire qu'elle sera indiquée pour un âge donné et pourra être jouée par n'importe quel utilisateur.

PtiClic 1.2 : Côté système

Acquisition automatique

Le nouveau PtiClic construit son réseau lexical à partir de deux méthodes JDM et lui-même. L'idée est de construire des relations entre les mots à travers les parties joués avec les données de JDM. Il s'agit d'une part d'enregistrer les relations récupérées à partir de ce dernier ; d'autre part celles provenant de PtiClic. Les données récupérées à partir de PtiClic sont collectées après l'enregistrement des réponses aux parties créées tant automatiquement que par l'enseignant (création mixte ou manuelle) et réinjectées régulièrement dans le réseau JDM.

Connaissances sur l'âge d'acquisition

Lors de l'inscription d'un joueur, des données sur son âge sont fournies et stockées au sein d'une base de données. Nous pouvons donc, pour reprendre l'exemple donné en introduction, savoir que a% des enfants de b ans pensent que souris « est un » animal mais que seulement c% de la même tranche d'âge pensent que souris « est une partie de » ordinateur et d%. Ce sont ces informations qui seront importantes pour savoir, une fois le seuil fixé, quelle relation est « acquise » à quel âge par rapport à un couple de termes donnés. A l'heure actuelle les données sont affichées dans la page « statistiques » car elles ne sont pas très importantes. Dès qu'une grande campagne de test aura lieu, nous pourrons mettre en place un système permettant d'insérer dans la base de données l'âge d'acquisition des termes joués.

Chapitre 8 – Perspectives

Nous souhaitons avoir un retour vis-à-vis de la plateforme. Nous avons donc décidé de faire suivre un questionnaire aux personnes qui allaient la tester afin d'avoir un retour uniformisé. Il était important pour nous d'avoir un retour en matière d'ergonomie d'une part, et d'autre part au niveau des fonctionnalités telle que la possibilité de créer des parties.

Dans un premier temps nous allons aborder les améliorations proposées par les utilisateurs que nous avons choisi de mettre en place. Nous traiterons dans une seconde section des expérimentations à venir et plus particulièrement des journées TICE ALPES. Avant de conclure ce chapitre par les perspectives d'évolution pour PtiClic.

Questionnaire

A partir des questionnaires plusieurs améliorations ont été proposées par les utilisateurs. Certaines de ces modifications ont été apportées.

Au niveau des modifications ergonomiques, la première suggestion était de mettre en place des tutoriels d'aide sur les différentes pages de l'application. Nous avons donc décidé d'intégrer des vidéos dans la partie droite de la page.

La deuxième suggestion était d'ajouter un onglet règles du jeu présent pour l'ensemble des statuts d'utilisateur. Un testeur nous a également proposé de colorer les onglets de la page courante, pour plus de clarté dans la navigation au sein du jeu.

Certains des noms de catégories n'étaient pas très clairs, les utilisateurs ont suggéré de les expliciter. En effet, la catégorie « Hyponyme » a par exemple été explicité de la façon suivante : « Hyponyme de [MOT] ». Nous avons également ajouté une légende pour la catégorie « Corbeille » qui n'a pas semblé très transparente aux premières utilisations.

La gestion de l'interface était la quatrième suggestion. Le principe était de permettre aux utilisateurs de choisir l'interface qu'il souhaitait utiliser. Pour ce faire, nous avons créé trois interfaces en supplément de l'interface classique (voir chapitre ... Refonte de PtiClic).

La dernière suggestion en matière d'ergonomie concernait l'authentification automatique après l'inscription. Ce dernier point a également été réglé.

Enfin, certains utilisateurs se sont rendus compte d'un biais au niveau de la correction des parties. En effet, le feed back à la fin des parties ne correspondait pas toujours aux réponses du joueur. Ces biais ont donc été corrigés

Nous avons aussi arrondi les scores des parties jouées à l'entier supérieur (28.6 devient 29).

Expérimentation à venir

Dans le cadre des journées Tice Alpes nous avons l'occasion de présenter notre projet. Notre participation à ces journées se fera en deux temps ; exposé et atelier.

Le but de l'exposé est de présenter la plateforme suivi d'une discussion. L'exposé a une durée de 20-25 minutes et d'une discussion de 15-20 minutes. Cet exposé a pour objectif de faire connaître l'application aux utilisateurs potentiels (chercheurs comme enseignants). Cet exposé est complété par un atelier.

Le but de cet atelier est de faire découvrir PtiClic en le mettant entre les mains des enseignants. Ainsi, nous pourrons les aider et obtenir leur avis sur cette application. Ceci nous permettra aussi d'avoir des retours d'utilisateurs potentiels et de pouvoir faire des modifications en fonctions de leurs demandes/attentes. Notre atelier aura lieu en salle informatique, d'une durée de 90 minutes lors desquelles nous ferons tester aux participants PtiClic. Le second objectif visé par cette expérimentation est d'inciter les enseignants à utiliser PtiClic avec leurs élèves. Ceci nous permettra notamment de récupérer un certain nombre de données quant à l'âge d'acquisition des termes chez les enfants.

Perspectives d'évolution pour PtiClic

Les retours que nous aurons grâce aux journées Tice Alpes nous permettront également de récupérer un retour critique sur l'outil. D'autres améliorations pourront ainsi y être apportées. Nous pouvons cependant prévoir un certain nombre de perspectives d'évolution pour PtiClic.

Nous pouvons dans un premier temps noter quelques-unes des améliorations possibles à court terme.

Parmi les améliorations envisagées, nous pouvons citer tout d'abord la proposition d'un utilisateur d'avoir le choix du mot que l'on souhaite jouer. Puis, d'un point de vue ergonomique, il serait intéressant de permettre aux utilisateurs d'organiser les éléments de

la page comme ils le souhaitent. Une autre suggestion est de pouvoir changer de statut (enseignant, joueur, etc.) en cas d'erreur lors de la création de compte. La dernière amélioration que nous proposons serait de mettre en place une page d'actualité destinée aux enseignants dans laquelle ils pourraient voir les nouvelles parties créées par les autres enseignants, et les mots qu'ils ont utilisés.

Enfin il serait judicieux d'adapter PtiClic aux nouvelles technologies (Smartphone et Tablette). En effet, nombre de jeux en ligne possèdent maintenant une version mobile du jeu. Le glisser-déposer que nous avons mis en place rend le jeu d'autant plus intuitif pour ce type de technologies.

Nous pouvons ensuite observer dans quelle mesure il est possible d'améliorer le suivi de l'élève sur la plateforme. Cette partie comporte essentiellement les améliorations possibles vis-à-vis du suivi des enfants. Nous pourrions lui afficher le score qu'il a obtenu pour chaque partie ainsi que des statistiques par rapport aux parties jouées (graphiques pour le nombre de partie par date, les scores par rapport aux parties, etc.).

Enfin, il serait intéressant de pouvoir extraire des informations par rapport aux réponses des élèves. Ce point peut à terme nous permettre d'améliorer la reconnaissance de l'âge d'acquisition en fonction d'un texte, mais aussi évaluer le niveau d'un élève à partir du texte qu'il a écrit.

Conclusion

Le mémoire que nous venons de présenter s'inscrit dans le projet PtiClic, nous avons pu voir que le logiciel existait mais qu'il ne rencontrait pas un grand succès. Nous avons également pu observer qu'il contenait de nombreuses lacunes. Nous avons donc décidé d'améliorer l'application. Le second logiciel avec lequel nous allons la fusionner est PtiClicKids. Ce dernier n'avait pas été réalisé, il était donc envisageable de fusionner ces deux projets en un seul et de le tester avec une plus grande quantité de utilisateurs.

