



**HAL**  
open science

## Conception d'un environnement informatique fondé sur la linguistique textuelle pour la formation initiale des futurs enseignants de FLE en Colombie

Jorge Mauricio Molina Mejia, Georges Antoniadis

### ► To cite this version:

Jorge Mauricio Molina Mejia, Georges Antoniadis. Conception d'un environnement informatique fondé sur la linguistique textuelle pour la formation initiale des futurs enseignants de FLE en Colombie. Adjectif: analyses et recherches sur les TICE, 2014, <http://www.adjectif.net/spip/spip.php?article293>. hal-01091866

**HAL Id: hal-01091866**

**<https://hal.science/hal-01091866v1>**

Submitted on 7 Dec 2014

**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.

# Conception d'un environnement informatique fondé sur la linguistique textuelle pour la formation initiale des futurs enseignants de FLE en Colombie

par Jorge Mauricio Molina Mejia & Georges Antoniadis

LIDILEM, Université Grenoble-Alpes

*Document de travail, le 05/22/14*

**Chapeau :** Notre article retrace la construction d'un projet de recherche pluridisciplinaire, autour d'un environnement informatique d'aide à la formation fondé sur les principes de la linguistique textuelle. Il concerne la formation initiale des futurs enseignants de Français Langue Etrangère (FLE) en Colombie.

**Mots clés :** Formation des enseignants, environnement informatique, linguistique textuelle, Traitement automatique des langues, FLE.

## Introduction

L'utilisation d'instruments informatiques dans l'enseignement/apprentissage des langues étrangères se fait depuis plusieurs décennies. Ces instruments visent, de manière générale, un public assez large d'apprenants de langue en situation d'apprentissage, souvent en autonomie. Néanmoins, leur emploi dans la formation des futurs enseignants de langue et notamment en FLE dans le contexte de l'éducation supérieure colombienne, a été beaucoup moins exploré. Comme nous le montrons (Molina & Antoniadis, 2014) l'utilisation de technologies informatiques de manière hybride (modules de formation qui peuvent être utilisés dans les cours en présentiel et des exercices, pouvant être utilisés en dehors de la salle de classe, en autonomie) peut s'avérer une aide précieuse pour la formation des futurs enseignants de FLE en Colombie (Salamanca Lamouroux, 2014).

Nous avons donc cherché à élaborer un projet fondé sur l'idée de fournir aux formateurs un instrument informatique qui soit aisément utilisable et paramétrable, et qui leur permettrait la préparation et la mise en place de séquences de formation au travers d'activités mobilisant des technologies destinées aux étudiants en formation initiale. Ce projet s'appuie sur un corpus textuel explorant différentes notions issues de la linguistique textuelle grâce aux techniques du Traitement Automatique des Langues (TAL).

En effet, nous considérons, à l'instar de Chaudiron (2007) et Antoniadis (2010), que les techniques et procédures du TAL permettent de tenir compte et de représenter des informations langagières à faire apprendre aux étudiants. Le projet devrait être aussi développé en fonction des besoins des usagers (enseignants/formateurs et étudiants). En ce sens, le support théorique de la linguistique textuelle puise ses sources dans l'analyse linguistique des textes à partir de l'idée du *contexte* et du *co-texte*, l'analyse du discours (Adam, 2011) et la notion de *discursivité textuelle* (Maingueneau, 1991). Nous avons cherché, à l'instar de ce que proposent Chiss et David (2012), à amener notre public cible à mieux maîtriser les contraintes d'ordre extralinguistique - mais aussi linguistiques - leur permettant d'interpréter et de construire des textes cohérents et, par ce moyen, d'améliorer aussi leur discours.

## Genèse du projet

### ***La formation des formateurs de FLE dans le contexte colombien***

Notre projet est né en 2010, à partir de nos observations et de notre expérience d'enseignant de FLE dans plusieurs universités colombiennes (notamment l'université d'Antioquia) : nous avons identifié quelques problèmes concernant la formation linguistique des futurs enseignants de FLE. Nous avons donc décidé de mettre en œuvre un projet qui, depuis cette année-là, fait l'objet d'une thèse de doctorat financée par le Ministère de Technologies de l'Information et des Communications du gouvernement colombien.

