

« Former des enseignants de FLE en Colombie à travers un dispositif informatique fondé sur la linguistique textuelle ».

Molina Mejia, Jorge Mauricio y Antoniadis, Georges.

Cita:

Molina Mejia, Jorge Mauricio y Antoniadis, Georges (2016). « *Former des enseignants de FLE en Colombie à travers un dispositif informatique fondé sur la linguistique textuelle* ». Actes du colloque « L'enseignement du français à l'ère du numérique », 1, 310-326.

Dirección estable: <https://www.aacademica.org/jorge.mauricio.molina.mejia/27>

ARK: <https://n2t.net/ark:/13683/pqc6/3kr>



Esta obra está bajo una licencia de Creative Commons.
Para ver una copia de esta licencia, visite
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es>.

Acta Académica es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto. Acta Académica fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: <https://www.aacademica.org>.

Actes du 12^e colloque de l'AiRDF

L'enseignement du français à l'ère informatique

Colloque tenu du 29 au 31 août 2013
à la Haute école pédagogique du canton
de Vaud (Lausanne)

Édités par Mathieu Depeursinge, Sonya Florey,
Noël Cordonier, avec la collaboration de Sandrine Aeby
Daghé et Jean-François de Pietro

Soutiens

Association suisse de linguistique appliquée (Vals-Asla)
Institut universitaire de formation des enseignants, Université de Genève
Faculté des lettres, Université de Lausanne
Faculté de psychologie et des sciences de l'éducation (FAPSE), Université de Genève
Fondation Aebli-Näf
Institut de recherche et de documentation pédagogique (IRDp)
Ville de Lausanne



L a u s a n n e

Actes du 12^e colloque de l'AiRDF

L'enseignement du français à l'ère informatique

Colloque tenu du 29 au 31 août 2013
à la Haute école pédagogique du canton
de Vaud (Lausanne)

Édités par Mathieu Depeursinge, Sonya Florey,
Noël Cordonier, avec la collaboration de Sandrine Aeby
Daghé et Jean-François de Pietro

Pour citer une communication de ces actes

[Nom], [Prénom]. (2016), [Titre de la communication]. Dans M. Depeursinge, S. Florey, N. Cordonier, S. Aeby Daghé, J.-F. de Pietro (éds.), Actes du colloque «L'enseignement du français à l'ère du numérique», 29, 30, 31 août 2013 [(pp. xx-xx)], Lausanne: Haute école pédagogique du canton de Vaud.

Cet ouvrage rassemble une partie des contributions au colloque «L'enseignement du français à l'ère informatique» qui a eu lieu du 29 au 31 août 2013 à la Haute école pédagogique du canton de Vaud (Lausanne).

Le colloque a été soutenu par la Haute école pédagogique du canton de Vaud, ainsi que par la Fondation Aebli-Näf, la FAPSE (Université de Genève), l'IRD, l'IUFE (Université de Genève), l'Université de Lausanne, la Vals-Asla et la Ville de Lausanne.

Sans oublier les participants-es au colloque.

Le Comité d'organisation tient à les remercier pour leur précieux soutien.

Le comité scientifique du colloque était composé de:

Bernard Baumberger	HEP Vaud
Pierre-François Coen	HEP Fribourg
Glaís Cordeiro	Université de Genève
Jean-Louis Dufays	Université catholique de Louvain
Jean-Louis Dumortier	Université de Liège
Catherine Dupuy	Université Montpellier 2
Daniel Elmiger	Université de Genève
Erick Falardeau	Université Laval
Carole Fisher	Université du Québec à Chicoutimi
Claudine Garcia-Debanco	Université Toulouse 2
Nathalie Lacelle	Université du Québec à Trois-Rivières
Marlène Lebrun	HEP BEJUNE
Jean-Paul Mabillard	HEP Valais
Véronique Marmy	HEP Fribourg
Caroline Masseron	Université de Lorraine
Florence Quinche	HEP Vaud
Yves Reuter	Université Lille 3
Christophe Ronveaux	Université de Genève
Bruno Védrières	Université de Genève
David Vrydaghs	Université de Namur
Yann Vuillet	HEP Valais

D'autres personnes ont contribué à l'organisation du colloque: Marie Cantoni-Uldry (organisation des événements), Boris Martin (réservation des salles), Philippe Garbillo, Davide Sbriga et Davide Ziccarelli, (intendance), Carmine Buompreda (informatique et médias), Anouk Zbinden et Bertrand Mure (mise en ligne des contenus) et Thomas Zoller (confection des documents), Jan Olof Strinning.

L'édition des présents actes a bénéficié du concours de Barbara Fournier (coordination de la publication), Anca Mérot (suivi et mise en forme des contenus), Fabrizio Rossi (Ross graphic design, conception de la maquette et réalisation de la mise en page) et Sonia Rihs (relecture et correction).

Qu'ils soient tous remerciés chaleureusement.

Table des matières

Avant-propos	10
 Partie 1: Lire-écrire	
Utilisation des TIC hors de l'école et compréhension de texte français au secondaire 2	19
BERNARD BAUMBERGER	
Les cartes mentales, un outil de connaissance des processus scripturaux à l'œuvre dans le cadre de l'écriture du commentaire littéraire	28
JEAN-YVES BOUTON	
Lire, comprendre, collecter des mots clés et rappeler le récit oralement en utilisant un média <i>De la compréhension au rappel de texte</i>	43
CLAUDE BURDET ET SONIA GUILLEMIN	
Contrôle orthographique en situation de production d'écrits <i>Le brouillon collaboratif au cycle 3 de l'école primaire française</i>	55
THIERRY GEOFFRE	
La lecture sur écran: ce que les enquêtes peuvent apprendre à l'enseignant de français	68
MONIQUE LEBRUN	
Intégrer les nouvelles technologies à l'enseignement du français au secondaire: le dispositif Radiobus sous la loupe, analyse des perceptions des enseignants utilisateurs	77
JOSÉ TICON ET DENIS BADAN	
Quel français enseigner à l'heure de l'informatique? Le défi est dans le tri	86
JORGE JUAN VEGA Y VEGA	
 Partie 2: Orthographe/structuration	
La synthèse vocale pour écrire	99
CLAUDE BEUCHER-MARSAL ET FLORENCE CHARLES	
Un modèle théorique d'articulation de l'enseignement de la grammaire et de l'écriture	108
MARIE-CLAUDE BOVIN ET REINE PINSONNEAULT	
Apprendre l'orthographe en autonomie avec le logiciel Progresser en orthographe, dictées codées: une expérimentation en CM2	121
CATHERINE BRISSAUD ET PATRICK LUYAT	

Partie 3 : Littérature

Corpus littéraires, entre tradition et nouveauté. Jeux et enjeux de la transmission des textes	133
SANDRINE AEBY DAGHÉ	
L'enseignement de la lecture des œuvres littéraires : portrait de pratiques en première année du collégial québécois	146
JULIE BABIN ET OLIVIER DEZUTTER	
Actualisation de l'enseignement de la littérature au secondaire II État des lieux en formation initiale	154
SYLVIE JEANNERET	
Forum et blogue pour le cours de littérature au lycée : entretenir en ligne une présence pédagogique	164
PIERRE MOINARD	
Développer des compétences communicationnelles en intégrant la robotique dans la mise en scène théâtrale	174
FLORENCE QUINCHE ET JOHN DIDIER	
Le commentaire littéraire avec Google Drive (anciennement Google documents): Les TICE (technologies de l'information et de la communication pour l'enseignement) au service de l'apprentissage collaboratif	183
VALÉRY RION	

Partie 4 : Français/langue étrangère

Lectures et écritures partagées sur la toile	199
CHIARA BEMPORAD ET CAMILLE VORGER	
Agir ensemble en contexte de projets d'échanges à distance à l'école primaire <i>Entre correspondance scolaire internationale et projets d'échanges à distance : continuum d'une pratique pédagogique</i>	210
ANNE CHOFFAT-DÜR	
Outils linguistiques en formation hybride : lesquels facilitent l'apprentissage?	219
DORA LOIZIDOU	
Système d'aide à la rédaction ScribPlus en classe de FLE Apports, limites et pistes d'amélioration	229
MONTYIA PHOUNGSUB	
Les supports numériques en Français Langue étrangère : un atout pédagogique? <i>Quels choix opérer parmi les offres du marché des TICE?</i>	241
MARIE-LAURE WULLEN	

Partie 5: Formation des enseignants

Les TIC dans l'enseignement du français en France: les difficultés d'un enseignement intégré en formation des enseignants	251
MAGALI BRUNEL	
Pratiques et usages du TBI au service de l'oral en cyberclasse primaire Exemples en Fédération Wallonie-Bruxelles (FWB)	262
FRANÇOISE MARIA CAPACCHI	
Quels usages numériques pour le cours de français en secondaire? Expérimentations à l'occasion du projet «École numérique»	272
FRANÇOISE CHATELAIN	
Analyse des effets du TBI sur les pratiques d'écriture et de réécriture des étudiants de Master «Professorat des écoles» De la «rédaction» à l'album typographique	285
DOMINIQUE DUVIN-PARMENTIER	
Savoir-faire communicationnels et communication des compétences en Master de spécialisation ès Lettres	301
GILLES MERMINOD ET MARCEL BURGER	
Former des enseignants de FLE en Colombie à travers un dispositif informatique fondé sur la linguistique textuelle	310
JORGE MAURICIO MOLINA MEJIA ET GEORGES ANTONIADIS	
Les outils dans la formation des enseignants de français sur la production écrite Quelle place pour les Technologies de l'Information et de la Communication pour l'Enseignement (TICE)?	327
VERÓNICA SÁNCHEZ ABCHI, CARLA MESSIAS RIBEIRO ET JOAQUIM DOLZ	

Former des enseignants de FLE en Colombie à travers un dispositif informatique fondé sur la linguistique textuelle

JORGE MAURICIO MOLINA MEJIA ET GEORGES ANTONIADIS

LIDILEM, Université Grenoble Alpes

Jorge-Mauricio.Molina-Mejia@e.u-grenoble3.fr

Georges.Antoniadis@u-grenoble3.fr

Résumé

Cet article montre les possibilités offertes par un système informatique de type hybride fondé sur la linguistique textuelle, dans le cadre de la formation des futurs enseignants de FLE en Colombie.

