

Le TALN au service de la didactique du français langue étrangère écrit.

Isabelle Audras, Jean-Gabriel Ganascia

▶ **To cite this version:**

Isabelle Audras, Jean-Gabriel Ganascia. Le TALN au service de la didactique du français langue étrangère écrit.. 13ème conférence sur le Traitement Automatique des Langues Naturelles, Apr 2006, Leuven, Belgique. pp.825-834. edutice-00086950

HAL Id: edutice-00086950
https://edutice.hal.science/edutice-00086950

Submitted on 20 Jul 2006

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.

Le TALN au service de la didactique du français langue étrangère écrit

Isabelle Audras, Jean-Gabriel Ganascia

LIP 6 – Université Pierre et Marie Curie, 8 rue du Capitaine Scott 75015 Paris
Isabelle.Audras@lip6.fr
Jean-Gabriel.Ganascia@lip6.fr

Résumé

De nouveaux logiciels d'analyse textuelle tirent partie des progrès récents effectués en apprentissage symbolique et dans le traitement automatique des langues naturelles. Conçu au LIP6 par Jean-Gabriel Ganascia, le *Littératron* est l'un d'entre eux ; il extrait automatiquement des motifs syntaxiques[1] à partir de textes écrits en langage naturel. Plus exactement, le *Littératron* prend comme entrée un arbre d'analyse syntaxique et donne en sortie un certain nombre de motifs syntaxiques récurrents. Associé à un analyseur de textes, qui engendre l'arbre d'analyse syntaxique à partir de textes écrits en langage naturel, il révèle les singularités stylistiques de ces textes.

Nous allons voir qu'utilisé en sciences du langage, dans le domaine de l'acquisition du français écrit, le *Littératron* permet d'effectuer un diagnostic linguistique de l'apprenant, que celui-ci provienne d'une classe de langue hétérogène (différentes langues maternelles) ou homogène (une seule langue maternelle, en l'occurrence ici l'arabe). L'intérêt de cette approche concerne trois domaines : d'une part la didactique des langues, à titre éducatif ; d'autre part, la linguistique computationnelle, et enfin l'enseignement assisté par ordinateur.

Abstract

New text analysis softwares issued from fields of research such as Machine Learning and Natural Languages Processing prove to be relevant tools for the language sciences. *Littératron* is a new data-processing tool for the automatic extraction of syntactic patterns, designed at LIP6 by Jean-Gabriel Ganascia. Associated with a linear text analyser, it reveals the stylistic peculiarities of a text.

We will see that *Littératron* carries out a linguistic diagnosis of learners if used in language sciences, especially in the field of acquisition of written French as a foreign language. The learner can be from a heterogeneous group (various language levels and various mother tongues) or from a homogeneous group (only one language level and one mother tongue, here, Arabic). The interest of this approach is related to three fields: first, language didactics, on a purely educational basis; next, computational linguistics; finally, computer-assisted learning.

[1] Un motif syntaxique est une association d'unités linguistiques cohérentes

Mots-clés : acquisition d'une langue étrangère écrit, didactique de l'écrit en langue étrangère, TALN, extraction de motifs récurrents, stylistique, diagnostic linguistique

Keywords: foreign-language acquisition, foreign-language written didactic, NLP, stylistics, extraction of recurrent patterns, linguistic diagnosis

---

1. Cadre théorique

Les recherches sur l'acquisition de l'écrit en langue étrangère sont récentes. Cependant elles ont bénéficié des résultats des recherches concernant l'acquisition de l'écrit en langue maternelle. Prise de recul par rapport à la langue, aide à la mémorisation : les vertus cognitives du passage à l'écrit ne sont plus à démontrer (Mangenot 1998).

De plus les résultats des recherches en linguistique textuelle croisent également les intérêts de la didactique de l'écrit. En effet les notions de cohérence textuelle et de pragmatique, au centre de la linguistique textuelle, sont porteurs du développement chez l'apprenant d'une compétence textuelle qui lui rend disponibles des outils d'articulation en vue de construire un discours. Autrement dit, « les structures textuelles formelles [...] guident le scripteur dans la construction d'un texte et le lecteur dans sa compréhension" (Scardamalia & Bereiter, 1986)

Selon Tuffs (Tuffs, 1993), travailler sur des genres textuels différents facilite l'acquisition des langues étrangères. De façon générale, l'écrit en classe de langue est associé à une consigne qui prévoit l'intention de communication, même à l'extérieur d'un genre. En effet, le cadre narratif choisi, par le genre ou la consigne, définit un objectif de communication précis. Celui-ci appelle des objectifs fonctionnels dont l'expression morphosyntaxique et lexicale est vue en classe. Ce contenu linguistique, découvert à l'intérieur d'une situation de communication, est automatisé lors de réemplois, et ceci est d'autant plus vrai si celui-ci se trouve dans un contexte similaire. Enfin, l'analyse des besoins communicatifs du cadre narratif aide l'apprenant à s'adapter face à une nouvelle situation de communication dans laquelle il doit réagir (Tagliante, 1994).