#### **Hypothèses de départ**

La première étape du projet, après les observations, a été de définir les hypothèses de travail et d'élaborer un questionnaire de recherche, informatisé, à destination du public cible colombien (étudiants en formation initiale visant à devenir professeurs de FLE et leurs formateurs). Ce questionnaire avait deux objectifs : il devait nous permettre de valider nos hypothèses et de mieux saisir les notions devant être enseignées au sein de notre environnement informatique. L'hypothèse générale de travail était que l'insécurité linguistique fait que les étudiants sont généralement gênés pour enseigner la langue française en utilisant un métalangage adéquat. Lorsque nous parlons d'insécurité linguistique, nous reprenons les mots de Klinkenberg (1993 : 185) :

*Il y a insécurité dès que l'on a une image assez nette de la norme, mais que l'on 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 où son usage n'est pas légitime, mais sans qu'il ait une conscience nette de la non-conformité.*

Cette hypothèse générale était enrichie par l'hypothèse que l'insécurité linguistique provient d'une connaissance un peu superficielle des phénomènes linguistiques inhérents à la langue cible. Ce que nous avons nommé la « norme linguistique » et qui, selon E. Coseriu (2001 : 246-247), pourrait être conçue 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.*

Nous considérons qu'un futur enseignant de FLE se doit non seulement de connaître la norme d'usage de la langue cible mais aussi d'être en capacité d'adapter son répertoire langagier et le métalangage auquel il devrait avoir recours lorsqu'il est en situation d'explication des faits de langue à des apprenants.

#### **Méthodologie de recherche**

Trois types de publics ont été enquêtés :

- les enseignants/formateurs (Ens-Form),
- les étudiants en formation (Etud-Form)
- et les enseignants de FLE (Ens-FLE) formés dans une université colombienne.

Nous avons reçu au total 172 réponses des enquêtés appartenant à 12 des 18 institutions d'éducation supérieure ayant un programme de formation pour les enseignants de FLE.

Afin de vérifier l'hypothèse générale signalée et de pouvoir définir les activités à développer dans notre système informatique, nous avons réalisé des entretiens semi-directifs ainsi qu'un questionnaire de recherche informatisé de notre public<sup>1</sup>, sur quatre grands axes, à savoir :

- **Celui de la formation didactique** : dans cette section nous avons voulu enquêter sur les pratiques d'enseignement dominante ou les priorités au sein de la formation des futurs enseignants de FLE. Selon notre enquête, nous avons pu constater que les aspects didactiques privilégiés sont la didactique de l'oral suivie de la didactique de l'écrit.
- **Celui de la formation linguistique** : dans cette section nous avons voulu enquêter sous les compétences linguistiques privilégiées au niveau de la formation (compétence grammaticale, sémantique, phonologique, orthographique et lexicale). La sous-compétence privilégiée lors de la formation des futurs enseignants de FLE est la compétence grammaticale (environ 80% pour les enseignants/formateurs, 60% pour les enseignants de FLE et 86% pour les étudiants), suivie de la compétence sémantique (environ 70% pour les étudiants et enseignants de FLE, et 50% pour les enseignants formateurs).
- **Celui de la connaissance et de l'utilisation de la linguistique textuelle** : en ayant demandé au public enquêté sur l'utilisation de la linguistique textuelle au niveau de la formation. Nous avons pu constater que bien que les enseignants/formateurs déclarent utiliser des textes et les analyser dans les cours, les étudiants en formation disent ne pas connaître cette approche théorique<sup>2</sup>. Pour les trois publics les aspects à être améliorés au travers de la linguistique textuelle sont : la cohérence textuelle, la structuration textuelle et les éléments constituants du texte.
- **Celui de l'emploi des instruments informatiques au sein de la formation initiale** : nous avons pu constater que l'utilisation des instruments informatiques intéresse les trois publics au niveau de la formation des futurs enseignants de FLE. Nous avons notamment constaté que ce sont les blogs pédagogiques et les sites ou plateformes e-learning qui relèvent les plus d'intérêt au niveau de la formation.

## Un projet fondé sur la linguistique textuelle

La linguistique textuelle est une approche théorique d'analyse linguistique relativement récente. Certains auteurs (Adam, 1990) situent son apparition au milieu des années 1950, d'autres (Charaudeau & Maingueneau, 2002) voient son émergence dans les années 1960. Ce qui nous semble important dans cette approche est le fait qu'au lieu de prendre la phrase comme unité d'analyse et de travail, elle s'intéresse à une unité beaucoup plus large : le texte (Adam, 1990 ; Lundquist, 1990). Ceci permet de faire des analyses différentes et complémentaires de ce que l'on faisait auparavant avec l'analyse structuraliste ou l'analyse de la Grammaire Générative Transformationnelle (GGT).