Mots clés:

dispositif informatique – linguistique textuelle – formation de formateurs didactique du FLE – traitement automatique des langues

1 Introduction

Si nous les comparons à d'autres domaines tels que la linguistique ou la didactique, la linguistique textuelle et l'informatique (notamment ses versants TAL et ALAO²⁸) sont des champs disciplinaires relativement récents. Malgré cette nouveauté disciplinaire, elles (chacune de leur côté) se sont assez bien positionnées ces dernières décennies. D'un côté nous pouvons dire qu'il existe des systèmes informatiques pour l'apprentissage/enseignement des langues fondés sur les procédures et les résultats du Traitement Automatique des Langues (TAL)²⁹ (Chaudiron, 2007; Antoniadis, 2010). D'autre part, plusieurs travaux mettent en avant l'apport de la linguistique textuelle dans l'enseignement et l'apprentissage de la langue française (Combettes, 1983; Chiss & David, 2012; Lundquist, 1990; 2013). Cependant, le terrain que nous abordons ici, celui de la formation au niveau universitaire des futurs formateurs au travers des didacticiels fondés sur des corpus issus de la linguistique textuelle, n'a pas été, à notre connaissance, vraiment exploré jusqu'ici³⁰.

Le système en élaboration, issu de notre démarche, est fondé sur la linguistique textuelle. Il vise à utiliser un corpus de textes créé et exploité à l'aide des techniques du TAL. Le système est destiné à la formation des formateurs en FLE au sein des universités colombiennes. Il a pour objectif l'amélioration du niveau linguistique des futurs enseignants de Français Langue Étrangère à travers les différentes approches de la linguistique textuelle, notamment l'analyse des phénomènes linguistiques à partir du contexte et du cotexte (Adam, 2011 a) et la notion de discursivité textuelle (Maingueneau, 1991; Pop, 2000; Adam, 2011a).

Projet inter et pluridisciplinaire, notre système se veut un outil d'aide à la formation. Il s'agit de la proposition d'un outil d'apprentissage à visée formative fondé sur la linguistique textuelle, un domaine de recherche où les travaux, et à fortiori les systèmes, sont relativement rares (Mangenot, 1998; Couto et al., 2005).

28 Apprentissage des Langues assisté par ordinateur.

29 En particulier pour la forme écrite des langues.

30 Notons néanmoins le projet du «Français en première ligne» développé par C. Develotte et F. Mangenot de 2002 à 2007, qui cherche à améliorer la performance des futurs enseignants de FLE à travers l'emploi des TICE.

2 Contexte de recherche : la formation des formateurs en Colombie.

En 2010, nous avons commencé à réfléchir autour d'un projet de formation des futurs enseignants de FLE. Nous avons pris pied sur notre expérience d'enseignement dans une université colombienne qui forme ce type d'enseignants. L'Université d'Antioquia (située à Medellín) qui possède un programme de formation des enseignants en langues étrangères (anglais-français) nous a servi comme terrain d'observation pendant plusieurs années. À partir de ces premières observations, nous avons proposé un projet qui a été financé par le «*Ministerio de Tecnologías de la Información y de la Comunicación de la República de Colombia*³¹». Peu de temps après le début de notre projet, une autre université a montré aussi de l'intérêt pour ce projet, l'université nationale de Colombie qui a son siège principal à Bogotá.

Notre travail au sein du laboratoire LIDILEM nous a permis de mieux cerner notre objet d'étude, et de donner naissance à notre projet qui, depuis l'année 2010, fait l'objet d'une thèse en Informatique et Sciences du Langage. Plusieurs échanges avec des universités en Colombie ont permis de mieux adapter notre projet aux particularités et exigences du public visé.

2.1 Problématique et hypothèses de travail

Notre problématique découle de la question que nous nous sommes posée : comment réduire l'insécurité linguistique en améliorant le niveau de compétence linguistique des futurs enseignants de FLE à travers un dispositif informatique ?

Nous partons de l'hypothèse qu'il y a une sorte d'insécurité linguistique chez les futurs enseignants qui finissent leurs études de FLE en Colombie. Nous considérons que cette insécurité pourrait être atténuée en utilisant un dispositif informatique spécifique pour la formation de la compétence linguistique. Selon Klinkenberg (1993), il y a insécurité linguistique lorsque l'individu a une image assez nette de la norme, mais que celui-ci n'est pas sûr d'avoir la maîtrise de cette variété légitime. « Il y a au contraire sécurité dans le cas où la production d'un usager est conforme à la norme qu'il reconnaît, et dans celui-ci où son usage n'est pas légitime, mais sans qu'il ait une conscience nette de la non-conformité. » (Klinkenberg, 1993: 185)

Cette norme dont parle cet auteur pourrait être expliquée selon E. Coseriu comme :

« La norme comprend tout ce qui, dans la "technique du discours", n'est pas nécessairement fonctionnel (distinctif), mais qui est tout de même traditionnellement (socialement) fixé, qui est usage commun et courant de la communauté linguistique. Le système, par contre, comprend tout ce qui est objectivement fonctionnel (distinctif). La norme correspond à peu près à la langue en tant qu'"institution sociale"; le système est la langue en tant qu'ensemble de fonctions distinctives (structures oppositionnelles). Comme corollaire, la norme est un ensemble formalisé de réalisations traditionnelles; elle comprend ce qui "existe" déjà, ce qui se trouve réalisé dans la tradition linguistique; le système, par contre, est un ensemble de possibilités de réalisation: il comprend aussi ce qui n'a pas été réalisé, mais qui est virtuellement existant, ce qui est "possible", c'est-à-dire ce qui peut être créé selon les règles fonctionnelles de la langue. »

(Coseriu, 2001: 246-247)

31 Ministère des Technologies de l'information et de la communication de la République de Colombie (MinTic).
Site: <http://www.mintic.gov.co/index.php>

Nous considérons qu'un formateur se doit non seulement de connaître la norme d'usage, mais aussi d'être en capacité d'adapter son répertoire langagier et le métalangage auquel il a recours pour expliquer les faits de langue aux apprenants. Nous supposons que ceci pourrait être fait grâce à notre système informatique de formation et de la linguistique textuelle.

2.2 Méthodologie de recherche

Nous avons cherché à valider nos hypothèses par le biais d'entretiens sur place et d'un questionnaire en ligne. Au total 172 personnes appartenant à 12 des 18 institutions d'éducation supérieure ont répondu à notre questionnaire. Celui-ci s'adressait à trois publics différents :

1. Enseignants formateurs (EnsForm) : 12,8% des enquêtés.
2. Étudiants en formation (EtudForm) : 72,7% des enquêtés.
3. Enseignants de FLE (EnsFLE) ayant subi une formation dans une université colombienne : 14,5% des enquêtés.

Le questionnaire était divisé en quatre sections à savoir :

2.2.1 Aspects concernant la formation didactique

Par rapport à la formation didactique, nous cherchions à savoir quels aspects didactiques étaient privilégiés dans le processus de formation des futurs enseignants de FLE. Comme nous pouvons l'observer (figure 1), il apparaît que les aspects privilégiés sont la didactique de l'oral, suivie de la didactique de l'écrit. Ceci va dans le sens des réponses à la question concernant la méthodologie d'enseignement des langues la plus utilisée au niveau de la formation, puisque 94,7% des enseignants formateurs, 77% des étudiants en formation et 91,7% des enseignants de FLE déclarent que l'approche communicative est la méthode d'enseignement la plus utilisée au niveau de la formation. Cette dernière est également la plus recommandée par les enseignants formateurs.

