

Sous la direction de **Roxana Anca TROFIN**

L'enseignement des langues de spécialité et l'apport du TAL

en collaboration avec **Georges ANTONIADIS**



**Editura UNIVERSITARIA
CRAIOVA, 2015**

Table des matières

« Le français de spécialité et le TAL - quels défis à relever ? », Introduction par Roxana Anca TROFIN	7
---	---

Les enjeux du discours scientifique

Roxana Anca TROFIN « Pour une utilisation des outils informatiques et du TAL dans l'analyse de textes techniques : la plate-forme du français technique et des technologies de la langue »	21
Alexandra CUNIȚĂ « L'analyse de la référence dans les discours techniques peut-elle s'arrêter aux seuls noms-termes qui y sont employés? »	39

L'apport du TAL à l'étude et à la didactique des langues

Victoria BOBICEV, Liviu Carcea « L'apport du TAL dans l'enseignement et l'apprentissage des langues »	59
Georges ANTONIADIS , Les acquêts et les écueils de l'intégration du TAL aux systèmes d'ALAO : L'exemple du projet MIRTO	77

L'enseignement des langues de spécialité

Lorena DEDJA, Esmeralda KROMIDHA « La conception des formations en fonction des besoins linguistiques du marché de travail (le domaine des assurances) »	103
Alba FRASHERI, Ardiana HYSO (KASTRATI) « Pour un enseignement équilibré entre FOS et FLE dans nos universités »	119
Rodica STANCIU-CAPOTĂ « Les enjeux de l'enseignement des langues de spécialité à l'université. Étude de cas : le(s) français de l'économie » .	133
Sonia ZAOUALI « Les écrits disciplinaires du domaine informatique : le cas des étudiants du Mastère à l'Institut Supérieur d'Informatique de Tunis »	141
Maria Ana OPRESCU « L'interculturel dans le cours de FLE »	163

Respect pentru oameni și cărți

Regina JUSKIENE, Olivia RUSU, Nicolas SAMSON « La traduction comme exercice pédagogique dans l'enseignement des technolèctes »	181
Corina VELEANU « Le français et l'anglais juridique à l'épreuve de la coopération internationale »	191
Les auteurs	209

Pour une utilisation des outils informatiques et du TAL dans l'analyse de textes techniques : la plate-forme du français technique et des technologies de la langue

Roxana Anca Trofin

Université Politehnica de Bucarest

Introduction

Depuis plusieurs décennies l'informatique traverse les frontières des disciplines de plus en plus variées fournissant des outils précieux pour les analyses spécialisées. La linguistique et la didactique des langues n'en font pas exception. Le Tal est devenu actuellement une réalité incontournable dans le processus d'enseignement apprentissage des langues. Considérant la langue comme un code, décomposable en informations susceptibles d'être identifiées et analysées, le TAL permet de traiter la langue à plusieurs paliers. Il entretient en même temps une relation complexe avec la linguistique située en amont et en aval du travail basé sur le TAL dans la mesure où pour extraire des informations il faut formuler des requêtes linguistiques exactes capables d'être intégrées dans les logiciels et dans la mesure où les informations fournies par l'analyse n'acquièrent de valeur qu'intégrées dans une approche linguistique et/ou didactique. Le TAL quelque incitant qu'il paraisse n'a pas grand intérêt en soi, les outils informatiques sont en revanche un instrument très utile dans certains types d'analyses linguistiques et dans l'apprentissage et/ou enseignement des langues. De par la fiabilité des informations fournies sur la langue et de par son caractère granulaire dans le sens qu'il permet le traitement d'éléments minimaux susceptibles par la suite d'être intégrés dans des ensembles plus complexes il est utile aussi bien aux apprenants qu'aux enseignants. Si les premiers peuvent grâce aux exercices générés à l'aide des instruments TAL faire une acquisition heuristique des

connaissances et procéder à une évaluation-validation des savoirs acquis les enseignants peuvent développer des activités et les intégrer dans des scenarii didactiques.

La place toujours plus importante qu'occupe le TAL dans le processus d'enseignement-apprentissage des langues, les utilisations que l'on peut en faire dans le développement des compétences non seulement linguistiques mais aussi langagières des apprenants ainsi que la nécessité de former de futurs spécialistes dans ce domaine nous ont menés à développer le projet *Pour le renforcement de la recherche en vue du développement des formations francophones niveau master - La Plate-forme du français technique et des technologies de la langue*.

Je m'arrêterai dans cet article aux principes qui nous ont guidés dans la mise en place du projet ainsi qu'aux enjeux du projet et de manière particulière à l'exploitation des outils informatiques dans l'étude du texte technique avec une focalisation sur la dimension épistémologique et l'analyse textuelle.