Par ailleurs, l'apprentissage du FLE est sanctionné par une certification appelée DELF (Diplôme d'Etudes en Langue Française) aligné sur le cadre européen commun de référence dans l'apprentissage des langues. Les épreuves écrites A1, A2 et A3 ont pour cadre narratif, respectivement : la carte postale, la lettre amicale, la lettre de motivation.
Dans ces deux types de production écrite en français langue étrangère, le niveau de l'apprenant est validé par rapport à sa capacité à exprimer un message à travers un modèle appris et reconnu et non simplement par rapport à ses compétences grammaticales.

Autrement dit, la production écrite en classe de langue est le reflet des compétences de l'apprenant lors du passage à l'écrit. Ses compétences se révèlent à la fois dans la fréquence des expressions observées, dans ses prises de risques et dans l'originalité de ses idées (Carroll, M. & Stutterheim Ch., 1997).

C'est pourquoi nous souhaitons repérer, grâce aux techniques actuelles du traitement automatique des langues, les erreurs écrites usuelles d'une population d'apprenants, ce qui permettra de mettre l'accent, au cours de l'enseignement, sur la correction de ces erreurs.
Ce repérage des erreurs peut se faire soit dans l'absolu, par détection des erreurs syntaxiques, soit par rapport aux usages, par une étude des tournures propres à une catégorie d'apprenants dans un cadre narratif précis, celles-ci se trouvant absentes ou peu usitées chez les locuteurs

natifs. C'est cette seconde approche que nous avons adoptée, sachant que le rôle des enseignants de langue n'est pas d'enseigner une langue abstraite parfaite mais de transmettre les usages d'une langue.
Plus exactement, le travail présenté ici recourt à l'emploi d'outils d'analyse stylistique pour dégager les caractéristiques des apprenants, selon leur niveau, et les distinguer des locuteurs natifs. Des études empiriques conduites autour de trois populations d'apprenants, l'une à Paris, à l'Alliance Française, l'autre à l'université de Naplouse (Territoires Palestiniens), auprès d'un public arabophone et la troisième à l'Ecole Normale de Port-au-Prince (Haïti) auprès d'étudiants créolophones valident l'approche proposée.

2. Présentation des outils informatiques utilisés

Un motif syntaxique se définit comme une association d'unités linguistiques cohérentes, par exemple : [préposition + pronom personnel réfléchi + verbe à l'infinitif]. Chaque motif, par exemple le motif précédent extrait automatiquement à partir des outils mis en œuvre, peut recevoir un nombre plus ou moins grand de réalisation dans un texte donné. Voici un exemple des réalisations de ce motif dans des textes que nous avons traités : "de vous adresser", "afin de vous donner", "de m'investir", "de vous donner". Ces quatre réalisations ont été extraites ensemble d'un même groupe de scripteurs de lettres de motivation.

Deux outils informatiques sont nécessaires pour extraire les motifs syntaxiques caractéristiques des textes écrits par les différentes populations. Le premier outil informatique requis est un analyseur morphosyntaxique du français qui construit des arbres syntaxiques à partir de productions écrites. Nous avons eu recourt à l'analyseur linéaire avec dictionnaire partiel Vergne qui a été élaboré par Jacques Vergne de l'Université de Caen, en 1998 (Vergne, 2001). Le second est l'analyseur stylistique *Littératron*, mis au point au LIP6 par Jean-Gabriel Ganascia (Ganascia, 2001) qui dégage les motifs syntaxiques récurrents présents dans ces arbres.
Plus exactement, à chaque mot ou groupe de mots l'analyseur de Vergne associe une étiquette ; un arbre stratifié est donc une partition d'étiquettes dont les classes dépendent de la profondeur du nœud dans l'arbre d'analyse. Etant donnée une structure d'ASO, le *Littératron* calcule une mesure de similarité entre plusieurs ASO, fondée sur la notion de distance d'édition, et génère un graphe de similarité enregistrant les sous-arbres les plus proches de l'ASO donné en entrée.

