

Descriptif du projet SimpleApprenant

Responsable scientifique : Amalia Todirascu, LiLPa, Université de Strasbourg

Financement Idex Attractivité Université de Strasbourg (juin 2017-décembre 2018)

Le projet SimpleApprenant est multidisciplinaire, se situant au carrefour du traitement automatique des langues (TAL), de la linguistique et de la didactique des langues. Le projet a comme objectif d'étudier les problèmes que les apprenants du français langue étrangère peuvent rencontrer dans le processus de la compréhension et de la production écrite. Partant de ces constats, le projet développera des outils d'accompagnement aux apprenants du FLE, répondant ainsi à un des enjeux majeurs de l'université d'aujourd'hui : combattre l'échec et mieux accompagner les étudiants dans leur cursus. En effet, nous visons la construction des ressources et des outils pour aider les apprenants d'une langue étrangère à mieux apprendre une langue étrangère. Ainsi, nous développerons des outils de traitement automatique des langues, intégrés dans une plate-forme Web, permettant d'adapter le contenu des textes proposés au niveau de l'apprenant et d'améliorer la qualité de ses productions écrites en langue cible.

La demande des contenus adaptés en fonction du niveau de connaissances du public visé est grandissante dans plusieurs domaines, plus particulièrement dans le domaine de l'apprentissage d'une langue étrangère. Les étudiants qui rencontrent des difficultés de lecture ou de compréhension dans la langue étrangère sont souvent démotivés par ces difficultés, ce qui peut les conduire à abandonner leurs études. Parfois les manuels ne sont pas toujours adaptés aux connaissances du lecteur: les textes sont trop difficiles pour leur niveau de compétence dans la langue cible; le vocabulaire est trop complexe (certains mots ou termes ne sont pas connus par l'utilisateur); l'emploi d'expressions polylexicales et des figures de style peut-être compliqué pour un locuteur non-natif. Pour pallier à ces problèmes, les apprenants d'une langue étrangère doivent bénéficier d'une aide personnalisée par le biais des outils pédagogiques spécifiques, pour mieux apprendre ces expressions difficiles et pour améliorer la qualité de la production écrite.

En particulier, les expressions polylexicales représentent une source majeure de difficultés pour les apprenants d'une langue étrangère. Les expressions polylexicales couvrent une multitude d'expressions linguistiques : expressions idiomatiques, collocations, termes spécifiques à un domaine, entités nommées. Elles sont caractérisés par un degré de figement plus ou moins important (kick the bucket), des préférences lexicales fortes (poser une question mais pas *demander une question), des propriétés morphosyntaxiques spécifiques (faire **la fête** : le nom est toujours au singulier). Leur traduction n'est pas toujours possible par une stratégie de traduction mot-à-mot.

Malgré leur importance et les difficultés de compréhension et d'utilisation en écriture, peu de ressources numériques en ligne ou manuels de FLE proposent des leçons spécifiques pour l'apprentissage de ces expressions. Parfois, on les retrouve dans la rubrique 'vocabulaire', d'autre fois dans certains exercices. Il existe relativement peu de ressources Web destinés à l'apprentissage de ces expressions : DIRE autrement (<http://web5.uottawa.ca/direautrement/index.html#>). Quelques listes de mots typiques par niveau, mais il y a peu de dictionnaires ou lexiques gradués contenant une liste riche d'expressions idiomatiques ou collocations destinés aux apprenants d'un certain niveau.

Nous proposons des ressources linguistiques et des outils de TAL pour palier à ce problème. Nous constituons un site Web proposant des exercices permettant l'apprentissage des expressions polylexicales et des transformations du texte de l'apprenant en cas d'erreur. Pour ce faire nous construisons une base de données d'expressions polylexicales annotés en niveau CECRL (A1-C2). La construction de la base se fait à partir de plusieurs ressources disponibles (Wiktionnaire, Lexique-Grammaire etc.) et de manuels de FLE. Les exercices permettent d'apprendre le sens d'une expression et l'usage correct, de créer des textes contenant une liste d'expressions adaptées au niveau de l'apprenant. L'annotation des expressions se fait partiellement automatiquement, sur la

base d'un modèle construit à partir des corpus de manuels et sites à destination pédagogique. La base est complétée manuellement sur la base des manuels.

D'autre part, l'analyse d'erreurs dans les production écrites des apprenants permet la mise en place de stratégies de remédiation des production écrites : nous proposons des règles de correction des erreurs (erreurs des accord sujet-prédicat, erreurs d'orthographe, sujet manquant, combinaisons lexicales erronées), reformulation en cas de phrase trop longue (relatives qui sont imbriquées). Ces règles sont définies manuellement et s'appliquent aux textes analysés avec un analyseur syntaxique Mind the Gap (Coavoux et Crabbé, 2017).