### *Notions liées à la linguistique textuelle*

Parmi les notions qui font partie de cette approche théorique, nous en avons choisi trois en les sous-divisant en sous-notions :

- **l'analyse de la structure textuelle** comprend celles des phrases principales<sup>3</sup>, de la *structure logique d'un texte* et du *type et des séquences textuelles* ;
- **l'analyse de la cohérence et cohésion textuelle** comprend celles de la *coréférence textuelle* (anaphores et cataphores) et l'utilisation des *marqueurs et connecteurs logico-temporels et argumentatifs* ;
- **l'analyse de la progression thématique** comprend celles des notions de *thème* et de *rhème*.

C'est à partir de ces trois grandes notions que la conception d'un environnement informatisé pour l'apprentissage pourrait améliorer la performance linguistique des futurs enseignants colombiens de FLE. Nous pensons qu'en travaillant ces trois axes, un individu pourra mieux structurer son discours au moment d'enseigner.

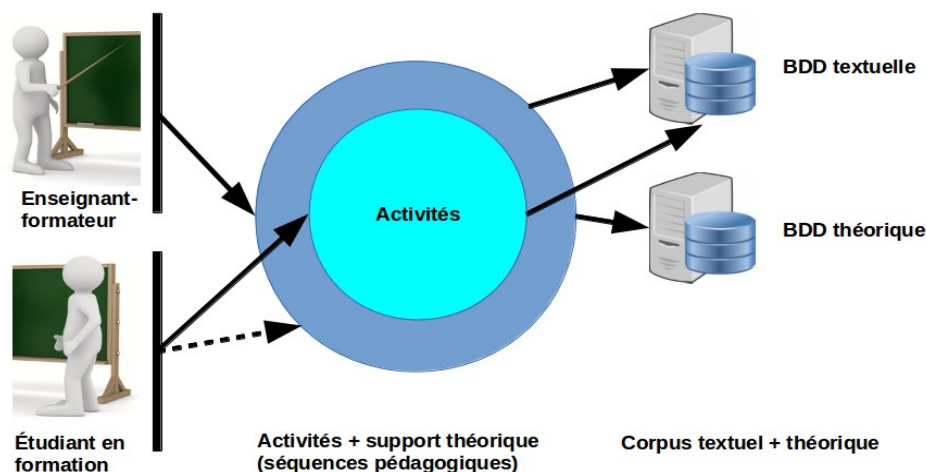
### ***L'intérêt d'un dispositif informatique fondé sur la linguistique textuelle***

Dans le domaine de la formation de formateurs de FLE en Colombie, les instruments informatiques de type Acquisition des Langues Assistée par Ordinateur (ALAO) destinés spécifiquement à un public d'étudiants en formation initiale et conçus à partir des besoins des étudiants sont inexistantes. Il n'y a pas, à notre connaissance, dans ce contexte, de dispositifs informatiques qui forment les étudiants aux aspects linguistiques de la langue cible, ainsi qu'à leur didactisation<sup>4</sup>.

Certains dispositifs de formation que nous avons identifiés en France, comme le « Français en première ligne » (Develotte & Mangenot, 2010), sont plutôt intéressés par les interactions d'étudiants en formation. Les instruments informatiques cherchant à former à partir de la *textualité* s'avèrent presque inexistantes (Mangenot, 1998 ; Couto *et al.*, 2005).

### **Fonctionnalités et architecture du système informatique**

L'environnement informatique que nous concevons est constitué de deux interfaces (figure 1). La première interface devrait permettre aux enseignants-formateurs de constituer et de paramétrer des séquences pédagogiques de formation. Une fois créées, ces séquences sont envoyées aux étudiants en formation qui y auront accès par une autre interface leur permettant de réaliser les activités programmées par les enseignants. Les étudiants en formation pourront également accéder au support théorique et didactique des différentes séquences pédagogiques, si l'enseignant-formateur décide de leur donner les droits pour consulter la séquence. Le système dispose de deux bases de données (BDD) : la première est destinée au corpus textuel et aux activités et la deuxième au support théorique et aux schémas linguistiques et didactiques.