C'est ce graphe de similarité qui sert ensuite d'entrée à l'algorithme de classification du *Littératron*, appelé 'centre-étoiles', qui construit des classes de motifs similaires et leur attribue un nom significatif. En effet, l'algorithme centre-étoile évalue d'abord l'ensemble des étoiles centrées sur les différents nœuds puis il prend, pour chacune, la somme des valeurs de similarité des nœuds de chaque étoile au centre. Une fois calculée la valeur de chaque étoile, l'algorithme 'centre-étoiles' prend celle qui à la plus forte évaluation. On marque ensuite, les nœuds qui appartiennent à cette première étoile, avant d'appliquer récursivement le même algorithme sur les nœuds non marqués, jusqu'à épuisement des nœuds non marqués.
En résumé, toute étoile est un sous-graphe du graphe de similarité centré sur un nœud. Pour chaque classe ainsi construite, l'algorithme choisit les motifs les plus similaires au centre de l'étoile, pour illustrer la signification de l'étoile. Il indique aussi le texte source couvert par chacun des motifs.

Voici l'exemple d'un centre d'étoile, illustré par la figure 1 :
[PREP ["de"]] + [GN [ART ["la"]] + [NOM ["forêt"]]] (texte : "de la forêt"), auquel sont associés les 5 motifs syntaxiques suivants :
- [PREP ["à"]] + [GN [ART ["l"]] + [NOM ["auberge"]]] (texte : "à l\'auberge ") ;
- [PREP ["d\"]] + [GN [ART ["un"]] + [NOM ["hiver"]]] (texte : "d\'un hiver") ;
- [PREP ["dans"]] + [GN [ART ["le"]] + [NOM ["monde"]]] (texte : "dans le monde") ;
- [PREP ["avec"]] + [GN [ART ["les"]] + [NOM ["chiens"]]] (texte : "avec les chiens") ;
- [PREP ["depuis"]] + [GN [ADJ ["quelques"]] + [NOM ["jours"]]] (texte : "depuis quelques jours").
Ceci signifie que la mesure de similarité entre le premier motif ('de la forêt') et l'un des arbres dérivés des arbres syntaxiques de chacun de ces cinq groupes nominaux est supérieure à un certain seuil. Ces cinq motifs font partie de la même étoile dont le centre est de la forme : [PREP ["de"]] + [GN [ART ["la"]] + [NOM ["forêt"]]].

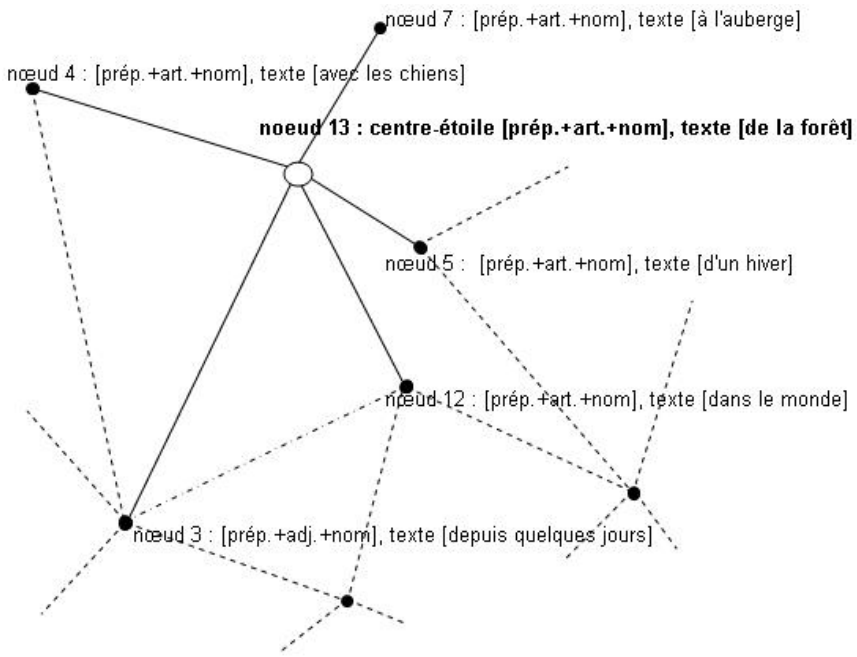


Figure 1 : Graphe du centre-étoile présenté en exemple.