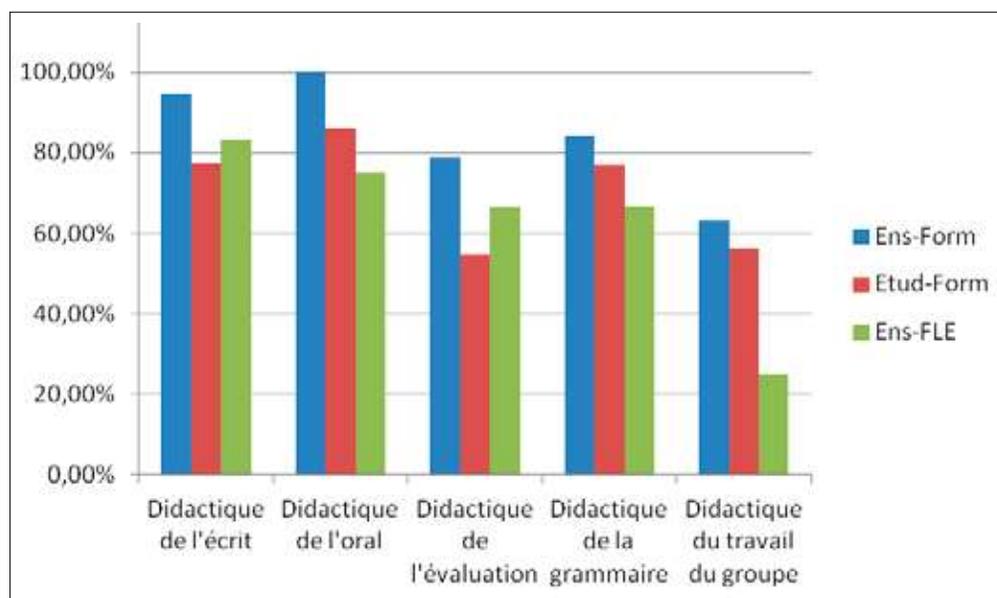


FIGURE 1 Aspects privilégiés de la formation didactique (Molina Mejia & Antoniadis, 2014b).

2.2.2 Aspects concernant la formation à la compétence linguistique

Quant à la formation à la compétence linguistique (figure 2), selon le CECR, elle est formée de cinq sous-compétences (lexicale, grammaticale, sémantique, phonologique et orthographique). Nous avons voulu savoir lesquelles étaient privilégiées dans la formation des enseignants. Nous avons trouvé que la sous-compétence privilégiée lors de la formation des futurs enseignants de FLE est la compétence grammaticale, suivie de la compétence sémantique. Les compétences les plus délaissées concerneraient la phonologie et l'orthographe. Nous voyons dans cette démarche l'idée de vouloir pallier un certain manque de connaissances linguistiques, dû à l'approche communicative utilisée principalement pour les enseignements de ce public, cette approche a un regard particulier sur la grammaire, la considérant comme un outil subalterne au service des compétences (Halté, 2004), voire parfois en remettant en question son utilisation (Beacco, 2010).

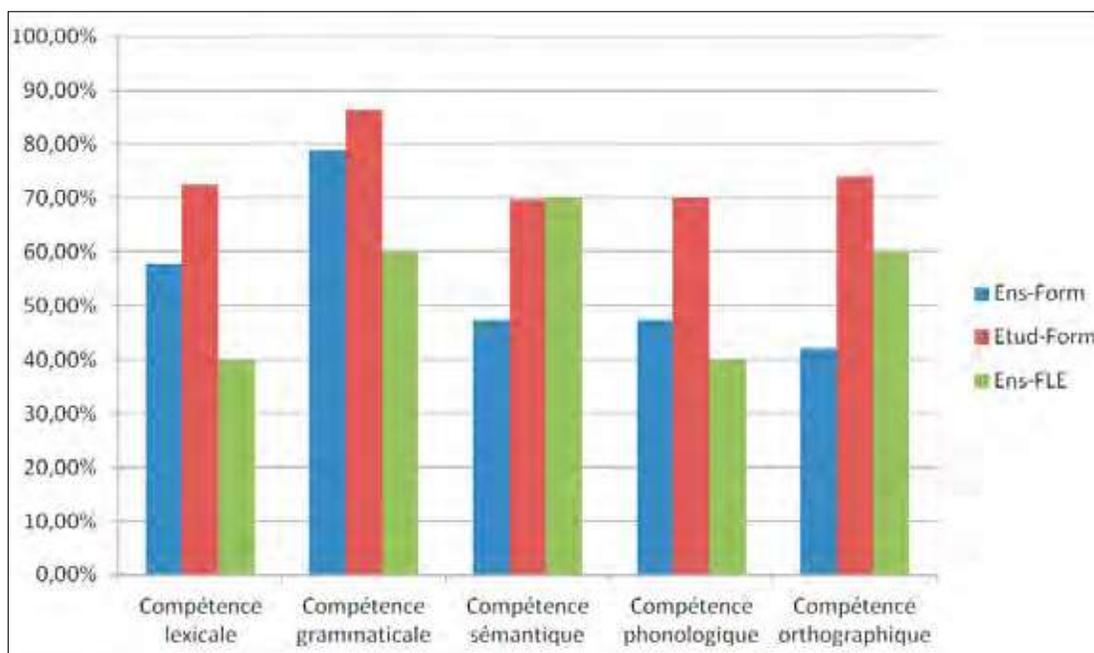


FIGURE 2 Aspects privilégiés de la formation linguistique (Molina Mejia & Antoniadis, 2014b).

2.2.3 Aspects concernant la linguistique textuelle

Du côté de la linguistique textuelle (figure 3), alors que les enseignants formateurs disent la connaître et l'utiliser pour leurs cours, leurs étudiants déclarent ignorer cette approche méthodologique. Pour les trois publics, les aspects qui pourraient être améliorés à travers la linguistique textuelle sont la cohérence textuelle, la structuration textuelle et les éléments constitutants du texte.

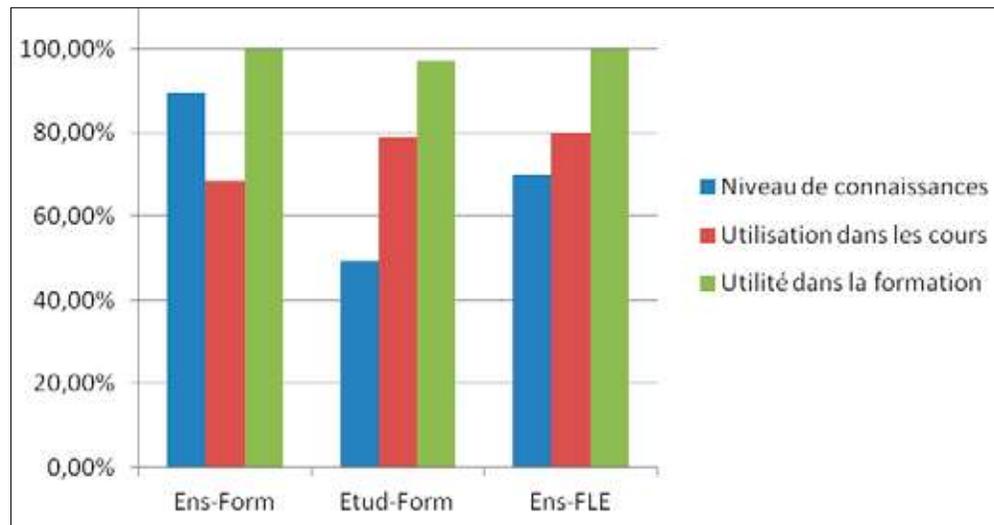


FIGURE 3 Utilisation et utilité de la linguistique textuelle dans la formation (Molina Mejia & Antoniadis, 2014b).

2.2.4 Emploi des TICE dans la formation des futurs enseignants de FLE

Enfin, en ce qui concerne l'emploi des nouvelles technologies pour l'enseignement, nous avons pu constater que ce sont les blogs pédagogiques et les sites et plateformes e-learning qui sont les plus utilisés dans la formation. Elles sont utilisées pour deux aspects, les exercices de compréhension orale et l'étude de la grammaire. Les trois publics pensent que l'utilisation des outils informatiques joue un rôle essentiel dans la formation des futurs enseignants de FLE. Ils trouvent qu'un projet informatique à caractère formatif utilisant les TICE et la linguistique textuelle pourrait pallier les problèmes de formation des aspects linguistiques.

3 Les fondements du projet

3.1 Améliorer le niveau de la compétence linguistique à travers la linguistique textuelle

Des trois grandes composantes (linguistique, sociolinguistique et pragmatique) qui permettent à un individu de communiquer langagièrement, d'après le CECR (Conseil de l'Europe, 2000), nous avons choisi la compétence linguistique. Ce choix n'a pas été anodin, étant donné que, d'après l'analyse de nos données (cf. figure 2), un des problèmes principaux au niveau de la formation est lié au manque de confiance des étudiants en formation à bien manier des concepts linguistiques. Ceci se traduit justement par l'insécurité linguistique dont nous avons parlé auparavant. Le CECR définit la compétence linguistique comme :

«[...] celle qui a trait aux savoirs et savoir-faire relatifs au lexique, à la phonétique, à la syntaxe et aux autres dimensions du système d'une langue, pris en tant que tel, indépendamment de la valeur sociolinguistique de ses variations et des fonctions pragmatiques de ses réalisations.»

(Conseil de l'Europe, 2000: 17)

Compte tenu de cette définition, nous nous intéressons notamment au rapport entre la compétence linguistique et son apport à la formation des futurs enseignants de FLE par le biais de la linguistique textuelle. Nous nous fondons sur l'idée que notre public cible (les futurs enseignants) sera confronté à l'utilisation d'un savoir-faire à la fois linguistique et métalinguistique. C'est-à-dire qu'ils devront savoir expliquer à des apprenants de français comment fonctionne le système de la langue. Que ce soit en s'appuyant sur l'utilisation du métalangage linguistique ou non, une connaissance des sous-compétences liées à cette compétence linguistique est nécessaire. Nous pensons que l'on peut mieux faire apprendre la langue lorsqu'on comprend mieux son fonctionnement interne, ainsi que sa forme.