Le projet La plate-forme des technologies de la langue et du français technique

Le projet mené par un consortium d'universités regroupant l'Université Politehnica de Bucarest, l'Université Stendhal de Grenoble, l'Université Technique de Moldavie, l'Université de Tirana, la Nouvelle Université Bulgare est issu des besoins de formation en industries de la langue ainsi que des besoins d'analyse des textes de spécialité en l'occurrence textes techniques.

Il a eu comme principaux objectifs de créer un instrument capable de favoriser la mutualisation des connaissances et des informations et d'entraîner les enseignants chercheurs et les étudiants niveau master et/ou doctorat dans la recherche sur la langue et son traitement par les machines. Nous avons cherché de mener une réflexion d'une part sur le développement et l'exploitation pédagogique d'outils spécifiques basés sur le TAL et d'autre part sur la spécificité du texte technique et son exploitation pédagogique mais

également de concrétiser cette réflexion dans la création d'une plate-forme afin de mettre à disposition des étudiants et des enseignants des activités pédagogiques et des notions de TAL. Ainsi la plate-forme comprend-elle la génération et la correction automatique d'exercices, offre la possibilité d'exploiter tout texte proposé par les enseignants sans traitement préalable, la possibilité d'exporter les exercices créés afin de les utiliser en présentiel ou à distance sans recourir à la plate-forme.

Genèse du projet :

Partant de la prémisse que la méthodologie est conditionnée par le contexte dans lequel elle a été proposée et pour lequel elle est proposée (Puren) nous avons dans un premier temps établi le public cible. Ainsi la plate-forme orientée utilisateur s'adresse-t-elle aussi bien :

- * aux étudiants en FLE, niveau B2 ou C1 qui pourront perfectionner leur niveau de langue à travers les activités proposées et les fiches pédagogiques ;

- * aux étudiants en industries de la langue qui pourront approfondir les connaissances dans le domaine du traitement de la langue à travers la prise de contact avec les scripts de création d'exercices et les activités basées sur des programmes informatiques ;

- * aux enseignants en FLE qui souhaitent travailler sur le français technique.

En même temps compte tenu du fait que les étudiants en FLE et les enseignants FLE ont en principe peu de connaissances informatiques la plate-forme permet d'accéder à des activités élaborées en exploitant des outils informatiques sans nécessité d'une formation renforcée dans ce domaine.

Situations didactiques

Elle pourra être exploitée en classe ou en auto-apprentissage tuteuré vu qu'elle offrira le corrigé des exercices

Prémises théoriques : linguistiques, didactiques, informatiques

Le principe didactique qui a sous-tendu la création de la plate-forme est la centration sur l'apprenant. Pilier de l'approche communicative, je fais référence ici aux thèses de Besse, Porcher 1995 Galisson et Puren¹, la centration sur l'apprenant considéré plus récemment dans l'approche actionnelle comme acteur social à même de réaliser des tâches et d'interagir efficacement est destinée à mener à une autonomisation de celui-ci. Nous avons donc cherché de mettre l'apprenant en situation d'exploiter la langue faisant appel aussi bien aux compétences grammaticales et lexicales qu'aux compétences discursives et sociolinguistiques.

La théorie actionnelle postule que les actes de parole ne remplissent pleinement leurs fonctions qu'à travers l'action. Ainsi la réussite communicationnelle se fonde-t-elle sur la maîtrise linguistique de la langue qui inclut les connaissances lexicales et grammaticales (morphosyntaxiques, sémantiques, orthographiques et phonologiques) mais également sur la maîtrise des composantes sociolinguistiques et pragmatiques ainsi que sur la maîtrise d'une série de compétences telles les savoir-faire, les savoir-être, les savoir-apprendre car il ne s'agit plus d'« agir sur » mais d'« agir en accord avec »² (Bourguignon, 2007). Toutes ces compétences contribuent à la réussite de l'action de l'usager-acteur social. Les activités proposées sur la plate-forme visent donc à fournir à l'apprenant des informations théoriques à travers les fiches pédagogiques, l'amènent à travailler sur la grammaire, le vocabulaire et le texte en tant que objet d'un processus de schématisation.

-
- 1 Voir PUREN 1995a. « La problématique de la centration sur l'apprenant en contexte scolaire », *Études de Linguistique Appliquée* n° 100, oct.-déc. 1995, pp. 129-149. Paris : Klincksieck.
 - 2 BOURGUIGNON C, « Apprendre et enseigner les langues dans la perspective actionnelle : le scénario d'apprentissage -action » in *Les Langues Modernes*, 10 mars 2007, <http://www.aplv-languesmodernes.org> consulté le 3 mai 2013