Afin de prendre en compte l'ensemble des paramètres nécessaires à l'amélioration et à la fusion des logiciels PtiClic et PtiClicKids, nous avons réalisé un état de l'art. Ce dernier nous a permis de faire des choix méthodologiques et informatiques. Nous avons notamment pu déterminer quelle était la meilleure méthode pour l'acquisition automatique lexicale à savoir dans notre cas : un outil d'acquisition lexicale contributif par l'intermédiaire de JeuxDeMots.

Nous avons ensuite conçu une nouvelle application incluant PtiClic et PtiClicKids. Une fois qu'un premier prototype a été implanté, nous avons pu diffuser le jeu afin d'avoir un premier retour des utilisateurs. Ce retour positif a tout de même mis en lumière certaines améliorations nécessaires à la poursuite de PtiClic (correction des parties, aide dans la navigation au sein de la plateforme, etc.).

La prochaine étape dans le développement de PtiClic est de l'expérimenter avec une certaine catégorie d'utilisateurs : les enseignants. Ce test se fera lors des journées Tice-Alpes 2013. Une fois que les enseignants auront expérimenté le jeu, ils pourront le faire passer à leurs élèves ce qui nous permettra notamment de collecter un maximum d'informations sur l'âge d'acquisition des termes chez les enfants.

Bibliographie

[LAFOURCADE 2007] M. Lafourcade, Making people play for Lexical Acquisition. In *Proc. SNLP 2007, 7th Symposium on Natural Language Processing*. Pattaya, Thaïlande, 13-15 December 2007, p.8

[LAFOURCADE, JOUBERT & RIOU 2009] Lafourcade M., Joubert A. & Riou S.. Sens et usages d'un terme dans un réseau lexical évolutif. *TALN*, Senlis (France). 2009

[LANDAUER & DUMAIS 1997] Landauer, T.K. & Dumais, S.T. A solution to Plato's problem: the Latent Semantic Analysis theory of acquisition, induction and representation of knowledge. *Psychological Review*, 104(2), 211-240. 1997.

[LIEBERMAN *et al*, 2007] 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, January 2007.

[LEON & MILLON 2005] Leon, S. et Millon, C. Acquisition semi-automatique de relations lexicales bilingues (français-anglais) à partir du Web. *RECITAL*, Dourdan (France), vol. 1 p. 595. 2005

[MESSIANT, NAKAMURA & VOYATZI 2009] Messiant, C., Nakamura, T. & Voyatzi, S. La complémentarité des approches manuelle et automatique en acquisition lexicale. *TALN*. Senlis (France) 2009

[MESNAGER, 2011] MESNAGER Jean Le vocabulaire et son enseignement Évaluation de la difficulté des textes La part du lexique, Novembre 2011

[MESNAGER 2002] MESNAGER Jean Pour une étude de la difficulté des textes, ou la lisibilité revisitée, *Le Français aujourd'hui*, n° 137, avril 2002

[MESNAGER 1989] MESNAGER Jean. Lisibilité des textes pour enfants : un nouvel outil ? In: *Communication et langages*. N°79, 1er trimestre 1989. pp. 18-38.

[MILLER, BECKWITH & FELLBAUM 1990], Miller, G, Beckwith R et Fellbaum C. Introduction to WordNet: An On-line Lexical Database, révision Août 1993

[SAADANI & BERTRAND-GASTALDY 2000], Saadani, L et Bertrand-Gastaldy,S. Cartes Conceptuelles et Thésaurus : Essai de Comparaison Entre Deux Modèles de Représentation Issus de Différentes Traditions Disciplinaires.2000

[SAGOT et al 2011], Sagot, B, Fort, K, Adda, G, Mariani, J et Lang, B. Un turc mécanique pour les ressources linguistiques : critique de la myriadisation du travail parcellisé. *TALN Montpellier* (France) 2011

[ZACKLAD, 2007] Zacklad, M : Classification, thésaurus, ontologies, folksonomies : comparaisons du point de vue de la recherche ouverte d'information (ROI), 2007

[ZAMPA & LAFOURCADE, 2012] Zampa, V et Lafourcade, M. PtiClic et PtiClic-Kids : jeux avec les mots permettant une acquisition lexicale par le joueur et par la machine. *STICEF*, 18, 135-157, 2012.

[VON AHN, KEDIA & BLUM, 2006] Von, Ahn, L, Kedi, M & Blum, M. Verbosity: A Game for Collecting Common-Sense Facts. In Proceedings of ACM CHI 2006 Conference on Human Factors in Computing Systems (ACM Press), volume 1 of Games p. 75-78. 2006

[VON AHN & DABBISH 2004] Von Ahn, L. & Dabbish, L. Labelling Images with a Computer Game. *ACM Conference on Human Factors in Computing Systems (CHI)*. pp. 319-326, New York, NY, USA. ACM Press. 2004

[VON AHN, 2006] von Ahn, L. (2006). Games with a purpose. *Computer*, 39(6):92–94.

[VON AHN *et al.*, 2006a] von Ahn, L., Liu, R., and Blum, M. (2006) Peekaboom: a game for locating objects in images. In *Proceedings of CHI '06*, pages 55–64.

[VON AHN *et al.*, 2006b] von Ahn, L., Kedia M., and Blum M. (2006) Verbosity: a game for collecting common-sense facts. In Proceedings of ACM CHI 2006 *Conference on Human Factors in Computing Systems*, volume 1 of Games, pages 75–78, 2006.

[VON AHN & DABBISH., 2008] von Ahn L., Dabbish L. Designing Games With a Purpose. *Communication of the ACM*, 51(8):58–67 August 2008.

Table des annexes

| | |
|---|----|
| ANNEXE 1 DIAGRAMMES DE CAS D'UTILISATION | 60 |
| ANNEXE 2 DIAGRAMME DE CLASSES | 64 |
| ANNEXE 3 DIAGRAMME DE SEQUENCES | 67 |
| ANNEXE 4 LES DIFFERENTES OPERATIONS POSSIBLES AU SEIN D'UN GROUPE | 68 |
| ANNEXE 5 LES DIFFERENTES OPERATIONS POSSIBLES SUR LES COMPTES « ELEVES »..... | 70 |