C'est justement à travers la linguistique textuelle que nous voulons explorer cette forme interne de la langue. La linguistique textuelle se situe historiquement à la fin des années 1960 (Charaudeau & Maingueneau, 2002); d'autres auteurs (Adam, 2011a) la situent au milieu des années 1950 (travaux de E. Coeriu publiés en 1969). Ce sera surtout dans les années 1990 et 2000 que cette approche théorique commence vraiment à se positionner, grâce notamment aux travaux de Jean-Michel Adam (1990, 2011a, 2011 b), et à ceux de Lita Lundquist (1990, 2013), travaux qui cherchent à montrer de quelle manière la linguistique textuelle peut être utilisée comme un outil d'aide dans l'analyse du discours et comme une aide pour la didactisation des phénomènes issus de cette approche théorique.

Nous nous sommes intéressés à la linguistique textuelle comme approche théorique d'aide à la formation, car celle-ci ne se présente pas comme une simple théorie de la phrase étendue au texte, mais plutôt comme une «translinguistique» (Bakhtine-Todorov, 1981; Benveniste, 1974). Nous pensons, comme Charaudeau et Maingueneau (2002), que l'étude de la langue doit rendre compte des phénomènes de cohérence et de cohésion au niveau textuel. Cela change l'optique de la phrase comme outil de travail et de son analyse structuraliste qui ne tient compte ni du contexte ni des aspects énonciatifs qu'elle peut porter lorsqu'elle est liée à un cotexte.

C'est en partant des textes et de leur caractère «discursif» (Adam, 2011a) que nous voulons améliorer le niveau de compétence linguistique des futurs enseignants de FLE colombiens. Il faut cependant noter que cette approche théorique n'est pas très connue ni très répandue au niveau de la formation des futurs enseignants de FLE dans le contexte colombien³². Il y a encore une assez forte tendance à faire apprendre la langue, et donc à se former, par le biais de l'approche communicative. Selon certains auteurs, cette approche théorique est trop complexe à mettre en œuvre (Puren, 1995 ; Demouguin, 2008).

En voulant utiliser la linguistique textuelle et les phénomènes déjà cités, nous voulons considérer la langue dans son intégralité, en analysant les caractères linguistiques au travers des textes. Nous cherchons, d'un côté, à faire apprendre la langue aux futurs enseignants et, d'un autre côté, leur apprendre à l'enseigner. Notre idée est de former les futurs enseignants de FLE à l'utilisation des phénomènes liés à la linguistique textuelle, mais aussi de les former en même temps à leur didactisation. Cela comporte un double processus cognitif: d'un côté, l'apprentissage de la langue, de l'autre, la transmission de celle-ci. D'après Chiss et David (2012: 166):

³² Ayant analysé les plans de formation de la plupart des universités colombiennes, nous avons constaté que la linguistique textuelle ne fait pas partie du curriculum de ces universités. Nous avons également, lors de notre enquête voulu vérifier son existence, mais à part l'utilisation du texte (en version papier) dans un but d'enseignement, aucune référence spécifique à cette approche théorique n'a été constatée.

«[...] ces règles de composition et d'organisation textuelles constituent des objectifs d'apprentissage importants, à inscrire dans une didactique d'ensemble du français. L'acquisition du langage écrit sous tous ses aspects passe par l'assimilation des procédés linguistiques décrits ici et la maîtrise des opérations langagières sous-jacentes.»

Ce que nous voudrions développer, chez les futurs enseignants de FLE dans le contexte colombien, est leur capacité à analyser le discours, par le biais des notions linguistiques, telles que la cohésion, la cohérence, la progression thématique, etc.

3.2 L'intérêt de disposer d'un outil informatique fondé sur la linguistique textuelle

«L'école n'aime pas tout ce qui pourrait se substituer à la parole du maître, tout ce qui pourrait laisser entendre qu'elle n'est pas l'unique dépositaire de la connaissance, tout ce qui pourrait lui dérober son image de mère nourricière, d'alma mater, en matière de connaissance.»

(Guir, 2002: 9)

Malgré ces mots de Roger Guir, nous pouvons dire qu'il existe, en didactique des langues, depuis plusieurs années déjà, un intérêt croissant pour l'utilisation des nouvelles technologies. Ces outils informatiques se sont centrés notamment sur l'enseignement/apprentissage de la langue vis-à-vis d'un public dit «d'apprenants» de langues étrangères. En FLE, il y a eu beaucoup d'efforts dans ce domaine. Dans le cas de l'aide à la formation des formateurs, il n'y a pas eu, à notre connaissance, un très grand nombre d'outils informatiques; moins encore d'outils informatiques avec un but précis, celui de la formation en linguistique/didactique des futurs enseignants. Dans ce deuxième cas, nous avons pu constater l'intérêt que ces outils suscitent au niveau de la formation en Colombie, bien qu'historiquement, la formation ait été très centrée sur la parole du maître et de son savoir-faire.

Notre idée est de montrer que bien que l'analyse des textes puisse se faire avec des textes papier – et cela se fait toujours au niveau de la formation des enseignants en Colombie – l'utilisation de l'outil informatique offre des possibilités pertinentes pour l'enseignement. Nous commencerons par exposer les possibilités que nous voyons pour l'outil informatique au travers d'un travail de constitution/structuration d'activités didactiques fondées sur un corpus. Un corpus exemple, annoté et traité selon les activités que nous voulons faire faire à notre public cible sera ensuite indiqué pour nous mener à la proposition d'activités bien ciblées en accord avec les besoins exprimés par les étudiants en formation. Finalement, nous aborderons la possibilité de disposer d'un outil qui permettrait aux enseignants formateurs de pouvoir faire apprendre des notions linguistiques tout en disposant du support théorique à la fois linguistique et didactique.

4 Un environnement informatique d'aide à la formation de formateurs

Après avoir défini notre cadre théorique et avoir pris en compte les besoins de nos publics, nous avons décidé de mettre en œuvre la conception d'un dispositif informatique à utiliser comme outil d'aide à la formation. Cet outil vise deux types d'utilisation: 1) en présentiel avec l'accompagnement de l'enseignant formateur; 2) en dehors de la salle de classe en autonomie, afin d'être utilisé pour renforcer les notions acquises dans le cours en présentiel.

Nous proposons un environnement hybride, dans le sens de Charlier et al. (2006: 481), c'est-à-dire un dispositif dans lequel «se conjuguent les dimensions innovantes liées à la mise à distance

et l'utilisation d'un environnement technopédagogique qui repose sur des formes complexes de médiatisation et de médiation».

4.1 Fonctionnalités et architecture du système informatique

La figure 4 définit l'architecture globale de notre système au travers de son découpage fonctionnel et du point de vue des données qu'il contient. Le dispositif général est divisé en deux interfaces, une pour les enseignants-formateurs et l'autre destinée aux étudiants en formation. Le système donne accès à deux types de ressources qui sont stockées dans deux bases de données; la première comprend le corpus textuel, nous l'avons appelé «BDD Textuelle», la deuxième, «BDD théorique», contient les explications théoriques d'ordre linguistique et didactique. Des séquences didactiques sont intégrées au système même, elles permettent d'aller chercher les textes en fonction des activités ou des explications théoriques liées à ces activités dans les différentes BDD.

Au niveau des fonctionnalités, les deux interfaces ont été dotées d'activités; l'interface enseignant vise à composer des séquences d'activités, celle pour les apprenants à les jouer. Ces activités sont divisées en quatre phases, et chaque phase est divisée en deux sous-phases; ceci permet d'aller d'une sous-phase avec des exercices dits d'initiation vers une deuxième sous-phase avec des exercices d'approfondissement. Certaines sous-phases contiennent du matériel théorique avec des explications d'ordre linguistique et didactique, correspondant aux notions étudiées. L'idée étant de former des futurs enseignants de FLE, nous nous sommes empressés de leur fournir le support théorique leur permettant de concevoir des activités pour les réutiliser avec leurs propres apprenants, autrement dit, nous leur fournissons les instruments pour qu'ils se forment à la didactisation des notions acquises. Notre système devrait permettre la conception et la création des séquences didactiques de la part de l'enseignant formateur. Pour ce faire, il dispose d'un ensemble d'outils informatiques et didactiques d'utilisation simple permettant la mise en place de ces séquences.