Outre la construction d'étoiles et l'extraction de motifs, le *Littératron* procède à un second type d'opérations qui consiste à comparer les étoiles issues de plusieurs textes afin de repérer les étoiles présentes dans l'un et absentes dans l'autre. Ceci permet de discriminer, parmi les motifs présents dans une production, ceux qui le distinguent d'autres productions. C'est à partir de ce type de discrimination que l'on construira les tournures caractéristiques de populations d'apprenants.
Par exemple, voici l'une des sorties du *Littératron* analysant trois groupes de scripteurs distincts à partir de lettres de motivation (LM app : lettres de motivation d'apprenants de langues maternelles diverses, LM appa : lettre de motivation d'apprenants arabophones, LM frcph : lettres de motivation de francophones).

Patron N°46

- Indépendante
  S
  P p 1 s
  V
  COD
  N c f s
- Exemples:
  Fichier LMapp: Je parle la langue anglaise et française
  Fichier LMfrcph: Je maîtrise la mise en place de l'organisation de l'archivage
  Fichier LMappra: J'apprends la presse à l'université de Naplouse

Le patron décrit la structure syntaxique du motif extrait. Ici il s'agit d'une proposition indépendante de forme sujet S + verbe V + COD. Il y est précisé que le sujet est un pronom personnel à la première personne du singulier (Pp 1 s) et que le COD est un nom commun féminin singulier (Nc f s).
Les exemples extraits de chaque groupe de scripteurs donnent un aperçu des différents textes en langage naturel que le *Littératron* détecte comme étant proches de cette structure centrale.

3. Premier type d'expérience : les apprenants sont de langues maternelles diverses.

L'idée de cette recherche est de comparer des productions écrites en classe de F.L.E. de différents niveaux avec des productions de francophones répondant aux mêmes consignes. Les scripteurs francophones sont des natifs français de niveau d'étude au moins équivalent à bac+4. Une autre étude, qui pourrait se révéler intéressante, travaillerait avec des francophones d'un niveau d'études moins élevé, voire des débutants scripteurs adultes (Morais J. & Kolinsky R., 2001). L'idée de cette perspective est de montrer que le critère du niveau d'éducation n'est pas négligé dans cette approche.

3.1 Présentation des productions écrites et méthodologie expérimentale

Quatre types de production ont été choisis : la carte postale (CP), la lettre amicale (LA), la lettre de motivation (LM), la description (Des). Chaque production correspond à un niveau d'apprentissage du français langue étrangère. Quant à la description, chaque apprenant, tout niveau confondu, est soumis à l'observation puis à la description écrite d'un même dessin en couleurs de format A3 (place de village, art naïf).
Toutes les productions d'apprenants ont été faites en classe, entre le mois d'avril et le mois de juin 2002. La plupart se sont déroulées à l'Alliance Française de Paris. Certaines descriptions ont été réalisées dans une formation en FLE et en alphabétisation dans le Foyer de travailleurs Pinel, à Saint Denis.
Le tableau 1 a une double fonction. Premièrement, il récapitule les expérimentations réalisées par genre textuel. Par exemple : en ce qui concerne la 'carte postale' (CP), vont être introduits simultanément dans les analyseurs les productions d'apprenants débutants et de francophones. Deuxièmement, il détaille le nombre total de production de chaque type.
Concernant la description, les productions des 4 groupes de scripteurs sont introduites en même temps dans les analyseurs.

	apprenants			francophones
	débutants (niveau A1 du CECR[2])	intermédiaires (A2 niveau du CECR)	avancés (A3 niveau du CECR)	
carte postale (CP)	6			6
lettre amicale (LA)		4		4
lettre de motivation (LM)			6	6
Description (Des)	5	5	5	5

Tableau 1 **:** Tableau récapitulatif des productions et leur nombre.

3.2 Résultats et commentaires

Les résultats obtenus sont de nature statistique, auxquels nous ajoutons des commentaires linguistiques sur les motifs extraits.

	CP déb.	CP frcph.	LA inter.	LA frcph.	LM av.	LM frcph.	Des deb.	Des inter.	Des. av.	Des frcph
nb étoiles	6	10	2	5	6	6	2	3	3	13
% texte	50	50	60	30	25	17	33	33	35	14

Tableau 2 **:** Nombre d'étoiles et pourcentage de texte représenté par celles-ci.