Annexe 1

Diagrammes de cas d'utilisation

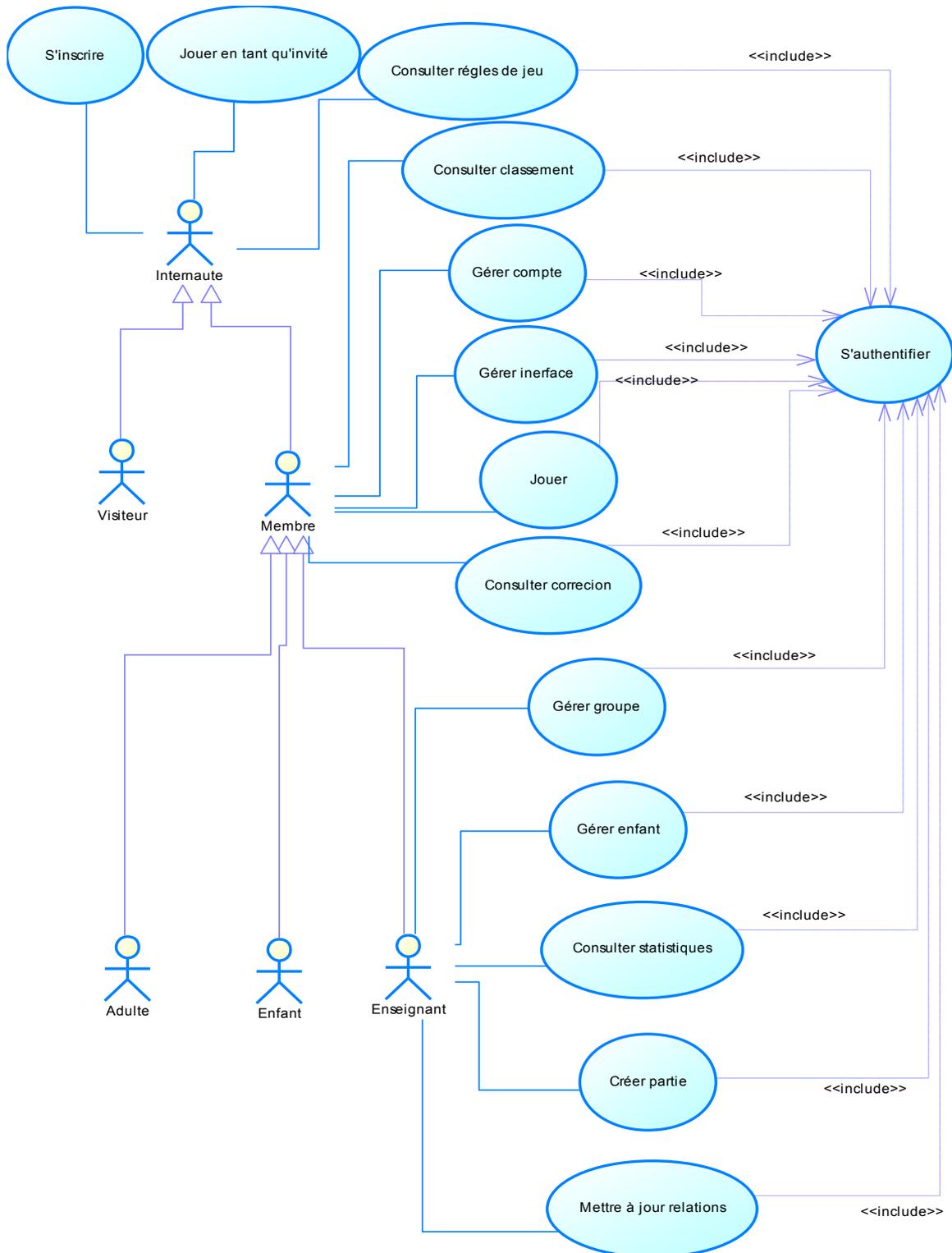


Figure 33 : Diagramme de cas d'utilisation général

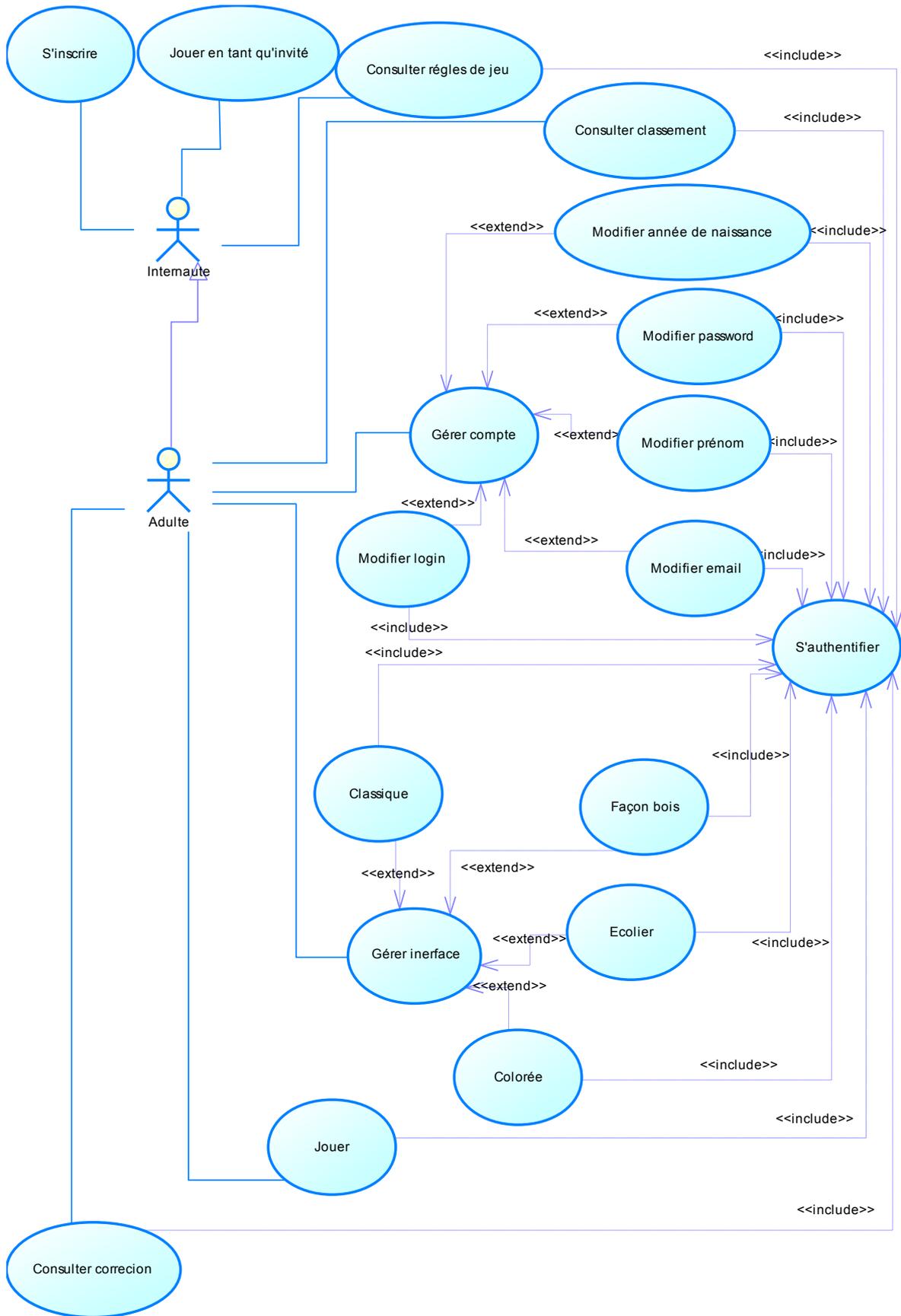


Figure 34 : Diagramme de cas d'utilisation pour l'adulte