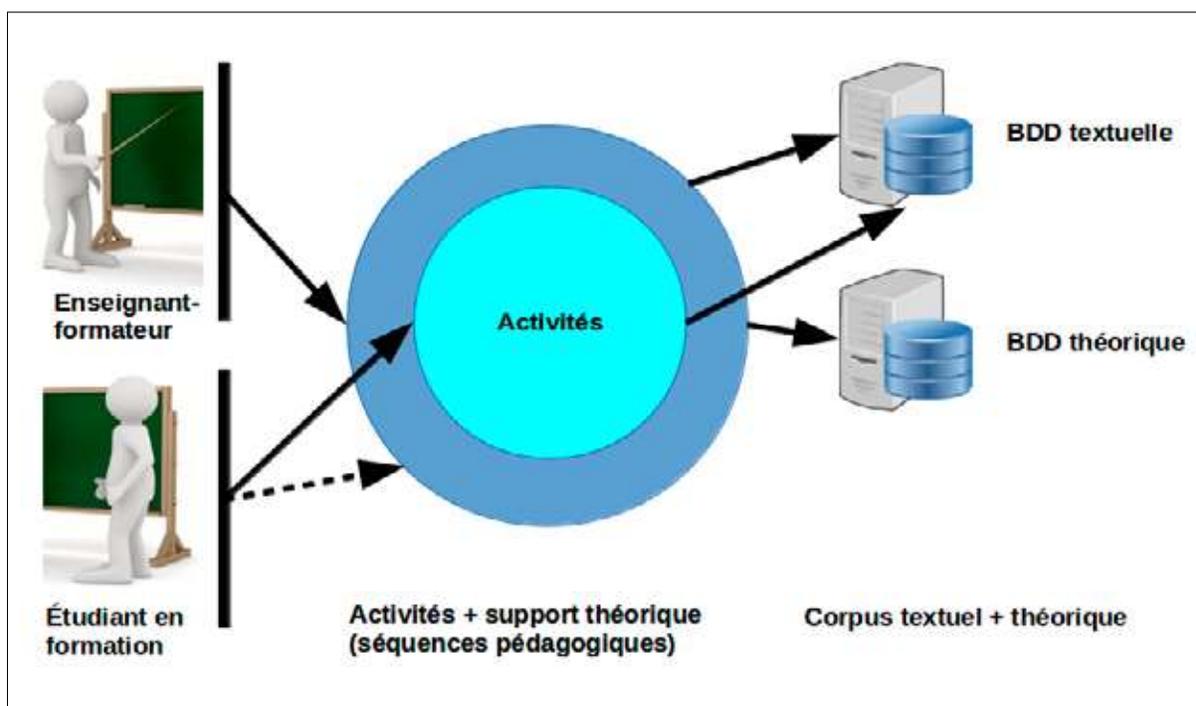


FIGURE 4 Architecture générale du système (Molina Mejia & Antoniadis, 2014a).

4.2 Constitution d'un corpus textuel

Pour pouvoir proposer des séquences didactiques, le système doit modéliser les phénomènes linguistiques à traiter :

«Le plus souvent, informaticiens et didacticiens des langues n'admettent pas la même acception du terme "langue". Pour les uns, il s'agit d'une séquence de codes, pour les autres d'un système de formes et de concepts. Cette différence peut s'expliquer aisément lorsque l'on constate que l'informatique [seule], par définition, ne peut considérer et ne peut traiter que la forme de la langue dénuée de toute interprétation, alors que pour la didactique des langues, la forme n'existe que par ses propriétés et des concepts qu'elle est censée représenter.»

(Antoniadis et al., 2005: 65)

C'est là l'enjeu de la constitution d'un corpus puisque ce sera sur ces annotations que se construira l'exploitation des textes dans le système.

4.2.1 Type de textes

Notre corpus «exemple» est un corpus de petite taille (environ 3800 mots) composé de textes journalistiques, appartenant aux niveaux B1, B2 et C1 du CECR. La caractéristique essentielle de ce corpus est qu'il s'agit de textes authentiques faisant partie d'anciens examens du DELF et du DALF qui apparaissent sur le site du CIEP³³. En choisissant ce type de textes, notre intention était, d'une part, la préparation des étudiants aux examens DELF et DALF, et d'autre part, l'amélioration de leur niveau de langue à l'aide de tâches fondées sur des documents authentiques, comme le préconise le CECR.

Dans une première étape, nous avons commencé par constituer un petit corpus de textes, notamment des extraits de journaux ou d'hebdomadaires, mais ceci n'interdit pas d'aller dans un second temps vers d'autres types de textes. Nous envisageons, dans une deuxième étape, d'élargir notre corpus textuel, de recueillir puis d'annoter des textes littéraires et de vulgarisation scientifique. Il faudra par contre que ces textes soient avérés de niveau B1, B2 et C1 (B2 et C1 sont les niveaux que les étudiants en formation doivent acquérir à la fin de leurs études).

4.2.2 Annotation des textes

Afin de pouvoir exploiter de manière informatique les informations portées par les textes, nous faisons deux types d'annotations (figure 5) : d'une part des annotations automatiques de type morphologique et syntaxique à l'aide de l'analyseur Cordial (Annexe A), d'autre part des annotations manuelles à travers le format XML (Annexe B). Une fois les textes annotés, ils sont déposés dans une base de données (BDD textuelle) et sont utilisés dans notre application.

Les annotations manuelles du corpus ont été élaborées suivant les trois axes de la linguistique textuelle que nous avons choisi de traiter pour cette phase de spécification : 1) structuration textuelle ; 2) cohérence et cohésion textuelles ; 3) progression thématique (cf. 4.3.2). Concernant les unités annotées, nous avons fait le choix d'annoter les textes au niveau de la proposition, sachant que toute phrase, en principe, est composée de propositions³⁴. Que ce soit des phrases dites com-

³³ Le CIEP (Centre International des Études Pédagogiques) est l'institution française chargée de la formation des enseignants et des évaluateurs en matière des examens internationaux en langue française (DELF, DALF, TCF). Site : <http://www.ciep.fr/>

³⁴ Même les phrases simples seraient constituées des propositions, celles que l'on nommerait des «propositions indépendantes», car selon Riegel et al. (2009: 784) «[...] elle [la proposition indépendante] ne dépend pas d'une autre proposition et aucune autre proposition ne dépend d'elle.»

posées ou simples, il y aura toujours une, deux ou plusieurs propositions. En effet, les propositions ne sont jamais isolées et elles correspondent (font référence) à d'autres propositions dans le texte et dans le discours (Adam, 2011a)³⁵.

Les unités textuelles (les connecteurs logico-temporels, argumentatifs et discursifs, certaines anaphores dans le cas des coréférents textuels) ne peuvent pas être repérées automatiquement. Nous nous sommes aperçus que les unités textuelles que nous avons signalées sont plus faciles à isoler et à identifier quand on les situe au niveau des relations établies entre des propositions, plutôt que lorsqu'on les analyse au niveau de la phrase.

Après avoir fini les annotations automatiques et l'étiquetage XML, nous procédons à leur vérification. Nous voulons nous assurer que l'étiquetage des unités textuelles soit précis et conforme et qu'il tienne compte des phénomènes associés à la linguistique textuelle que nous voulons faire étudier. À partir de ces annotations, nous avons réfléchi à la meilleure manière de les intégrer dans les activités et dans les explications d'ordre linguistique et didactique. Du fait de son affinité avec notre approche méthodologique, nous avons choisi comme grammaire de référence la «Grammaire méthodique du français» (Riegel *et al.*, 2009).

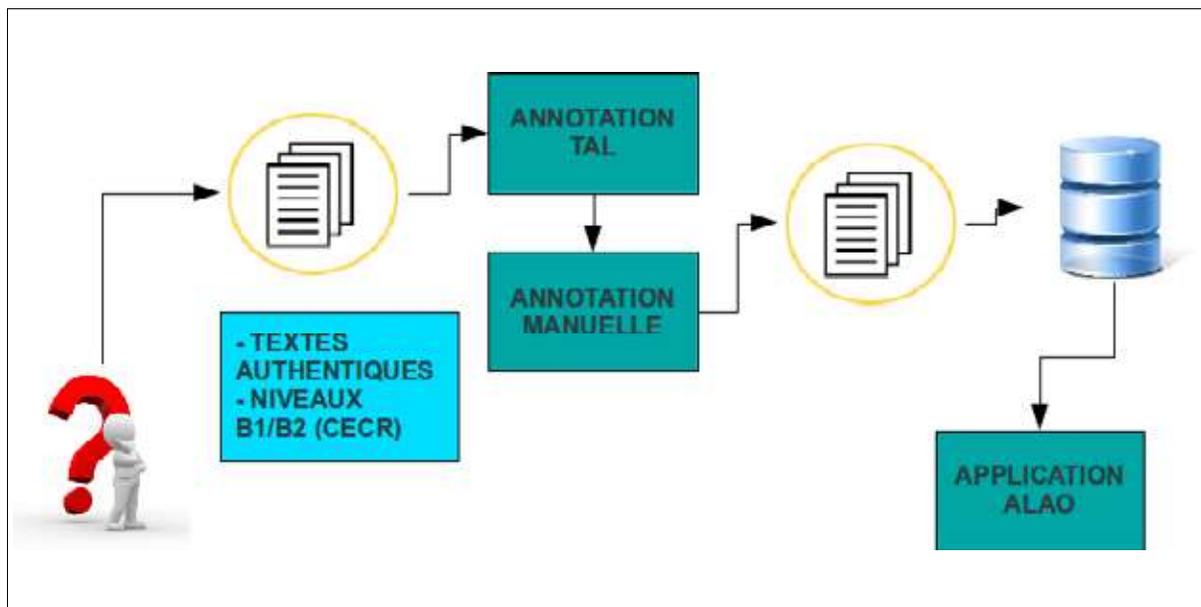


FIGURE 5 Schéma du traitement du corpus (Molina Mejia & Antoniadis, 2014a).