Le tableau 2 ci-dessus donne les résultats numériques des calculs statistiques effectués par l'analyseur. Il indique, pour chaque classe de scripteurs (francophones : frcph ; apprenants débutants : deb ; apprenants intermédiaires : inter ; apprenants avancés : av) et pour chaque type de production, le nombre d'étoiles détectées par le *Littératron* ainsi que le pourcentage de texte représenté par ces étoiles. Les paramètres d'entraînement du *Littératron* sont identiques sur tous ces ensembles de productions, en particulier les seuillages de l'algorithme centre étoile et du graphe de similarité. Autrement dit, Le nombre d'étoiles détectées est donc un bon indicateur de la richesse stylistique : plus il y a d'étoiles, plus le style est riche, c'est-à-dire moins les automatismes prévalent. Il en va de même pour le pourcentage de texte couvert par les étoiles détectées : plus celui-ci est faible, plus les patrons varient, ce qui signifie que le style est plus riche.

[2] CECR : Cadre Européen Commun de Référence

Notons que cette notion de richesse stylistique doit être relativisée ; en effet, un grand écrivain pourrait se caractériser par la singularité d'un style qui déclinerait une palette restreinte de patrons, tandis qu'un écrivain sans style les déploierait tous. En dépit de ces quelques réserves, dans le cas particulier de la didactique qui nous intéresse, nous assimilons la richesse d'un texte (ou d'un ensemble de textes) au nombre de figures syntaxiques employées.

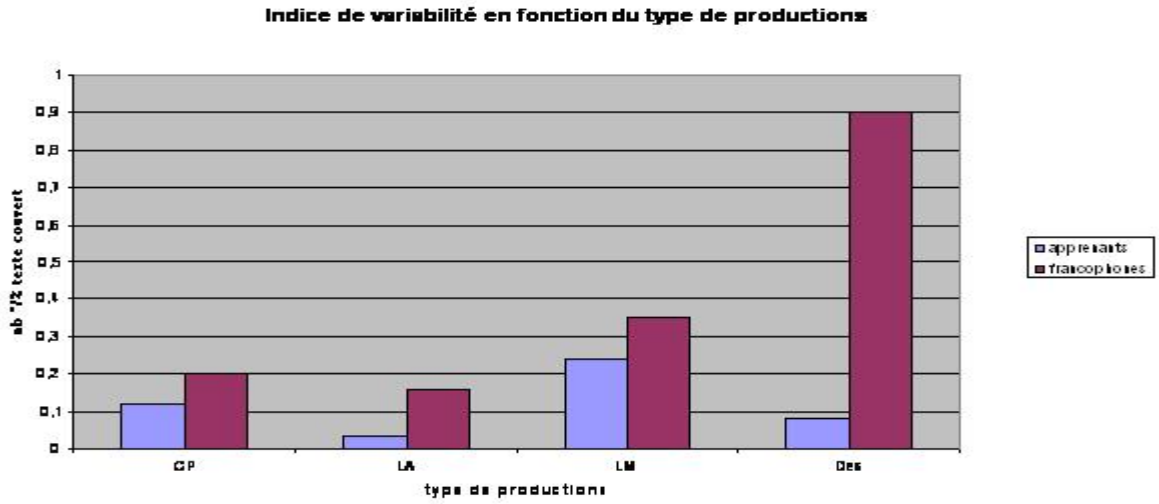


Figure 2 : Indice de variabilité en fonction du type de production.