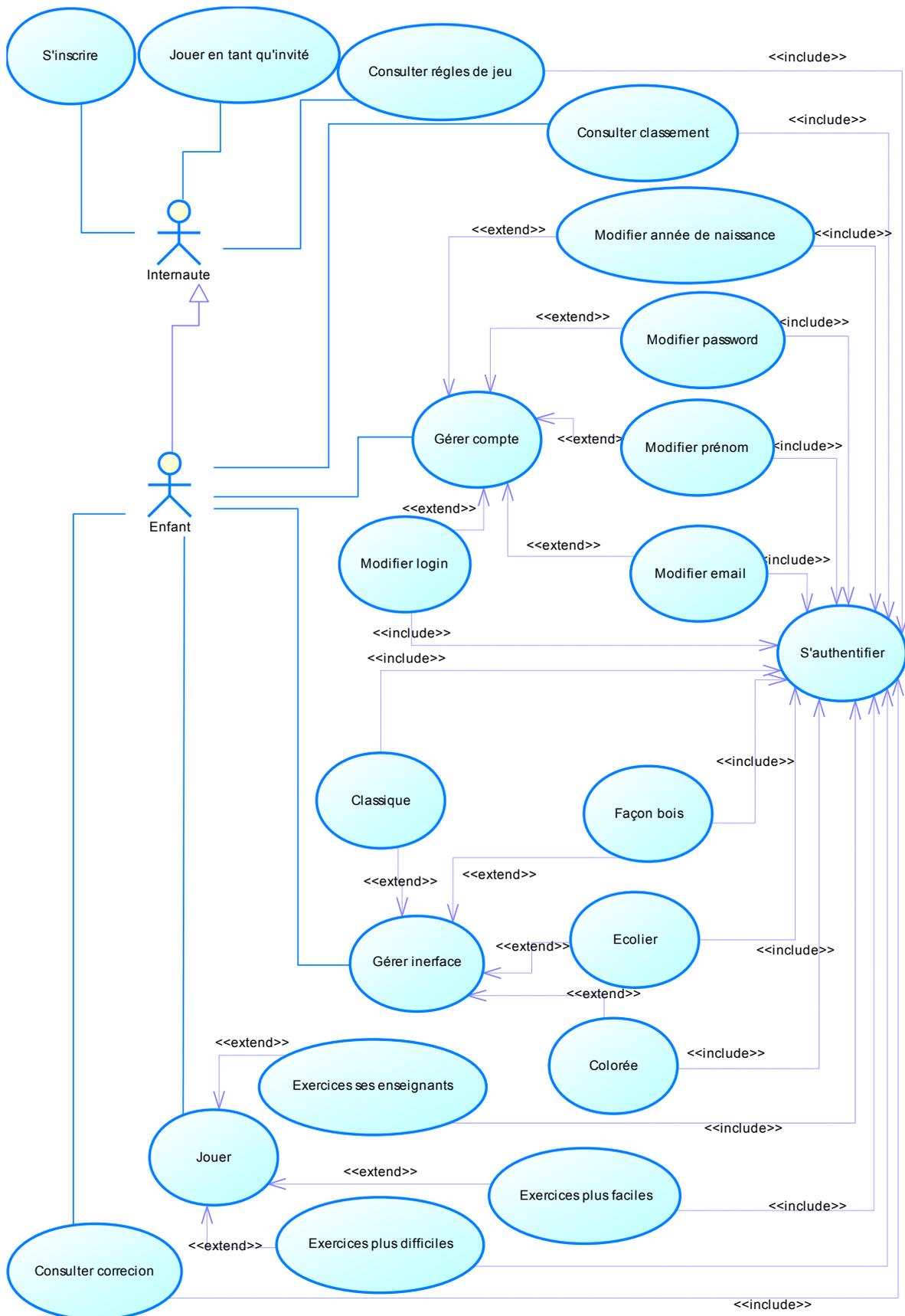


Figure 35 : Diagramme de cas d'utilisation pour l'enfant

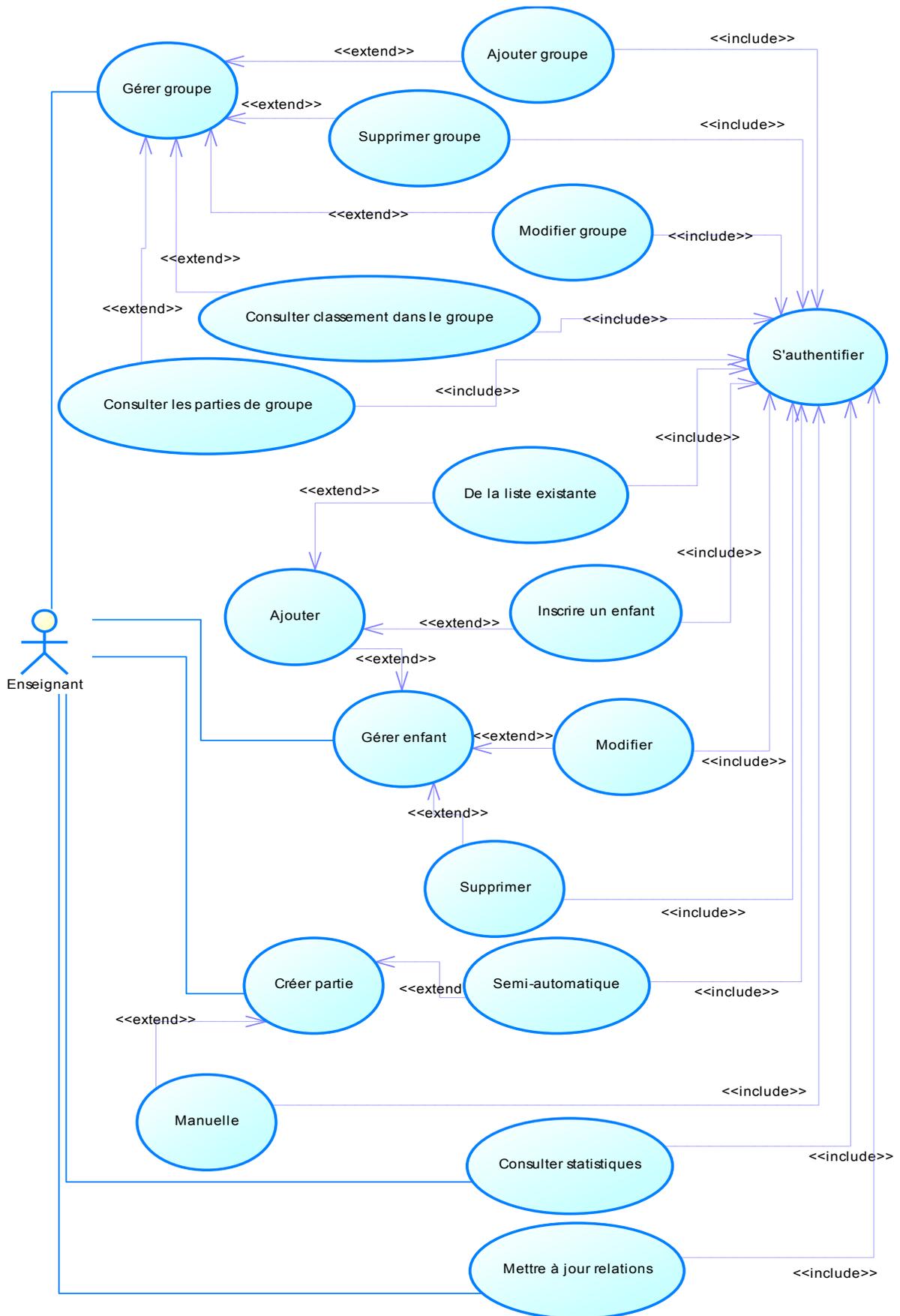


Figure 36 : Diagramme de cas d'utilisation des fonctionnalités propres à l'enseignant