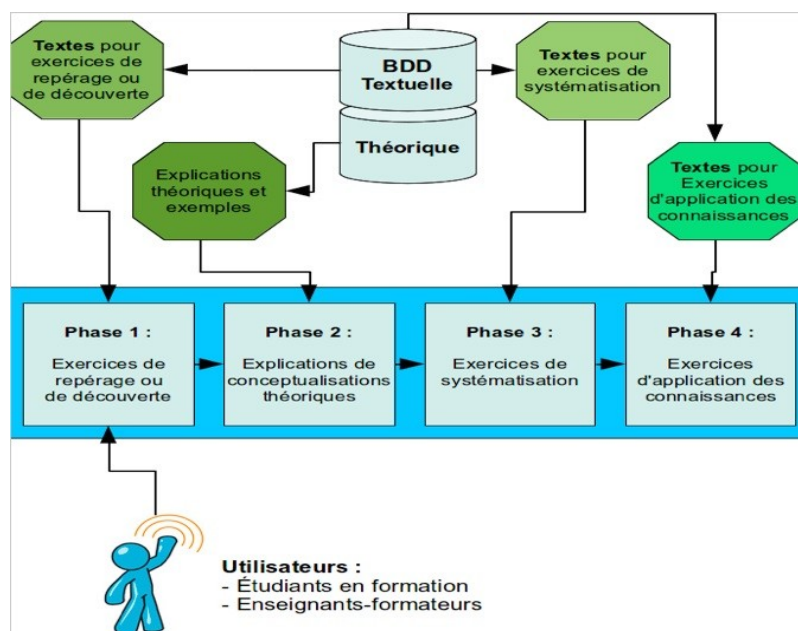
**Figure 1 : architecture générale du système informatique de formation**

### **Scénarisation des activités pédagogiques :**

Les deux BDD avant signalées permettent la mise en œuvre du scénario pédagogique de formation à partir des séquences didactiques de formation.

#### **Séquences didactiques**

Nous avons décidé de constituer notre système de formation à partir des séquences didactiques. Chacune est constituée de quatre phases (figure 2).



**Figure 2 : schéma représentant une séquence didactique**

Comme nous pouvons le voir, les phases 1, 3, et 4 utilisent la BDD textuelle (textes annotés avec les notions à travailler) ce qui permet le paramétrage des exercices. La phase 2 (le support théorique et les schémas) sera stockée dans la BDD théorique utilisée par l'enseignant-formateur afin d'expliquer les phénomènes linguistiques liés à la linguistique textuelle. Chaque phase comporte des exercices ou des contenus de formation différents :

- **Exercices de repérage ou de découverte (phase 1) :** ils permettent de découvrir un phénomène d'ordre linguistique, grâce à un travail de type déductif. Des exercices de permutation d'éléments et textes lacunaires permettant de découvrir les phénomènes linguistiques à partir d'activités d'ordre déductif.
- **Explications de conceptualisations théoriques et didactiques (phase 2) :** elles permettent aux étudiants en formation de mieux comprendre les phénomènes étudiés dans la phase précédente.
- **Exercices de systématisation (phase 3) :** ils permettent aux étudiants de mieux saisir les notions travaillées. Des exercices de type QCM, de textes lacunaires, permutation d'éléments, etc., permettent d'une manière assez simple de s'exercer aux notions travaillées pendant le processus de formation.
- **Exercices d'application de connaissances (phase 4) :** dans cette phase, les étudiants vont pouvoir créer leurs propres textes en utilisant les notions apprises au travers des séquences de formation. Les étudiants envoient des textes créés par eux-mêmes et qui seront envoyés à l'enseignant/formateur pour leur correction.