Sur le graphe de la figure 2, nous définissons un indice de variabilité qui est, pour chaque type de texte, le rapport du nombre d'étoiles détectées sur le pourcentage de texte utilisé par l'application.
Que conclure du nombre de motifs syntaxiques récurrents et du pourcentage de texte recouvert par ces motifs ? D'une part, les résultats statistiques représentés par l'indice de variabilité nous montre que pour un même genre de production écrite, les motifs syntaxiques retenus par l'application sont plus nombreux, divers et dans une proportion de texte plus petite chez les francophones que chez les apprenants. De plus, la partie de texte non recouvert par les motifs syntaxiques récurrents varie dans un rapport 2 (pour les CP, LM et LA) à 9 (pour la Des) fois plus important chez les francophones que chez les apprenants, même les plus avancés. Cette partie de texte, où le *Littératron* n'a pas détecté de motifs récurrents, pourrait être utilisée pour définir l'originalité du scripteur.
Cette analyse a révélé des automatismes de l'écrit à l'intérieur de certains types de production. Ces automatismes concernent aussi bien des textes d'apprenants du français que ceux des francophones. Il y a donc des matrices d'écriture de cartes postales, de lettres d'invitation ou de lettres de motivation. Pour ce qui concerne les descriptions, la comparaison entre les différents niveaux fait apparaître des fréquences de motifs qui évoluent vers une complexification dans la composition et les liens de dépendance, donc une aisance d'écriture qui s'installe au fur et à mesure que la compétence morpho-syntaxique s'acquiert.
Enfin, concernant la description, nous sommes en mesure de rajouter quelques commentaires sur la structure syntaxique des motifs extraits. Les motifs de base extraits en qualité de syntagme nominal et en qualité de syntagme verbal ont, respectivement, la composition suivante : préposition + substantif + adjectif qualificatif et pronom sujet + verbe + adverbe.

Ces motifs de base s'enrichissent progressivement en fonction de la maîtrise du français écrit. Par exemple, le syntagme nominal "des paysages variés" issu d'une production d'un scripteur débutant évolue en "un paysage bien vert" chez un scripteur natif francophone.

4. Deuxième type d'expérience : les apprenants sont de même langue maternelle

L'hypothèse est de déceler une spécificité de la syntaxe française au sein d'une classe homogène.

4.1. Population arabophone

4.1.1. Présentation des productions écrites

L'ensemble des productions analysées correspond aux examens d'histoire et de civilisation du 1er semestre 2004 d'étudiants en 3ème année du département de français de l'Université An-Najah de Naplouse (Territoires Palestiniens). Les étudiants de l'université sont tous de langue maternelle arabe, l'anglais est leur première langue étrangère, le français leur deuxième.
Les productions de chaque type sont au nombre de dix.

4.1.2. Résultats et commentaires

Trois mêmes motifs syntaxiques ressortent systématiquement des productions arabophones. Ces motifs recouvrent ¼ du texte analysé. Il s'agit de deux motifs nominaux et d'un verbal. Les deux motifs nominaux sont de construction : DE + adjectif + nom, comme dans les exemples : « de choses magnifiques » et « d'autres villages » :

Motifs présents dans la constellation issue de ::atelier:civilisatioarabph.prs
et absents de la constellation issue de ::atelier: civilisatioarabfrph.prs
Motif N°151 (étoile N°7) :
V s1
P k ne
C V s1 fatigue
P l pas
Expression couverte : ne fatigue pas
Motif N°184 (étoile N°11) :
N fp3
I o de
C S fp3 choses
P E fp3 magnifiques
Expression couverte : de choses magnifiques
Motif N°183 (étoile N°13) :
N ms3
I o d'
P a ms3 autre
C S ms3 village
Expression couverte : d' autre village

Figure 3 : Sortie du *Littératron.*

Ce motif syntaxique révèle une utilisation massive de groupes nominaux compléments du verbe de forme adjectif + nom commençant par DE, au détriment d'autres articles et d'autres prépositions. L'étudiant, à défaut de connaître la bonne rection d'un verbe, directe ou indirecte, ou le bon emploi de l'article partitif sur le défini ou l'indéfini, va utiliser systématiquement la préposition DE pour introduire ses compléments d'objet.
L'apprenant ne semble pas maîtriser une bonne utilisation des prépositions et des articles.
Cette expérience révèle pour une classe homogène d'apprenants arabophones, une spécificité de la syntaxe française dont l'acquisition nécessite un accompagnement particulier.

4.2. Population créolophone

Une expérimentation en cours, sur des productions de même consigne, auprès d'étudiants de l'Ecole Normale de Port-au-Prince (Haïti) semble converger vers la même conclusion, avec une flagrante récurrence de motifs extraits de la forme : de + adjectif + nom.

5. Conclusion

Utilisé en sciences du langage dans le domaine de l'acquisition en langue étrangère du français écrit, le *Littératron* est en mesure d'effectuer un diagnostic linguistique de l'apprenant sur des productions au cadre narratif contraignant. Ce diagnostic sert à l'accompagnement à l'acquisition de l'écrit.