Il conviendrait que je précise à ce stade de la communication les prémisses théoriques de ma démarche. Dans la perspective de l'analyse du discours et je me place dans le sillage de l'école française de l'analyse du discours pour produire un discours adapté à la situation de communication il faut satisfaire aux postulats de la situation, des préconstruits culturels, de finalité et de représentation formulés par Grize. En même temps en produisant un discours ou un texte le locuteur-producteur opère presque automatiquement un choix générique préalable car ainsi que le soulignait Bakhtine dans *Esthétique de la création verbale* «Lorsque nous choisissons un type donné de proposition, nous ne choisissons pas seulement une proposition donnée, en fonction de ce que nous voulons exprimer à l'aide de cette proposition, nous sélectionnons un type de proposition en fonction de tout l'énoncé fini qui se présente à notre imagination verbale et qui détermine notre opinion. L'idée que nous avons de la forme de notre énoncé, c'est-à-dire d'un genre précis de la parole, nous guide dans notre processus discursif.»³

Je considère le discours comme résultat d'une opération de schématisation « une production linguistique formant avec ses conditions de production socio-idéologiques un tout accessible à la description. », (Adam) et comme l'horizon de la langue (Saussure, Benveniste). Ainsi son étude suppose-t-elle l'analyse des conditions de sa production. Le discours est ainsi porteur d'une signification référée aux circonstances de communication alors que le sens des énoncés est défini en dehors de tout cadre énonciatif (Oswald Ducrot et Patrick Charaudeau) le producteur du discours étant un locuteur. Je reviendrai plus loin sur le conditionnement du discours scientifique opéré par le cadre de sa production.

Le texte est lui objet empirique et tout à la fois objet de la réception. Conditionné par deux types de liens d'une part « l'environnement extralinguistique : contexte ou situation d'interaction sociodiscursive, les situations d'énonciation et d'interprétation décalées ou non dans le temps et/ou dans l'espace »

3 BAKHTINE M *Esthétique de la création verbale*, Paris, Gallimard, 1984, p. 288

et d'autre part « l'environnement linguistique immédiat : co-texte et savoirs construits linguistiquement par le texte » (Adam), il porte les marques du cadre épistémologique de production, de l'intentionnalité de son producteur, de la position que celui-ci assume au moment de la production

De par sa fonction communicative (Rastier) le texte s'inscrit dans une pratique sociale.⁴

Or la plate-forme offre la possibilité comme on le verra plus loin de développer non seulement les compétences linguistiques mais de travailler les composantes pragmatiques et discursives des textes.

Nous avons poursuivi un renforcement linguistique, niveaux B2 ou C1, pour les étudiants utilisant la plate-forme mais également le développement des compétences de compréhension et d'expression écrites, l'enrichissement du vocabulaire technique et la sensibilisation des étudiants aux termes techniques en pratiquant le renforcement-fixation par des exercices d'association entre le terme et la définition, des QCM, la réutilisation des termes (dans le sens donné par Lérat) dans des contextes différents.

Du point de vue linguistique notre démarche est fondée sur la nécessité de traiter des textes techniques longs (articles de spécialité, mémoires, thèse) ainsi que sur la spécificité discursive et textuelle des textes techniques en fonction de leur type, du genre de discours, du public visé, sur la spécificité des collocations dans un texte technique avec un accent mis sur les contraintes de sélection.

Du point de vue informatique on a cherché de développer des instruments basés sur le TAL propres aux objectifs didactiques et de créer de scripts.

Nous avons créé la plate-forme sous forme de site Internet pour qu'elle soit accessible au large public. Elle a une section indications bibliographiques destinées à orienter les étudiants dans

4 Nous suivons les définitions de François RASTIER, *Sens et textualité*, Paris, Hachette, 1989 et « Pour une sémantique des textes » *Cahiers de l'ILSL* 6, 1995, p. 183-212.

la recherche documentaire en linguistique, informatique et TAL, une section regroupant des textes techniques et des activités, une section regroupant des fiches pédagogiques. En même temps les étudiants en industries de la langue pourront à travers un nom d'utilisateur et un mot de passe accéder à une plate-forme basée sur le TAL incluant des programmes capables de générer des exercices, créée à l'université Stendhal dans le cadre d'une thèse de doctorat réalisée par Abdelkarim Mars sous la direction de Georges Antoniadis. Cette section n'est pas accessible à tout public car le développement informatique en TAL n'est en général pas visible en libre accès, faisant l'objet des activités de recherche. Cet instrument basé sur le TAL permet à l'utilisateur de formuler la requête en fonction de l'objectif didactique poursuivi. Si par exemple il souhaite travailler les temps grammaticaux il indiquera le texte et le type de script par exemple exercice à trous, il lui suffit par la suite d'indiquer les catégories à enlever et les formes : verbes, formes du passé composé. La machine enlèvera automatiquement les formes du texte et fournira après la validation. On peut réaliser un exercice semblable si l'on veut travailler l'aspect sémantique en se basant sur les contraintes de sélection.