4.3 Scénarisation des séquences pédagogiques

Une fois les textes annotés et étiquetés, nous avons décidé de concevoir un scénario informatique dans lequel nous avons mis en place une série d'activités. Le but de ces activités est de faire travailler aux étudiants en formation des notions linguistiques à l'aide de textes, et à partir de ces textes les aider à structurer leur discours. L'idée en somme est de les faire travailler dans deux

³⁵ J.-M. Adam proposera plus précisément la notion de «proposition énoncée», la considérant comme «l'unité textuelle élémentaire». Il la définira par la suite comme: «[...] une unité de base, effectivement réalisée et produite par un acte d'énonciation, donc comme un énoncé minimal.» (Adam, 2011a: 83) Selon ce même auteur, dans le texte les propositions ne sont jamais isolées, car elles font référence à d'autres propositions.

sens: 1) sur les aspects linguistiques de la langue française; 2) sur la didactisation de ces aspects linguistiques. Afin d'atteindre cet objectif, nous avons mis en place des séquences pédagogiques ou séquences d'apprentissage (figure 6).

4.3.1 Élaboration des séquences

La structure des activités élaborées est celle qui apparaît chez Cuq (2003) et qui comporte trois phases: 1) activités de découverte; 2) activités de systématisation; 3) activités d'utilisation. La figure 8 montre la structure des activités que notre système est censé proposer, ainsi que les données nécessaires. Pour notre système, une activité est divisée en quatre phases: 1) exercices de repérage ou de découverte; 2) explications de conceptualisations théoriques; 3) exercices de systématisation; 4) exercices d'application de connaissances. Les textes utilisés proviennent du corpus évoqué plus haut; les explications de conceptualisations théoriques d'une base de données spécifique.

Globalement, nous avons défini les quatre phases de la manière suivante:

- **Exercices de repérage ou de découverte:** dans cette phase les étudiants en formation seront confrontés à des activités de type déductif dans lesquelles ils découvriront des règles de grammaire ou d'analyse textuelle au moyen de textes authentiques que leurs enseignants auront choisis dans la base de données grâce à une interface appropriée.
- **Explications de conceptualisations théoriques:** ici, l'enseignant formateur explicitera la règle ou le concept théorique aux étudiants grâce à un support théorique informatique et à des exemples. Il partira, cette fois-ci, d'une démarche inductive afin de mieux faire comprendre les règles et les concepts. Il donnera, dans un premier temps, les notions langagières et/ou textuelles, et dans un deuxième temps, il s'agira de leur montrer les exploitations pédagogiques possibles.
- **Exercices de systématisation:** dans cette troisième phase, et au travers d'exercices d'application des concepts théoriques appris dans la phase précédente, les étudiants pourront s'entraîner à des aspects plus spécifiques du sujet traité. Il s'agira donc de textes à trous, de QCM, d'unités à glisser et à mettre en place, etc.
- **Exercices d'application des connaissances:** dans cette dernière phase, les étudiants en formation vont rédiger des textes dans lesquels ils utiliseront les notions acquises aux phases précédentes (coréférence textuelle, progression thématique, etc.). Une première sous-phase consistera à envoyer un document écrit dans lequel il y aura l'étude qu'ils ont effectuée pour une notion à envoyer (via la plateforme). Après correction par l'enseignant, l'étudiant procédera à la didactisation de la notion, et enverra sa production à l'enseignant formateur. Ce type d'activité permettra d'évaluer, d'une part, l'acquisition des notions linguistiques étudiées, d'autre part, la capacité des futurs enseignants à didactiser ces notions.

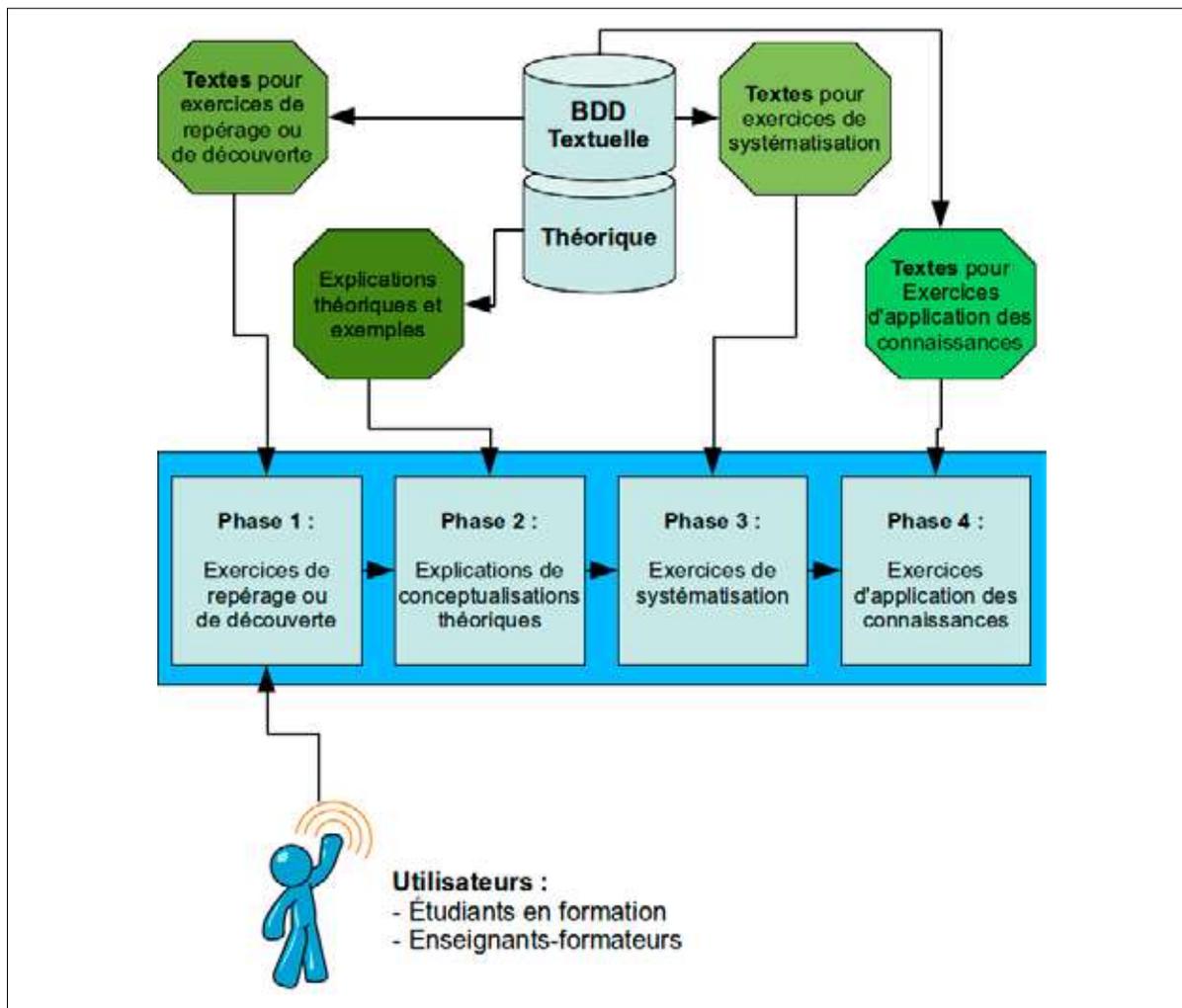


FIGURE 6 Développement d'une séquence pédagogique (Molina Mejia & Antoniadis, 2014a).

4.3.2 Types d'activités

Notre système propose des activités qui reprennent des notions liées à la linguistique textuelle et à l'analyse textuelle, ainsi qu'indirectement à l'analyse du discours. Trois classes d'activités sont proposées :

- **Activités autour de la structuration textuelle :** ce type d'activités peut permettre, comme son nom l'indique, de structurer un texte afin de le rendre compréhensible, que ce soit au niveau du texte en général, des paragraphes, au niveau des phrases et de leur importance, etc. Nous avons sous-divisé ce premier type d'activité en trois sous-types d'activités.
 - *Activités d'analyse des phrases principales :* nous demandons d'analyser le texte et ses composantes à partir de l'emploi de phrases principales et secondaires. Cette analyse permet aux apprenants de structurer leur pensée et de mieux rédiger par la suite.
 - *Activités d'analyse de la structure logique des textes :* dans le cas des textes de type argumentatif ou explicatif, il y a une structure qui suit une logique. Cette structure est donnée, normalement, par un titre principal, un ou plusieurs sous-titres, une introduction, un ou plu-

sieurs aspects (paragraphe de développement ou d'argumentation), une conclusion, une signature, une source et une date. Repérer ces éléments permet aux étudiants en formation de structurer leurs textes et leurs discours d'une manière plus logique et cohérente.

- *Activités d'analyse du type ou des séquences textuelles*: le type de texte ou des séquences dans un texte est un sujet très étudié par la linguistique textuelle (Adam, 2011b). Ce type d'activité permet une analyse textuelle en fonction de la construction de celui-ci; ceci permet de déterminer la prise de position éventuelle de l'auteur.
- **Activités autour de la cohérence et de la cohésion textuelles**: les notions de cohérence et de cohésion textuelles sont d'une importance capitale en linguistique textuelle. Elles permettent de mieux structurer le texte et le discours. Parmi les différentes notions qui font partie de ces deux phénomènes, nous avons choisi de nous concentrer sur deux types d'activités:
 - *Activités sur la coréférence textuelle*: l'utilisation des endophores (anaphores et cataphores) permet de ne pas répéter dans un discours ou dans un texte donné les mêmes mots ou les mêmes expressions. Avec l'emploi de ce type d'éléments textuels ou discursifs, le locuteur peut enrichir son discours tout en restant dans une logique de la langue. L'étude des endophores permet de dépasser le stade de l'analyse traditionnelle de la phrase (simple ou complexe) pour aller vers une analyse intra et interphrastique qui tiendra compte du contexte et du cotexte.
 - *Activités sur l'emploi des connecteurs et des marqueurs logico-temporels et discursifs*: l'emploi de ce type d'éléments linguistiques permet de donner de la cohésion au texte et de la cohérence au discours que celui-ci véhicule. De la même façon que les coréférents, les connecteurs et les marqueurs logico-temporels et discursifs permettent d'établir les relations entre les phrases et les paragraphes et les textes.
- **Activités autour de la progression thématique**: les notions de thème et rhème sont les notions clés au niveau de la progression thématique (Combettes, 1983). Le thème représente ce dont on parle, et le rhème ce que l'on dit à propos du thème. Il y a trois types de progression thématique: 1) la progression à thème constant; 2) la progression linéaire simple; et 3) la progression à thèmes dérivés. Il est dans notre intérêt de faire apprendre ces notions, ainsi que de les articuler dans des textes.