### **Constitution du corpus textuel**

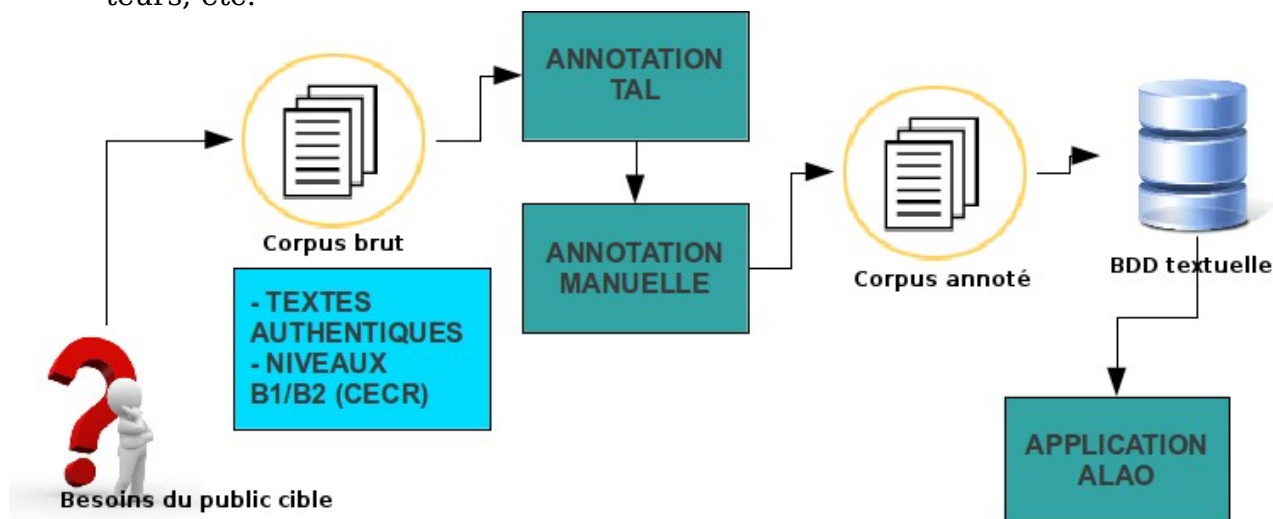
#### **Type de textes**

Notre corpus est constitué, actuellement, d'un petit échantillon de textes de type journalistique et littéraire. Pour le moment, il compte à peu près 4000 mots, correspondant niveaux B1, B2 et C1 du Cadre Européen Commun de Référence (CECR). Il s'agit de textes authentiques de la presse et de romans utilisés dans d'anciens examens du Diplôme d'Études en Langue Française (DEL F) et du Diplôme Approfondi de Langue Française (DAL F), trouvés sur le site du Centre International d'Études Pédagogiques (CIEP)<sup>5</sup>.

## Annotation des textes

Afin de pouvoir exploiter ces textes permettant de réaliser les activités signalées plus haut, nous avons annoté de manière fine les informations langagières portées par les textes. Deux types d'annotations ont été faits (Figure 3) :

- **Des annotations automatiques** au moyen de l'analyseur Cordial (Synapse)<sup>6</sup> : il génère des annotations, un étiquetage morphologique et syntaxique des unités textuelles. Les informations ainsi repérées nous ont permis de mieux saisir les éléments à annoter manuellement.
- **Annotations manuelles** : ce type d'annotations nous permet d'isoler les informations langagières que les analyseurs TAL n'arrivent pas à repérer. Nous avons isolé des phénomènes linguistiques tels que les coréférents, les connecteurs, etc.



**Figure 3 : schéma du traitement du corpus textuel**

## Conclusion et perspectives

Comme nous avons pu observer tout au long de cet article, l'élaboration de notre système devient le produit de la mise en œuvre des compétences et de connaissances issues de plusieurs disciplines ou domaines : linguistique générale, linguistique textuelle, linguistique de corpus, didactique des langues, TAL, ingénierie pédagogique, informatique, ALAO. Notre projet renvoie donc à un dispositif pédagogique visant l'enrichissement de la formation en langues, issu de l'inter et de la pluridisciplinarité, ce qui mérite sans doute d'être mis en relief.

Pour le moment, nous en sommes au stade de la phase finale du développement de l'environnement informatique et de l'implémentation de ses fonctionnalités. Le corpus est presque prêt : il manque seulement l'évaluation, la vérification des annotations ainsi que le développement final des activités.

Nous comptons finir d'ici peu la maquette du système qui sera évaluée et testée au sein des universités qui ont montré de l'intérêt pour le projet (l'université d'Antioquia et l'université nationale de Colombie à Bogotá). Nous comptons par la suite recueillir les apports des utilisateurs afin de développer une version plus complète à proposer à d'autres universités qui pourraient être intéressées par ce type de dispositif d'aide à la formation des futurs enseignants de FLE.

## Références bibliographiques

Adam, J.-M. (1990). *Éléments de linguistique textuelle : Théorie et pratique de l'analyse textuelle*. Pierre Mardaga : Liège.