Annexe 2 Diagramme de classes

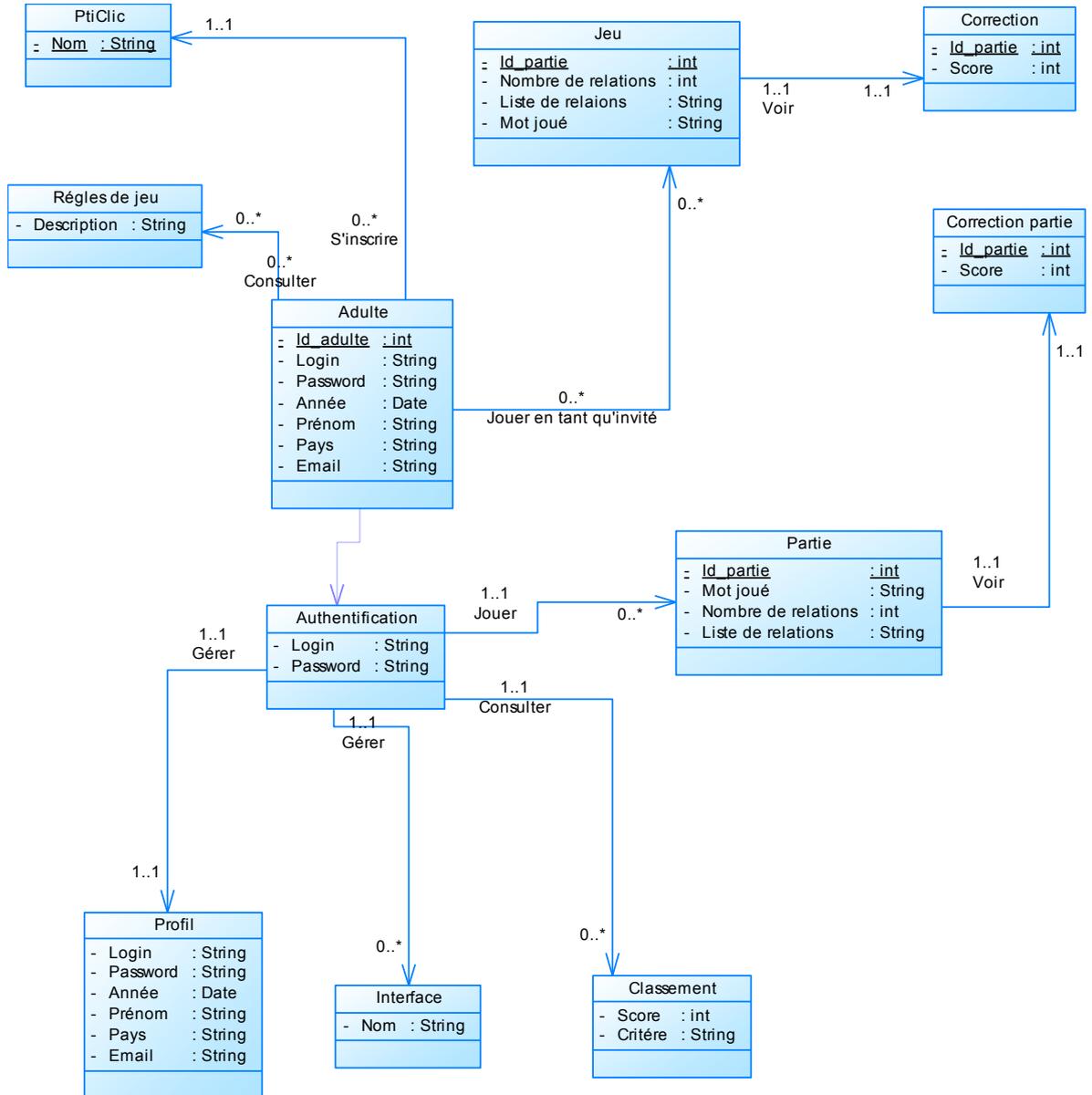


Figure 37 : Diagramme de classe pour l'adulte

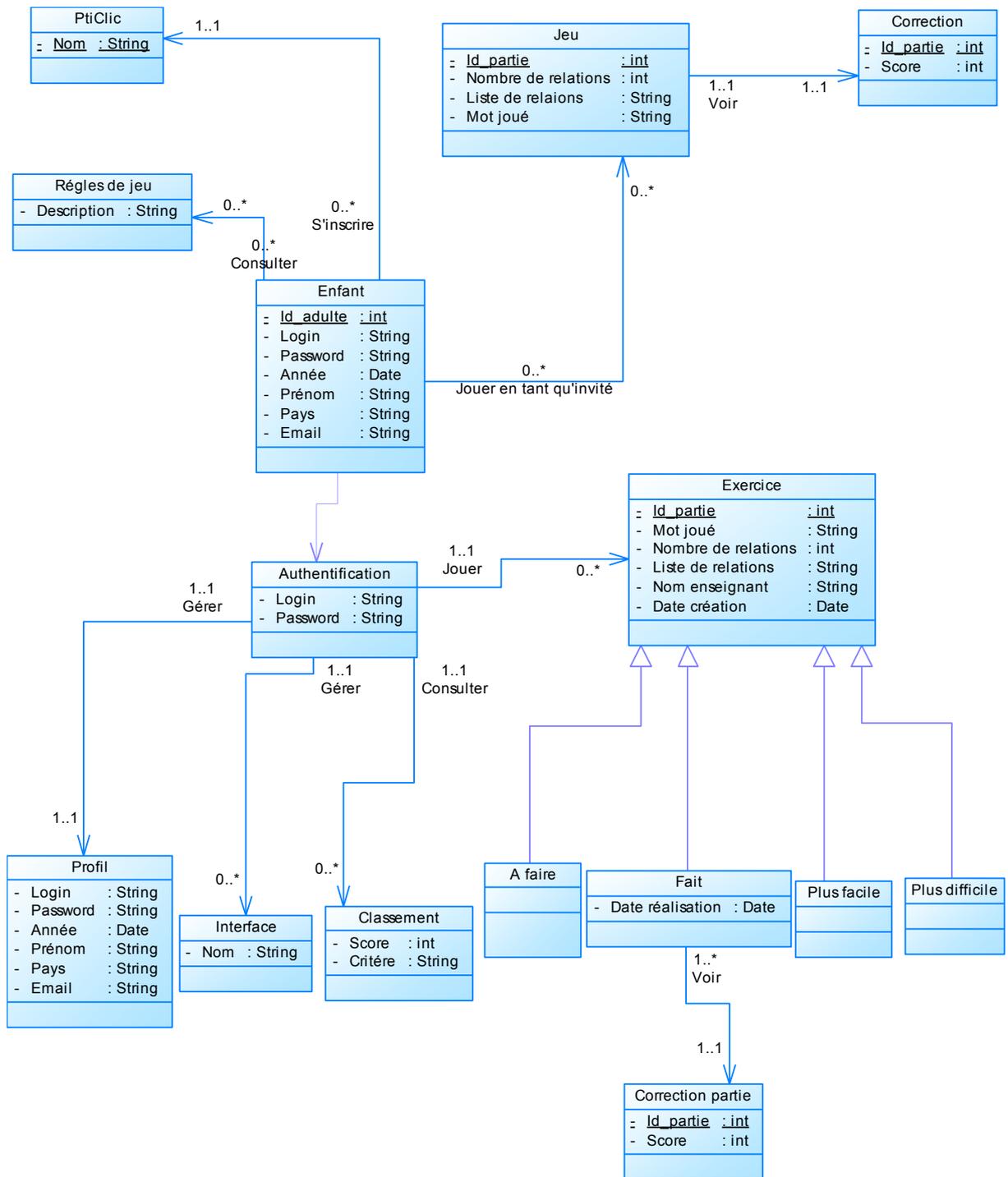


Figure 38 : Diagramme de classe pour l'enfant

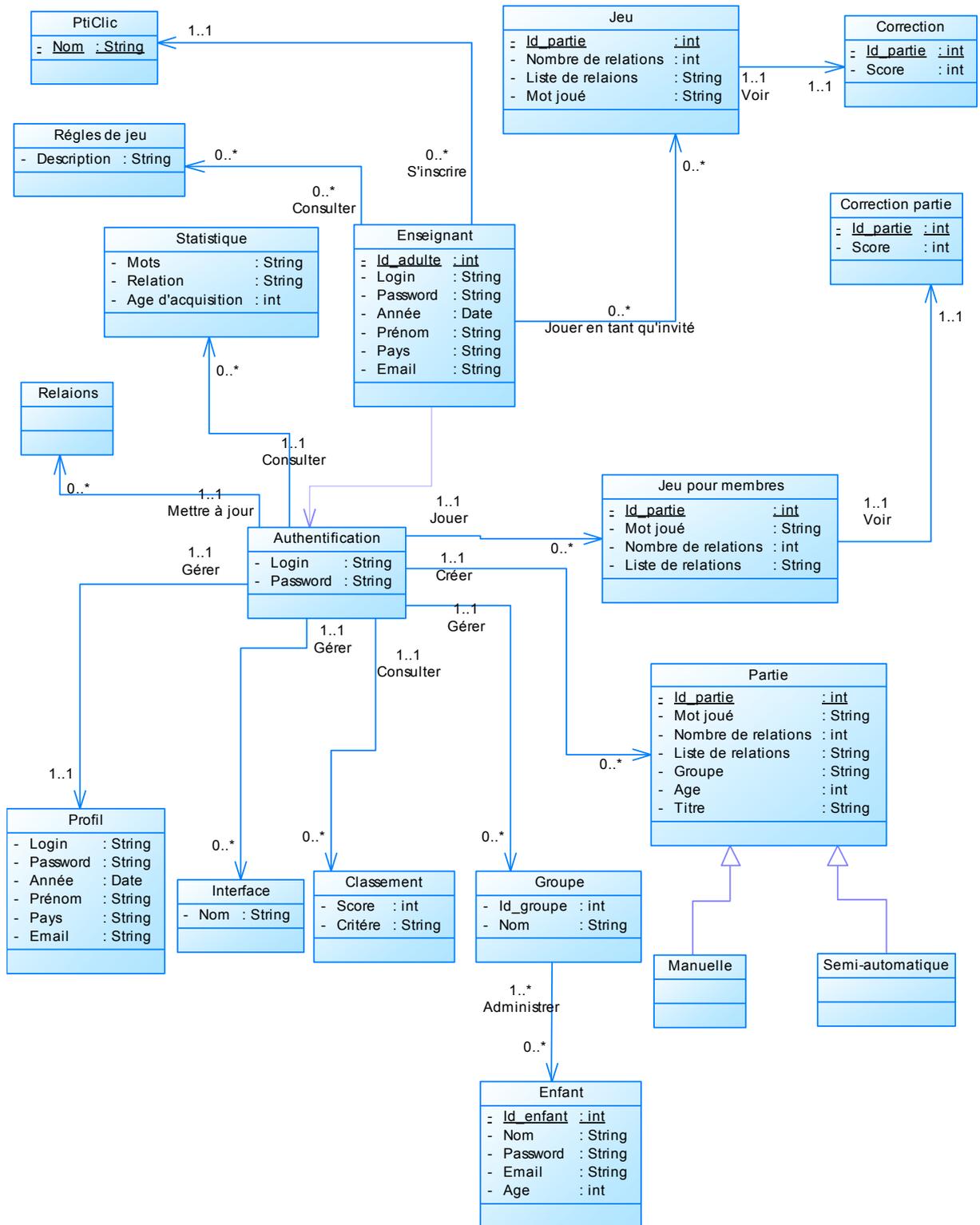


Figure 39 : Diagramme de classe pour l'enseignant

Annexe 3 Diagramme de séquences

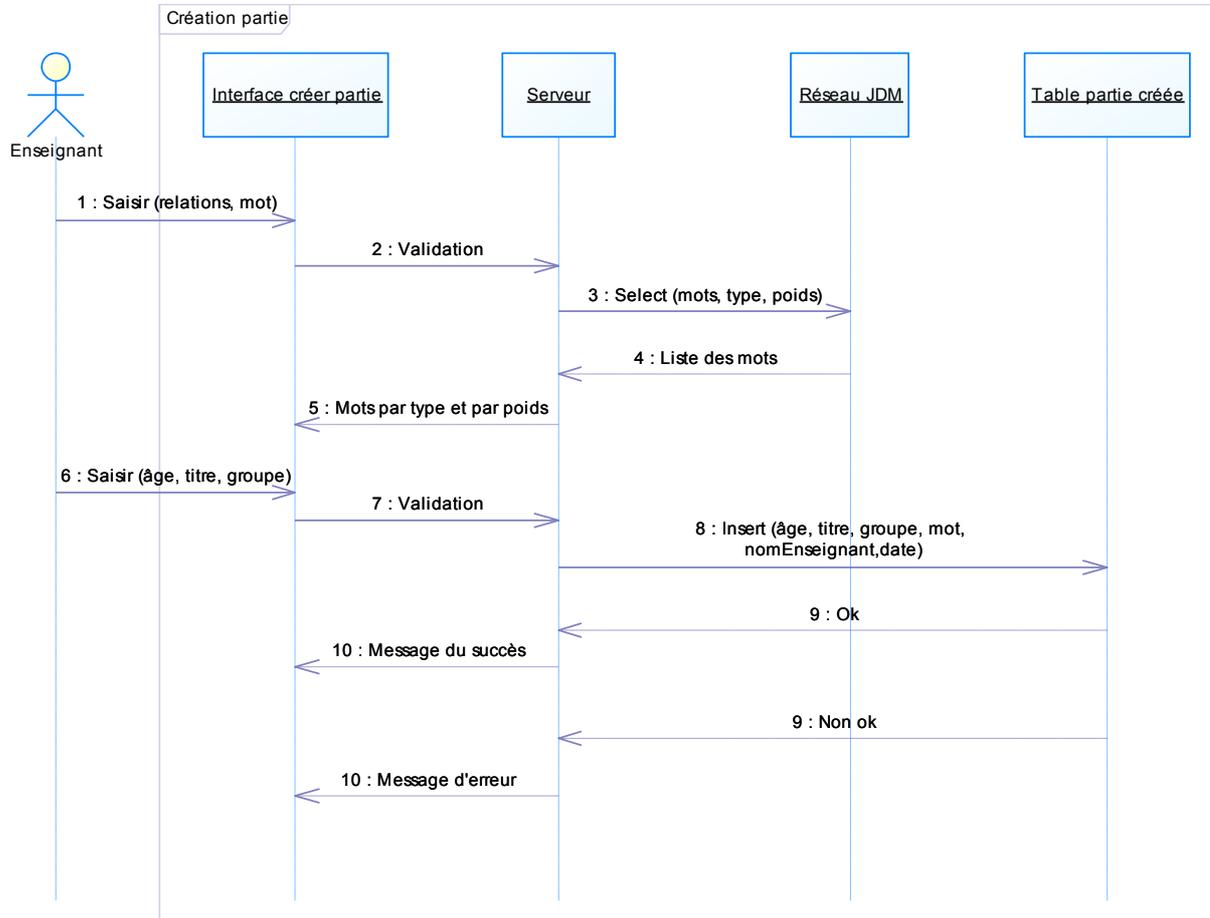


Figure 40 : Diagramme de séquence pour la création de partie propre à l'enseignant

Annexe 4

Les différentes opérations possibles au sein d'un groupe

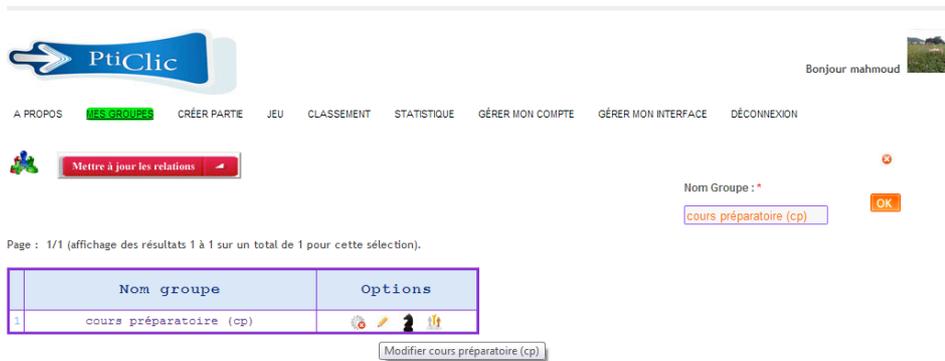


Figure 41 : capture d'écran pour la modification d'un groupe existant



Figure 42 : Page les parties créées pour le groupe

Le mot principal de partie est **chat**

Synonyme de chat

```

minet
matou
minou
chaton
dialogue
mistigri
MSN
messagerie instantanée
tchat
clavardage
    
```

Hyperonyme de chat

```

félin
animal
mammifère
félidé
être vivant
animal de compagnie
bête
vertébré
compagnon
carnivore
    
```

Figure 43 : Page de consultation des réponses pour une partie attribuée au groupe

Classement dans cours préparatoire (cp) :



| | Pseudo | Score général | Nombre de partie | Moyenne par partie | Score maximal | Score minimal |
|---|----------|---------------|------------------|--------------------|---------------|---------------|
| 1 | mohammed | 100 | 1 | 100.0000 | 100 | 100 |