Tout en ouvrant des voies intéressantes capables de dynamiser l'enseignement apprentissage des langues et d'offrir à l'étudiant une certaine autonomie du parcours d'apprentissage le traitement automatique de la langue a des limites dans la mesure où ainsi que le montre Antoniadis et autres il est tributaire à la rigidité des logiciels qui ne traitent que les données prédéfinies ne pouvant pas détecter l'écart par rapport à la règle introduite⁵

5 Voir Georges ANTONIADIS, Olivier KRIAF, Thomas LEBARBE, Claude PONTON et Sandra ECHINARD, Modélisation de l'intégration de ressources TAL pour l'apprentissage des langues: la plate-forme MIRTO, Alsic Apprentissage des langues et systèmes d'information et de communication, vol. 8 n° 2 2005 : spécial ATALA

Le texte technique et sa spécificité

Le texte, objet matériel, concret est un énoncé complet résultat d'une opération d'énonciation, analysable selon des critères contextuels de structure et texture mais aussi en vertu des catégories d'intentionnalité et d'acceptabilité (Adam). Ainsi que l'a montré Ali Bouacha l'étude des textes apparaît comme un sous-ensemble des pratiques discursives dans la mesure où le texte ne peut être appréhendé qu'à travers une activité discursive et dans la mesure où il se présente comme un ensemble de données linguistiques qu'il faut traiter en données discursives.

Le problème qui se pose est si on peut vraiment parler d'une spécificité des textes techniques. Plusieurs aspects sont à prendre en considération un premier a trait à la généricité. Appartenant au régime vériconditionnel car renvoyant à une référence concrète relevant de la réalité factuelle le texte technique appartient aux genres définis par Maingueneau comme routiniers (1999), « de régime institué » (2004) dans le sens qu'ils supposent des scripts et des rôles relativement stables ne laissant pas beaucoup de place à la créativité. Le discours suit un protocole prédéfini au sein duquel le producteur est le détenteur du savoir qu'il transmet au lecteur, les rapports entre les deux variants selon le type de texte et de stratégie mise en œuvre. Dans les textes scientifiques tels articles de spécialité, thèses le producteur et le destinataire ont une connaissance unitaire et partagée du domaine, de la référence. Ces textes véhiculent des connaissances destinées à modifier l'épistémè. En revanche dans les textes didactiques et les textes de vulgarisation il existe un décalage de savoir entre le producteur et le destinataire.

En rapport avec le genre la problématique de l'énonciation présente elle aussi des particularités. Dans les textes scientifiques le locuteur assume et valide par son assertion le contenu, il se porte garant de sa vérité. Je considère le locuteur suivant le modèle de Benveniste et Ducrot comme l'instance productrice du discours ancré

temporellement et historiquement et qui oriente le discours selon ses choix épistémologiques, affectifs et conformément aux buts exprimés ou implicites. Le locuteur est ainsi la source énonciative sans coïncider nécessairement avec l'instance d'écriture ou l'auteur du texte, il est ainsi que le montre Ducrot « un être qui dans le sens même de l'énoncé est présenté comme responsable, c'est-à-dire comme quelqu'un à qui l'on doit imputer la responsabilité de cet énoncé. C'est à lui, à qui réfèrent le pronom « je » et les autres marques de la première personne. »⁶ Il est ce que Ducrot appelle un être de discours qui dans le cas des textes de spécialité, techniques en l'occurrence se superpose le plus souvent au sujet parlant (« l'être empirique personne complète qui possède entre autres propriétés celle d'être l'origine de l'énoncé »).

L'énonciation scientifique est considérée en général neutre car objective ne nécessitant pas d'ajustement intersubjectif vu qu'elle repose sur une même représentation du contenu. Néanmoins elle est souvent polyphonique ainsi qu'on le verra plus loin. Et si j'admets avec Kocourek que dans la langue de spécialité et surtout dans la langue scientifique et technique, l'écart entre l'intention de l'émetteur du message et la compréhension de destinataire est minimum et que la communication scientifique échappe à l'interprétation plurielle car elle véhicule des concepts qui en plus, sont exprimés par d'autres systèmes sémiotiques (les ensembles signifiants de spécialité, non linéaires : modèles ou dessins, les langages artificiels de spécialité) et a comme principale fonction la fonction cognitive (référentielle), si je partage en même temps la thèse de Jacobi selon laquelle dans les textes appartenant au registre scientifique⁷ la polyphonie est moins évidente que dans les textes de vulgarisation, je ne peux toutefois pas ne pas constater après étude de corpus que la communication scientifique renferme une dramatisation de l'énonciation et qu'elle est souvent polyphonique .

6 DUCROT O., *Le dire et le dit*, Paris, Éditions du Seuil, 1984, p. 175.

7 JACOBI Daniel, *La communication scientifique Discours, figures, modèles*, Presses universitaires de Grenoble, 1999, Grenoble, p. 22.