- Adam, J.-M. (2011). *La linguistique textuelle*. Collection Linguistique Coursus, troisième édition. Armand Colin : Paris.
- Alkhatib, M. (2012). La cohérence et la cohésion textuelles : problème linguistique ou pédagogique ? In *Didáctica. Lengua y Literatura*, 2012, Vol. 24 pp 45-64.
- Antoniadis, G. (2010). « De l'apport pertinent du TAL pour les systèmes d'ALAO. L'exemple du projet MIRTO ». *2e Congrès Mondial de Linguistique Française (CMLF-2010)*. 12-15 juillet 2010, La Nouvelle Orléans, USA.
- Chaudiron, S. (2007). « Technologies linguistiques et modes de représentation de l'information textuelle ». In *Documentaliste-Sciences de l'information*, 2007/1 Vol. 44 pp 30-39.
- Charaudeau, P. & Maingueneau, D. (2002). *Dictionnaire d'analyse du discours*. Éditions du Seuil : Paris.
- 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.
- Coseriu, 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.
- Develotte, C. & Mangenot, F. 2010. Former aux tutorats synchrone et asynchrone en langues. *Distances et savoirs*, vol. 8 - n°3/2010, p. 345-359.
- 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.
- Mangenot, F. 1998. « Outils textuels pour l'apprentissage de l'écriture en L1 et en L2 ». Publié dans *Pratiques discursives et acquisition des langues étrangères (Université de Franche-Comté, 1998), actes du 10e colloque international FOCAL*, Besançon, 19-21 septembre 1996.
- Molina, J. M. & Antoniadis, G. (2014). « Toward the Constitution of a Hybrid Learning Environment for the FFL Teacher's Training in Colombian Universities Based on Text Linguistics ». In QUIROZ, Gabriel & PATIÑO, Pedro (Éditeurs) : *LSP in Colombia: Advances and Challenges*. Chapitre 15, Volume 175, pp 233-249. Collection Linguistic Insights. Éditions Peter Lang : Bern, Berlin, Bruxelles, Frankfurt am Main, New York, Oxford, Wien.
- Narcy-Combes, J.-P. (2003). « Dans quelle mesure l'apprentissage relève-t-il d'un transfert ? », In *ASp*, N° 39-40, pp 107-117 [En ligne] <http://www.asp.revues.org/1326> (Consulté le 21 mai 2014).
- Salamanca Lamouroux, C. (2014). Une recherche exploratoire sur la formation initiale des enseignants colombiens de langues vivantes aux TICE. In *Adjectif : Analyses recherches sur les TICE*. [En ligne] <http://www.adjectif.net/spip/spip.php?article281> (Consulté le 21 mai 2014).



- 1 En mai 2011, nous avons procédé à interviewer des enseignants/formateurs et des étudiants de deux universités colombiennes. Ces entretiens nous ont permis de mieux cibler le questionnaire que nous avons rédigé en France et envoyé par la suite au travers d'Internet à la fin 2011.
- 2 Nous avons également regardé les programmes de formation de la plupart des universités colombiennes et nous n'avons pas trouvé de cours relatifs à la linguistique textuelle.
- 3 Bien que la notion de « phrase principale » ne soit pas intrinsèquement liée à la linguistique textuelle, nous considérons qu'elle a sa place au niveau de l'analyse de la structuration textuelle, en tant qu'élément permettant la cohérence des textes. Comme est dit dans (Alkhatib, 2012), l'utilisation de ce type de phrases doit permettre l'analyse et la création des textes à partir de l'emploi des phrases constituants des idées directrices ou principales, des arguments et des exemples.
- 4 Il restera cependant à être établi la situation de transfert en situation didactique vis-à-vis de ce projet de dispositif informatique. Nous considérons à l'instar de ce qui est dit dans (Narcy-Combes, 2003), le transfert en situation didactique dépend en grande mesure du rôle de l'enseignant au niveau de la médiation et le rôle de tous les acteurs au niveau du contrat pédagogique. Étant donné que dans notre système informatique la médiation est faite par l'enseignant/formateur et que les contenus soulevés au travers des questionnaires de recherche ont été négociés avec tous les acteurs (enseignants/formateurs et étudiants), nous considérons que ce transfert sera assez positif pour l'ensemble d'acteurs.
- 5 <http://www.ciep.fr/>
- 6 [http://www.synapse-fr.com/Cordial\\_Analyseur/Presentation\\_Cordial\\_Analyseur.htm](http://www.synapse-fr.com/Cordial_Analyseur/Presentation_Cordial_Analyseur.htm)