5 Conclusion et perspectives

L'élaboration de notre système est le produit de la mise en œuvre de connaissances et de compétences issues de plusieurs disciplines ou domaines: didactique des langues, linguistique générale, linguistique textuelle, TAL, informatique, linguistique de corpus, ingénierie pédagogique. Il se veut principalement un outil pédagogique pour l'enseignement/apprentissage des langues. Comme le suggère Antoniadis (2010), nous ambitionnons de fournir aux étudiants en formation un outil qui soit pensé en fonction de leurs besoins et de leurs objectifs d'apprentissage. En ce sens, sa problématique est essentiellement didactique, ses fonctions uniquement pédagogiques, ses utilisateurs visés (les enseignants-formateurs et les étudiants en formation) sont considérés sans connaissances autres que celles de leur domaine d'apprentissage, son utilité est ainsi la plus-value pédagogique qu'il apporte.

La maquette sera évaluée et testée au sein des universités colombiennes qui ont montré de l'intérêt pour ce projet. Nous comptons par la suite recueillir les réactions des utilisateurs pour en développer une version plus complète à proposer à d'autres universités qui pourraient être intéressées par ce type d'outil d'aide à la formation des futurs enseignants de FLE.

Références

- ADAM J.-M.** (1990). *Éléments de linguistique textuelle: Théorie et pratique de l'analyse textuelle*. Pierre Mardaga: Liège.
- ADAM J.-M.** (2011a). *La linguistique textuelle*. Collection Linguistique Cursus, troisième édition. Armand Colin: Paris.
- ADAM J.-M.** (2011 b). *Les textes types et prototypes*. Collection Linguistique Cursus, troisième édition. Armand Colin: Paris.
- ANTONIADIS G.** (2010). «*De l'apport pertinent du TAL pour les systèmes d'ALAO. L'exemple du projet MIRTO*». 2^e Congrès Mondial de Linguistique Française (CMLF-2010). 12-15 juillet 2010, La Nouvelle Orléans, USA.
- ANTONIADIS G., ECHINARD S., KRAIF O., LEBARBÉ T. & PONTON C.** (2005). Modélisation de l'intégration de ressources TAL pour l'apprentissage des langues: la plateforme MIRTO. In *ALSIC*, 8 (Numéro spécial TALAL) pp. 65-79.
- BAKHTINE M.** (1981). *Écrits du cercle Bakhtine*. In T. Todorov (éditeur): *Mikhaïl Bakhtine, le principe dialogique*. Seuil: Paris.
- BEACCO J.-C.** (2010). *La didactique de la grammaire dans l'enseignement du français et des langues*. Langues et didactique. Didier: Paris.
- BENVENISTE E.** (1974). *Problèmes de linguistique générale II*. Gallimard: Paris.
- CHARLIER B., DESCHRYVER N. & PERAYA D.** (2006). «Apprendre en présence et à distance: une définition des dispositifs hybrides». In *Distances et savoirs*, 2006/4 Vol. 4 pp. 469-496.
- CHARAUDEAU P. & MAINGUENEAU D.** (2002). *Dictionnaire d'analyse du discours*. Éditions du Seuil: Paris.
- CHAUDIRON S.** (2007). «Technologies linguistiques et modes de représentation de l'information textuelle». *Documentaliste-Sciences de l'information*, 2007/1 Vol. 44 pp. 30-39.
- CHISS J.-L. & DAVID J.** (2012). Didactique du français et étude de la langue: Linguistique textuelle et didactique du français. In *Le Français aujourd'hui. Vol. Hors-série* pp. 155-173. Armand Colin: Paris.
- COMBETTES B.** (1983). *Pour une grammaire textuelle: la progression thématique*. Formation continue. De Boeck – Duculot: Bruxelles.
- COSERIEU E.** (2001). *L'homme et son langage*. Sterling-Virginia, Peeters: Louvain-Paris.
- COUTO J., LUNDQUIST L. & MINEL J.-L.** (2005). «*Navigation interactive pour l'apprentissage en linguistique textuelle*». Environnements Informatiques pour l'Apprentissage Humain, Montpellier 2005.
- CUQ J.-P.** (2003). *Dictionnaire de didactique du français: Langue étrangère et seconde*. CLE International: Paris.
- DEMOUGIN F.** (2008) «La didactique des langues – cultures à la croisée des méthodes», Tréma [En ligne], 30 | 2008, mis en ligne le 01 novembre 2010, Consulté le 22 janvier 2015. URL: <http://trema.revues.org/427>

GUIR R. (2002). *Pratiquer les TICE: Former les enseignants et les formateurs à des nouveaux usages*. Collection Pédagogies en Développement. De Boeck – Larcier S.A.: Bruxelles.

HALTÉ J.-F. (2004). La grammaire au cœur des apprentissages? In *Langue et études de la langue: Approches linguistiques et didactiques. Langues et écritures*. pp. 11-23. Publication de l'Université de Provence: Marseille.

KLINKENBERG J.-M. (1993). Le français: une langue en crise? In *Études françaises*, vol. 29, N° 1, pp. 171-190.

LUNDQUIST L. (1990). *L'analyse textuelle: méthode, exercices*. Deuxième édition, DJOF Publishing: Copenhague.

LUNDQUIST L. (2013). *Lire un texte académique en français*. Collection l'essentiel français. Éditions Ophrys: Paris.

MAINGUENEAU D. (1991). *L'analyse du discours: introduction aux lectures de l'archive*. Hachette Supérieur: Paris.

MAINGUENEAU D. (2009). *Les termes clés de l'analyse du discours*. Éditions du Seuil: Paris.

MANGENOT F. (1998). «Outils textuels pour l'apprentissage de l'écriture en L1 et en L2». In *Pratiques discursives et acquisition des langues étrangères* (Université de Franche-Comté, 1998), actes du 10^e colloque international FOCAL, Besançon, 19-21 septembre 1996.

MOLINA MEJIA J. M. (2014). Conception d'un environnement informatique fondé sur la linguistique textuelle et destiné à la formation des enseignants de FLE en Colombie. In Pierre-André CARON & Ronan CHAMPAGNAT (Éditeurs): *Actes des 5^e Rencontres des Jeunes Chercheurs en Environnements Informatiques pour l'Apprentissage Humain*. Université de la Rochelle, 16 - 17 juin 2014, pp. 23-29. La Rochelle – France.

MOLINA MEJIA J. M. & ANTONIADIS G. (2014a). Conception d'un environnement informatique fondé sur la linguistique textuelle pour la formation initiale des futurs enseignants de FLE en Colombie. In *Adjectif: Analyses recherches sur les TICE*. Article 293, 05 Juin 2014. [En ligne] <http://www.adjectif.net/spip/spip.php?article293>

MOLINA MEJIA J. M. & ANTONIADIS G. (2014 b). Conception et élaboration d'un environnement informatique fondé sur la linguistique textuelle et destiné à la formation des futurs enseignants de FLE en Colombie. In Françoise OLMO CAZEVIEILLE & Jean-Marc MANGIANTE (Éditeurs): *II Coloquio franco-español de análisis del discurso y enseñanza de lenguas para fines específicos. Lenguas, comunicación y tecnologías digitales*. Universitat Politècnica de Valencia, 03 - 05 septembre 2014, pp. 143-158. Colección Congresos. Editorial Universitat Politècnica de Valencia: Valencia - Espagne.

POP L. (2000). *Espaces discursifs: pour une représentation des hétérogénéités discursives*. Éditions Peeters: Louvain.

PUREN C. (1995). Des méthodologies constituées et de leur mise en question. In *Le Français dans le monde: Méthodes et méthodologies (Numéro spécial)* pp. 36-41.

REY-DEBOVE J. (1997). *Le métalangage*. Armand Colin/Masson: Paris.

RIEGEL M., PELLAT J.-C. & RIOUL R. (2009). *Grammaire méthodique du français*. Quadrige, quatrième édition. Presses Universitaires de France: Paris.

Annexe A : Extrait des annotations d'un texte à l'aide de l'analyseur CORDIAL (Molina Mejia, 2014).

Une génération inoxydable

C'est un phénomène unique dans l'histoire française, né d'une conjonction particulièrement favorable de la démographie et de l'économie plus rare que l'on sait déjà qu'il ne pourra pas durer. Ce phénomène, c'est l'apparition d'une génération inédite. Née entre 1936 et 1950, elle fête aujourd'hui ses 50-55 ou 65 ans et paraît en tout point hors norme. D'abord, parce qu'elle est fort nombreuse. Ces plus de 55 ans, qui furent les petits Français babillants du baby-boom de l'après-guerre, représentent aujourd'hui 16 millions de nos compatriotes, soit 20 % de la population ! (...)