A terme, ce travail doit faire l'objet de deux types de développements complémentaires, sur les plans technique et expérimental.
D'un côté, nous nous sommes limités ici à une décomposition en syntagmes, et à une étude de la structure de la phrase relativement à cette décomposition. Cela restreint assez fortement le type de motifs détectés. Nous allons faire appel à une décomposition plus riche qui prendra en compte la structure propositionnelle. L'algorithme d'extraction de motifs est identique, mais l'analyse syntaxique diffère ; surtout, l'arbre résultant de cette analyse doit être considérablement enrichi.
D'un autre côté, les résultats obtenus auprès d'étudiants arabophones et créolophone nous encourage à poursuivre plus loin l'étude des différences spécifiques auprès d'apprenants venant de différentes régions du monde, et de langues maternelles diverses, sur des productions de DELF.

Références.

ADAM J.-M. (1992). *Les textes : types et prototypes. Récit, description, argumentation, explication et dialogue*. Paris : Nathan.

AUDRAS I. GANASCIA J.-G. (2005). Analyses comparatives de productions écrites d'apprenants du français et de locuteurs francophones, à l'aide d'outils d'extraction automatique du langage. *Apprentissage des Langues et Système d'Information et de Communication (ALSIC)*, 8, 81-94. http://alsic.u-strasbg.fr/v08/audras/alsic_v08_16-rec10.htm

BESSE J.-M. (dir.) (2003). *Qui est illettré ? Décrire et évaluer les difficultés à se servir de l'Ecrit*. Paris : Retz.

CARROLL M., STUTTERHEIM CH. (1997). Relations entre grammaticalisation et conceptualisation et implications sur l'acquisition d'une langue étrangère. *Acquisition et Interaction en Langue Etrangère (AILE)* 9, 14-19.

GANASCIA J.-G. (2001). Extraction automatique de motifs syntaxiques. Actes de *Traitement Automatique du Langage Naturel 2001* (*TALN 2001).*

GANASCIA J.-G. (2001). Extraction of Recurrent Patterns from Stratified Ordered Trees. Actes de *Machine Learning : 12th European Conference of Machine Learning 2001 (ECML 2001),* 167-179.

GANASCIA J.-G. (2004). Detection of Statistically Abnormal Patterns from Stratified Ordered Trees. In Milutinovic, Vujovic Milutinovic (dir.) *Advances in the Internet Technology, Concepts and Systems.*

GIGUET E. (1998). *Méthode pour l'analyse automatique de structures formelles sur documents multilingues.* Thèse de doctorat en informatique, Université de Caen. http://users.info.unicaen.fr/~giguet/these/

KAHN G. (dir.) (1993). Des pratiques de l'écrit. Numéro spécial *Le Français dans le Monde*. Paris : Hachette.

MANGENOT F. (1998). Outils textuels pour l'apprentissage de l'écriture en L1 et en L2**.** In *Pratiques discursives et acquisition des langues étrangères,* Souchon M. (dir.). Actes du *Xe colloque international "Acquisition d'une langue étrangère : perspectives et recherches",* 515-525.

MOIRAND S. (1990). *Une grammaire des textes et des dialogues*. Paris : Hachette FLE.

MORAIS J., KOLINSKY R. (2001). The literate mind and the universal human mind. In Dupoux E. (ed.) *Language, brain and cognitive development : Essays in Honor of Jacques Mehle.* Cambridge, Mass. : MIT. 463-480.

PERY-WOODLEY M.-P. (1993). *Les écrits dans l'apprentissage*. *Clés pour analyser les productions des apprenants*. Paris: Hachette FLE.

SCARDAMALIA M., BEREITER, C. (1986). Research on written composition. In *Handbook of research on teaching*, Wittrock, M.C. (dir.). New York : McMillan. 778-803.

TAGLIANTE C. (1994). *La classe de langue.* Paris : CLE International.

TUFFS R. (1993) A genre approach to writing in the second language classroom : the use of direct mail letters. *Revue belge de philology et d'histoire,* 71, 691-721.

VERGNE, J. (1999). *Etude et modélisation de la syntaxe des langues à l'aide de l'ordinateur. Analyse syntaxique non combinatoire Synthèse et résultats*. Habilitation à Diriger des Recherches, Université de Caen. http://users.info.unicaen.fr/~jvergne/#HDR.

VERGNE, J. (2001). Analyse syntaxique automatique de langues : du combinatoire au calculatoire. Actes de *Traitement Automatique du Langage Naturel 2001* (*TALN 2001). http://users.info.unicaen.fr/~jvergne/Taln2001FR_JV.pdf*