Trier selon le pseudo

Trier selon le score générale

Trier selon les nombres des parties

Trier selon la moyenne par partie

Trier selon le score maximale

Trier selon le score minimale

Figure 44 : Page de classement au sein d'un groupe

Annexe 5

Les différentes opérations possibles sur les comptes « élèves »

Bonjour mahmoud

A PROPOS **LES GROUPES** CRÉER PARTE JEU CLASSEMENT STATISTIQUE GÉRER MON COMPTE GÉRER MON INTERFACE DÉCONNEXION

Page : 1/1 (affichage des résultats 1 à 3 sur un total de 3 pour cette sélection).

| | Login | Prénom | Mot de passe | Âge | Opérations |
|---|--------|---------|--------------|-----|------------|
| 1 | jack | reacher | ***** | 26 | |
| 2 | mym | Mehdi | ***** | 33 | |
| 3 | mathis | mathis | ***** | 6 | |

Figure 45 : Page de consultation des membres du groupe

Bonjour mahmoud

A PROPOS **LES GROUPES** CRÉER PARTE JEU CLASSEMENT STATISTIQUE GÉRER MON COMPTE GÉRER MON INTERFACE DÉCONNEXION

Page : 1/1 (affichage des résultats 1 à 3 sur un total de 3 pour cette sélection).

| | Login | Prénom | Mot de passe | Âge | Opérations |
|---|--------|---------|--------------|-----|-------------------------------|
| 1 | jack | reacher | ***** | 26 | |
| 2 | mym | Mehdi | ***** | 33 | Modifier jack |
| 3 | mathis | mathis | ***** | 6 | |

Login *

Mot De Passe *

Prénom *

Pays *

Email *

Figure 46 : Formulaire de modification d'informations sur l'élève

Page : 1/1 (affichage des résultats 1 à 3 sur un total de 3 pour cette sélection).

| | Login | Prénom | Mot de passe | Âge | Opérations |
|---|--------|---------|--------------|-----|------------|
| 1 | jack | reacher | ***** | 26 | |
| 2 | mym | Mehdi | ***** | 33 | |
| 3 | mathis | mathis | ***** | 6 | |

| | Login | Prénom | Âge |
|---|-----------|------------|-----|
| 1 | alesan | Alessandro | 8 |
| 2 | Charlotte | Charlotte | 9 |
| 3 | lancelot | Lancelot | 10 |

OK

Figure 47 : formulaire d'ajout d'élève existant

Page : 1/1 (affichage des résultats 1 à 3 sur un total de 3 pour cette sélection).

| | Login | Prénom | Mot de passe | Âge | Opérations |
|---|--------|---------|--------------|-----|------------|
| 1 | jack | reacher | ***** | 26 | |
| 2 | mym | Mehdi | ***** | 33 | |
| 3 | mathis | mathis | ***** | 6 | |

Login *

 Mot De Passe *

 Mot De Passe (Confirmation)*

 Année De Naissance *

 Prénom *

 Pays *

 Email *

 OK

Figure 48 : Formulaire d'ajout d'élève pour l'enseignant

Table des figures

| | |
|---|----|
| FIGURE 1 SCHEMA DE DESCRIPTION D'UN RESEAU LEXICAL [ZAMPA & LAFOURCADE 2011, p.7] | 13 |
| FIGURE 2 : COPIE D'ECRAN DU JEU <i>THE ESP GAME</i> (VON AHN & DABBISH, 2004, p.2) | 23 |
| FIGURE 3 : COPIE D'ECRAN DU JEU <i>VERBOSITY</i> (VON AHN, KEDIA & BLUM, p.2) | 24 |
| FIGURE 4 : EXEMPLE D'UNE PARTIE SUR JDM..... | 25 |
| FIGURE 5 : EXEMPLE DES RESULTATS COMMUNS A DEUX JOUEURS POUR LA MEME PARTIE SUR JDM | 26 |
| FIGURE 6 : <i>EXEMPLE DE JEU AVEC L'ANCIEN PTICLIC</i> | 30 |
| FIGURE 7 : STRUCTURE DE L'ARCHITECTURE TROIS TIERS | 33 |
| FIGURE 8 : DIAGRAMME DE CAS D'UTILISATION POUR LE VISITEUR | 34 |
| FIGURE 9 : DIAGRAMME DE CLASSES POUR LE VISITEUR | 35 |
| FIGURE 10 : DIAGRAMME DE SEQUENCE POUR L'INSCRIPTION | 36 |
| FIGURE 11 : PRESENTATION DES DIFFERENTES INTERFACES | 37 |
| FIGURE 12 : MENU D'UN INVITE | 38 |
| FIGURE 13 : PAGE D'ACCUEIL | 39 |
| FIGURE 14 : PAGE A PROPOS | 39 |
| FIGURE 15 : PAGE D'INSCRIPTION..... | 40 |
| FIGURE 16 : PAGE DE JEU | 41 |
| FIGURE 17 : FEED-BACK ET RESULTATS DU JOUEUR : LA CORRECTION | 42 |
| FIGURE 18 : PAGE CLASSEMENT..... | 42 |
| FIGURE 19 : PAGE GERER MON COMPTE | 43 |
| FIGURE 20 : PAGE GERER MON INTERFACE..... | 43 |
| FIGURE 21 : MENU ADULTE | 44 |
| FIGURE 22 : MENU ENFANT..... | 44 |
| FIGURE 23 : PAGE EXERCICES A FAIRE ET FAITS POUR UN ENFANT | 45 |
| FIGURE 24 : PARTIE EN COURS..... | 46 |
| FIGURE 25 : CORRECTION DONNEE PAR L'ENSEIGNANT | 46 |
| FIGURE 26 : PAGE PARTIES SUPPLEMENTAIRES | 47 |
| FIGURE 27 : MENU ENSEIGNANT | 47 |
| FIGURE 28 : OPERATIONS POSSIBLES SUR UN GROUPE | 48 |
| FIGURE 29 : PAGE MES GROUPES..... | 48 |
| FIGURE 30 : LISTE DES ELEVES D'UN GROUPE | 49 |
| FIGURE 31 : PAGE STATISTIQUE..... | 50 |
| FIGURE 32 : PAGE DE CREATION MIXTE D'UNE PARTIE | 51 |
| FIGURE 33 : DIAGRAMME DE CAS D'UTILISATION GENERAL..... | 60 |
| FIGURE 34 : DIAGRAMME DE CAS D'UTILISATION POUR L'ADULTE | 61 |
| FIGURE 35 : DIAGRAMME DE CAS D'UTILISATION POUR L'ENFANT | 62 |
| FIGURE 36 : DIAGRAMME DE CAS D'UTILISATION DES FONCTIONNALITES PROPRES A L'ENSEIGNANT | 63 |
| FIGURE 37 : DIAGRAMME DE CLASSE POUR L'ADULTE | 64 |
| FIGURE 38 : DIAGRAMME DE CLASSE POUR L'ENFANT | 65 |
| FIGURE 39 : DIAGRAMME DE CLASSE POUR L'ENSEIGNANT..... | 66 |
| FIGURE 40 : DIAGRAMME DE SEQUENCE POUR LA CREATION DE PARTIE PROPRE A L'ENSEIGNANT..... | 67 |
| FIGURE 41 : CAPTURE D'ECRAN POUR LA MODIFICATION D'UN GROUPE EXISTANT | 68 |
| FIGURE 42 : PAGE LES PARTIES CREEES POUR LE GROUPE..... | 68 |
| FIGURE 43 : PAGE DE CONSULTATION DES REPONSES POUR UNE PARTIE ATTRIBUEE AU GROUPE..... | 69 |
| FIGURE 44 : PAGE DE CLASSEMENT AU SEIN D'UN GROUPE | 69 |
| FIGURE 45 : PAGE DE CONSULTATION DES MEMBRES DU GROUPE | 70 |
| FIGURE 46 : FORMULAIRE DE MODIFICATION D'INFORMATIONS SUR L'ELEVE | 70 |
| FIGURE 47 : FORMULAIRE D'AJOUT D'ELEVE EXISTANT | 71 |
| FIGURE 48 : FORMULAIRE D'AJOUT D'ELEVE POUR L'ENSEIGNANT..... | 71 |