N°	MDT	LEMMES	P	Prop.	Franchon	Groupe	St-Gr.	Type	Type détaillé	Sémantique
1	C	cet	1	indépendante	Sujet	Groupe pronominal	1/1	PRO	PRON. Det. Sing.	existence / événement / vérité
2	est	être	1	indépendante	Verbe	Groupe nominal	2/2	VINDP3S	Indicateur PRESENT	poly-sémique - fait observable, événement / événement
3	un	un	1	indépendante	Attribut du sujet	Groupe nominal	4/4	NDM3S	ART. Ind. Masc. Sing.	poly-sémique - fait observable, événement / événement
4	phénomène	phénomène	1	indépendante	Attribut du sujet	Groupe nominal	4/4	NDM3S	NDM Masc. Sing.	poly-sémique - fait observable, événement / événement
5	unique	unique	1	indépendante	Attribut du sujet	Groupe nominal	4/4	ADJ3S	ADJ Sing. Inv. Genre	poly-sémique - original, spécial / original, spécial
6	dans	dans	1	indépendante	Attribut du sujet	Groupe nominal prépositionnel	4/8	PREP	PREPOSITION	poly-sémique - étude historique / possible, évolution continue
7	le	le	1	indépendante	Attribut du sujet	Groupe nominal prépositionnel	4/8	DETDF	ART. Det. Fem. Sing.	poly-sémique - étude historique / possible, évolution continue
8	histoire	histoire	1	indépendante	Attribut du sujet	Groupe nominal prépositionnel	4/8	NDF3S	NDM Fem. Sing.	poly-sémique - étude historique / possible, évolution continue
9	française	française	1	indépendante	Attribut du sujet	Groupe nominal prépositionnel	4/8	ADJF3S	ADJ Fem. Sing.	poly-sémique - étude historique / possible, évolution continue
10	né	naître	1	indépendante	prépositionnel	Groupe nominal prépositionnel	11/11	VPR3PMS	Participle PASSE M.	poly-sémique - manifestation / origine / être / être, begin
11	d'	de	1	indépendante	Complément d'objet indirect	Groupe nominal prépositionnel	14/14	PREP	PREPOSITION	poly-sémique - manifestation / origine / être / être, begin
12	de	de	1	indépendante	Complément d'objet indirect	Groupe nominal prépositionnel	14/14	DE TIPS	ART. Ind. Fem. Sing.	poly-sémique - manifestation / origine / être / être, begin
13	un	un	1	indépendante	Complément d'objet indirect	Groupe nominal prépositionnel	14/14	NDF3S	NDM Fem. Sing.	poly-sémique - manifestation / origine / être / être, begin
14	conjonction	conjonction	1	indépendante	Complément d'objet indirect	Groupe nominal prépositionnel	14/14	ADV	ADVERBE	poly-sémique - manifestation / origine / être / être, begin
15	particulièrement	particulièrement	1	indépendante	Complément d'objet indirect	Groupe nominal prépositionnel	14/14	ADV	ADVERBE	poly-sémique - manifestation / origine / être / être, begin
16	favorable	favorable	1	indépendante	Complément d'objet indirect	Groupe nominal prépositionnel	14/14	ADV	ADVERBE	poly-sémique - manifestation / origine / être / être, begin
17	de	de	1	indépendante	Complément d'objet indirect	Groupe nominal prépositionnel	14/19	PREP	ADJ Sing. Inv. Genre	poly-sémique - manifestation / origine / être / être, begin
18	le	le	1	indépendante	Complément d'objet indirect	Groupe nominal prépositionnel	14/19	DETDF	ART. Det. Fem. Sing.	poly-sémique - manifestation / origine / être / être, begin
19	démographique	démographique	1	indépendante	Complément d'objet indirect	Groupe nominal prépositionnel	14/19	NDF3S	NDM Fem. Sing.	poly-sémique - manifestation / origine / être / être, begin
20	et	et	2	coordonnée	Complément d'objet indirect	Groupe nominal prépositionnel	14/19	CONJ	CONJ. Coordon.	poly-sémique - manifestation / origine / être / être, begin
21	de	de	2	coordonnée	Complément d'objet indirect	Groupe nominal prépositionnel	18/23	PREP	PREPOSITION	poly-sémique - manifestation / origine / être / être, begin
22	le	le	2	coordonnée	Complément d'objet indirect	Groupe nominal prépositionnel	18/23	DETDF	ART. Det. Fem. Sing.	poly-sémique - manifestation / origine / être / être, begin
23	économie	économie	2	coordonnée	Complément d'objet indirect	Groupe nominal prépositionnel	18/23	NDF3S	NDM Fem. Sing.	poly-sémique - manifestation / origine / être / être, begin
24	plus	plus	2	coordonnée	Complément d'objet indirect	Groupe nominal prépositionnel	18/23	ADV	ADVERBE	poly-sémique - manifestation / origine / être / être, begin
25	rare	rare	2	coordonnée	Complément d'objet indirect	Groupe nominal prépositionnel	18/23	ADV	ADVERBE	poly-sémique - manifestation / origine / être / être, begin
26	que	que	2	coordonnée	Complément d'objet indirect	Groupe nominal prépositionnel	18/23	ADV	ADVERBE	poly-sémique - manifestation / origine / être / être, begin
27	f	f	2	coordonnée	Complément d'objet indirect	Groupe nominal prépositionnel	18/23	ADV	ADVERBE	poly-sémique - manifestation / origine / être / être, begin
28	en	en	2	coordonnée	Complément d'objet indirect	Groupe nominal prépositionnel	27/27	NDF3S	NDM Masc. Inv. Née	poly-sémique - manifestation / origine / être / être, begin
29	par	par	2	coordonnée	Complément d'objet indirect	Groupe nominal prépositionnel	28/28	PREP3S	PRON. Fem. 3e S	poly-sémique - manifestation / origine / être / être, begin
30	déjà	déjà	2	coordonnée	Complément d'objet indirect	Groupe nominal prépositionnel	29/29	VINDP3S	Indicateur PRESENT	poly-sémique - manifestation / origine / être / être, begin
31	qu'	qu'	3	subordonnée	Complément circonstanciel de...	Groupe pronominal	32/32	SUB	CONJ. Subord.	poly-sémique - manifestation / origine / être / être, begin
32	il	il	3	subordonnée	Complément circonstanciel de...	Groupe pronominal	34/34	ADV	PRON. Fem. 3e S	poly-sémique - manifestation / origine / être / être, begin
33	ne	ne	3	subordonnée	Complément circonstanciel de...	Groupe pronominal	34/34	ADV	ADVERBE	poly-sémique - manifestation / origine / être / être, begin
34	pourra	pourra	3	subordonnée	Complément circonstanciel de...	Groupe pronominal	34/34	VINDF3S	Indicateur FUTUR 3e	poly-sémique - manifestation / origine / être / être, begin
35	pas	pas	3	subordonnée	Complément circonstanciel de...	Groupe pronominal	34/34	ADV	ADVERBE	poly-sémique - manifestation / origine / être / être, begin
36	durer	durer	4	infinitive	Complément circonstanciel de...	Groupe pronominal	36/36	VINF	INFINITIF	poly-sémique - manifestation / origine / être / être, begin

Annexe B: Exemple d'annotations d'un texte à l'aide du format XML (Molina Mejia, 2014).

```

- <DOCUMENTCOREF>
- <ENTETE>
  <titre_doc>Une génération inoxydable</titre_doc>
  <auteur>Catherine Golliau et Émilie Lamez</auteur>
  <source type="magazine" periodicite="hebdomadaire">Le Point</source>
  <date>03/10/03</date>
  <type>Argumentatif</type>
  <niveau>B2-CECR</niveau>
  <activite>Corréférence textuelle</activite>
- <CONTENU>
  - <titre>
    - <phrase type="simple" >
      - <phrase type="verbiale" >
        - <referent type="Nom" sous-type="GN" id="1" chaine="1" >
          <element type="DETFS" >Une</element>
          <element type="NCFS" >génération</element>
          <element type="ADJSIG" >inoxidable</element>
        </referent>
        <prop>
          <phrase >
            <titre >
              - <paragraphe >
                - <phrase type="complexe" >
                  - <coreferent type="Pron" sous-type="Dem" id="2" chaine="1" position="Anaphore" reference="1" >
                    <element type="PDS" >C</element>
                  </coreferent>
                  - <coreferent type="Verb" temps="Present" mode="Ind" id="3" chaine="1" position="Anaphore" reference="2" >
                    <element type="VINDPSS" >est</element>
                  </coreferent>
                - <referent type="Nom" sous-type="GN" id="4" chaine="2" >
                  <element type="DETMS" >un</element>
                  <element type="NCMS" >phénomène</element>
                  <element type="ADJSIG" >unique</element>
                  <element type="PREP" >dans</element>
                  <element type="DETFDS" >I</element>
                  <element type="NCFS" >histoire</element>
                  <element type="ADJFS" >français</element>
                </referent>
                <element type="PCTFAB" ></element>
              </titre >
            </phrase >
          </prop >
        </phrase >
      </titre >
    </paragraphe >
  </titre >

```