Sigles et abréviations utilisés

| | |
|----------|--|
| TAL: | Traitement Automatique des Langues. |
| JDM : | JeuxDeMots. |
| LIDILEM: | Linguistique et Didactique des Langues Etrangères et Maternelles. |
| EIAH: | Environnements Informatiques pour les Apprentissages Humains. |
| LIRMM: | Laboratoire Informatique, de Robotique et de Microélectronique de Montpellier. |
| TALN: | Traitement Automatique des Langues Naturelles. |
| LSA: | Analyse Sémantique Latente. |
| AMT: | Amazon Mechanical Turk. |
| UML: | Unified Modeling Language. |
| W3C: | World Wide Web Consortium. |
| XML: | Extensible Markup Language. |
| HTTP: | Hypertext Transfer Protocol. |
| TICE: | Technologies de l'Information et de la Communication pour l'Enseignement. |

Table des matières

| | |
|--|-----------|
| REMERCIEMENTS | 6 |
| SOMMAIRE..... | 7 |
| INTRODUCTION..... | 8 |
| PARTIE 1 ETAT DE L'ART | 9 |
| <i>Chapitre 1 – Introduction.....</i> | <i>10</i> |
| <i>Chapitre 2 – Ressources lexicales et réseaux lexicaux sémantiques</i> | <i>12</i> |
| <i>Les ressources lexicales.....</i> | <i>12</i> |
| <i>Les réseaux lexico-sémantiques.....</i> | <i>12</i> |
| Définition..... | 13 |
| Exemples de réseaux lexicaux..... | 13 |
| <i>Les relations sémantiques.....</i> | <i>14</i> |
| Définition..... | 14 |
| Les relations sémantiques de Wordnet..... | 15 |
| Les relations sémantiques de JeuxDeMots..... | 16 |
| <i>Chapitre 3 – Construction de la ressource</i> | <i>19</i> |
| <i>Les outils d'acquisition automatique du lexique existants.....</i> | <i>19</i> |
| La méthode polysémique..... | 19 |
| LexSchem..... | 19 |
| LSA..... | 20 |
| <i>Les outils d'acquisition avec intervention humaine par des spécialistes</i> | <i>20</i> |
| WordNet..... | 20 |
| Autres ressources lexicales..... | 21 |
| <i>Les outils d'acquisition lexicale contributifs.....</i> | <i>21</i> |
| Volontaire..... | 21 |
| Jeux..... | 22 |
| Verbosity..... | 23 |
| JeuxdeMots..... | 24 |
| <i>Chapitre 4 – Conclusion.....</i> | <i>27</i> |
| PARTIE 2 CONCEPTION DE LA PLATEFORME..... | 28 |
| <i>Chapitre 5 – Les objectifs initiaux</i> | <i>29</i> |
| <i>Pticlic.....</i> | <i>29</i> |
| Interface initiale..... | 30 |
| <i>PticlicKids.....</i> | <i>30</i> |
| Dysfonctionnement..... | 31 |
| <i>Chapitre 6 – Refonte de PtiClic</i> | <i>32</i> |
| <i>Nos choix.....</i> | <i>32</i> |
| <i>Etude conceptuelle.....</i> | <i>33</i> |
| Outils de conception..... | 34 |
| Présentation..... | 34 |
| Vue statique..... | 34 |
| Diagrammes de cas d'utilisation..... | 34 |
| Diagrammes de classes..... | 35 |
| Vue dynamique..... | 35 |
| Diagrammes de séquence..... | 35 |
| <i>Aspects ergonomiques et sécurité.....</i> | <i>36</i> |
| Charte graphique..... | 36 |
| Police..... | 36 |
| Couleurs..... | 36 |
| Sécurisation des données..... | 37 |
| Inscription..... | 37 |
| <i>Public.....</i> | <i>37</i> |
| Invité..... | 38 |

| | |
|--|-----------|
| <i>Accueil</i> | 38 |
| <i>A propos</i> | 39 |
| <i>Page d'inscription</i> | 39 |
| <i>Page de jeu</i> | 40 |
| Pages communes..... | 42 |
| <i>Classement</i> | 42 |
| <i>Gérer mon compte</i> | 43 |
| <i>Gérer mon interface</i> | 43 |
| Spécificités propres à l'utilisateur « Adulte »..... | 44 |
| Spécificités propres à l'utilisateur « Enfant »..... | 44 |
| Spécificités propres à l'utilisateur « Enseignant »..... | 47 |
| <i>Mes groupes</i> | 47 |
| <i>Page statistique</i> | 49 |
| <i>Page création de partie</i> | 50 |
| PtiClic 1.2 : Côté système..... | 52 |
| <i>Acquisition automatique</i> | 52 |
| <i>Connaissances sur l'âge d'acquisition</i> | 52 |
| Chapitre 8 – Perspectives | 53 |
| <i>Questionnaire</i> | 53 |
| <i>Expérimentation à venir</i> | 54 |
| <i>Perspectives d'évolution pour PtiClic</i> | 54 |
| CONCLUSION | 56 |
| BIBLIOGRAPHIE | 57 |
| TABLE DES ANNEXES | 59 |
| TABLE DES FIGURES | 72 |
| SIGLES ET ABREVIATIONS UTILISES | 73 |
| TABLE DES MATIERES | 74 |

MOTS-CLÉS : Acquisition lexicale automatique, réseaux lexicaux, jeux sérieux, TAL

RÉSUMÉ

PtiClic est un logiciel conçu pour permettre des acquisitions lexicales par l'apprenant mais aussi par le système (enrichissement du réseau JeuxDeMots) et proposant une interface enseignant. Outre l'enrichissement du réseau lexical, PtiClic vise à obtenir l'âge moyen d'acquisition du lexique chez les enfants qu'il s'agisse du sens particulier d'un mot ou des relations qu'il entretient avec d'autres. Ces informations, inexistantes pour l'instant, sont importantes que ce soit pour les recherches en acquisition, pour le TAL, la linguistique, etc.,

KEYWORDS : Automatic lexical acquisition, lexical network, serious games, NLP

ABSTRACT

PtiClic is a software designed so that lexical acquisition would be possible not only for the learner but for the system as well, that is the implemented techniques automatically enriches the system's data base (JeuxDeMots network), along with a teacher interface. In addition to the lexical network's enrichment, Pticlic aims to evaluate the average age at which children are able to acquire lexicon with regard to the word's particular meaning or links it has with other words. This kind of information, so far nonexistent, is relevant and important when it comes to the acquisition field of research, NLP, linguistic, etc.



Déclaration anti-plagiat

Document à scanner après signature
et à joindre au mémoire électronique

DECLARATION

1. Ce travail est le fruit d'un travail personnel et constitue un document original.
2. Je sais que prétendre être l'auteur d'un travail écrit par une autre personne est une pratique sévèrement sanctionnée par la loi.
3. Personne d'autre que moi n'a le droit de faire valoir ce travail, en totalité ou en partie, comme le sien.
4. Les propos repris mot à mot à d'autres auteurs figurent entre guillemets (citations).
5. Les écrits sur lesquels je m'appuie dans ce mémoire sont systématiquement référencés selon un système de renvoi bibliographique clair et précis.

NOM : *Chafeddine* PRENOM : *Plahmoud*.....

DATE : *07/05/2013*.. SIGNATURE

