

# A navegação textual na didática da escrita : o exemplo das astruturas enumerativas

Fanny Rinck, Mustapha Mojahid

▶ **To cite this version:**

Fanny Rinck, Mustapha Mojahid. A navegação textual na didática da escrita : o exemplo das astruturas enumerativas. Linha d'Àgua, 2015, Col. "Linguagem e Tecnologia", Leitura, escrita e tecnologias da informação, Ana Elvira Luciano Gebara (28-1), pp.199-203. 10.11606/issn.2236-4242.v27i2p199-203 . hal-03190194

**HAL Id: hal-03190194**
**https://hal.science/hal-03190194**

Submitted on 24 Feb 2022

**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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

**Version auteur et en français**
**Référence :** Rinck F. et Mojahid M. (2015). A navegação textual na didática da escrita : o exemplo das astruturas enumerativas. In A.L. Tinoco Cabral, J.-L. Minel, S.C. Marquesi (eds) *Leitura, escrita e tecnologias da informação*, São Paolo, Terractota, Col. Linguagem e Tecnologia, p. 139-176. (La navigation textuelle en didactique de l'écrit : l'exemple des structures énumératives - Traduction de Silvia Augusta A.B. Bachur).

# La navigation textuelle en didactique de l'écrit : l'exemple des structures énumératives.

Fanny Rinck, IUFM de Grenoble et Laboratoire Lidilem, EA609, Université Grenoble Alpes.
Fanny.Rinck@ujf-grenoble.fr
Mustapha Mojahid, IRIT, UMR5505, CNRS/Université Paul Sabatier, Toulouse
Mustapha.Mojahid@irit.fr

## Introduction

Nous proposons dans ce chapitre une démarche pour la formation à l'écrit auprès d'étudiants et d'adultes, basée sur l'observation de corpus comme outil de prise de conscience pour l'aide à la compréhension et à la production de textes : l'enjeu est de favoriser la réflexion métalinguistique des apprenants pour développer leurs compétences littéraciques. Nous nous centrons sur une structure textuelle transversale aux genres et omniprésente dans les écrits académiques et professionnels : l'énumération. Notre objectif est de montrer en quoi la description de ce type de structure peut servir de fondement à un scénario didactique (ou plusieurs) intégré(s) au logiciel Navilire. Dans la perspective du développement de ressources pour la formation à l'écrit, il s'agit d'interroger le « teaching/learning with corpora » du point de vue de l'annotation des textes et des évolutions possibles d'un logiciel comme Navilire en regard des avancées de la linguistique de corpus et des contraintes de l'utilisation des technologies dans l'enseignement.
Dans une première partie, nous présentons notre cadre de référence, celui de la formation à l'écrit, en montrant l'intérêt de la navigation dans les textes comme outil didactique pour une formation basée sur le principe de développer la conscience métalinguistique des apprenants. Dans une seconde partie dédiée aux structures énumératives, nous nous attachons à leur description linguistique et montrons leur intérêt dans le cadre de la formation à l'écrit. Nous proposons alors dans une troisième partie les grandes lignes d'un scénario didactique visant la prise de conscience des caractéristiques des structures énumératives à partir d'une observation de corpus qui procède par étapes en fonction des critères d'analyse de ces structures issus des travaux de description linguistique. Nous discutons dans une dernière partie les annotations textuelles et discursives de corpus et les évolutions des corpus annotés et de la navigation textuelle en tant que ressources pour l'enseignement.

## 1. Formation à l'écrit et navigation didactique dans les textes

### 1.1. Les compétences littéraciques d'un point de vue linguistique

La notion de littéracie désigne la diversité des pratiques de lecture et d'écriture, des compétences qu'elles mobilisent et des contextes où elles s'exercent et permet ainsi de les envisager comme des acquisitions tout au long de la vie. Face aux écrits numériques et à l'omniprésence de l'écrit dans nos sociétés, il nous faut apprendre à lire et écrire dans des situations diverses et en évolution (dans le monde académique et, par la suite, dans des situations de travail), penser et agir à travers l'écrit. En écho aux travaux fondateurs de J. Goody sur la *Raison graphique* (1979), c'est ce que met en avant la

définition que donne l'Unesco de la littéracie : « It is key to communication and learning of all kinds and a fundamental condition of access to today's knowledge Societies »[1].
La maîtrise de l'écrit dans une société de la connaissance peut se définir à l'instar des multinationales qui se dotent de « bonnes pratiques » pour standardiser et rationaliser leurs documents, et du champ émergeant de la rédactologie comme science de la rédaction « experte ». Elle ne concerne donc pas le seul fait d'avoir une bonne orthographe mais appelle une attention générale à l'écrit. Il s'agit de questionner le développement littéracique aux deux échelles des individus et de la société, ainsi que les problèmes spécifiques à certains publics ou à certains domaines d'activité, les besoins de formation voire, dans une acception large, d'éducation ; ainsi de la *littéracie critique* qui, en référence à J. Dewey (1916) ou P. Freire (1970), est le titre d'un programme éducatif basé sur la capacité à décrypter un monde complexe et à agir en citoyen éclairé.

Sans prétendre couvrir les enjeux d'une formation à la littéracie et par la littéracie tout au long de la vie, notre approche relève de la didactique de l'écrit et prend comme point de départ les constats sur les difficultés à l'écrit en contexte universitaire et professionnel : ils font état de besoins de formation y compris chez des scripteurs et lecteurs qualifiés - qui autrement dit ne sont pas illettrés, sans être des « lettrés ». Pour répondre aux besoins de formation chez ce type de public (littéracie dite « avancée »), notre proposition est de mettre à profit les apports de l'analyse de textes, de l'analyse de discours et de la linguistique de l'écrit ; alors que la rhétorique fournit des modèles de texte et la grammaire des règles du bon usage de la langue, ces domaines peuvent permettre de bâtir une formation visant le développement de compétences de lecture et d'écriture à travers une conscience de l'écrit fondée sur la matérialité des textes : il s'agit de favoriser la réflexion métalinguistique des apprenants sur les caractéristiques des textes dans une perspective d'aide à la compréhension et à la production.
La démarche de formation peut consister à travailler sur un genre spécifique (la lettre de motivation, la dissertation, la thèse dans tel domaine disciplinaire) et à cerner les caractéristiques de ce genre, en tenant compte notamment de la variation au sein d'un même genre, voire de l'hybridité (i.e. les frontières entre genres). Elle peut consister aussi à travailler sur une entrée transversale aux genres (c'est le cas par exemple dans une formation à l'orthographe), et de voir éventuellement comment elle s'actualise dans différents genres ou dans différentes séquences textuelles (par exemple avec la notion de cohérence des textes). Ces entrées « transversales » vont du matériau graphique aux effets pragmatiques de l'écrit : on peut proposer une séquence sur la ponctuation comme sur la notion de destinataire, étant entendu que l'attention sera portée aux textes dans les deux cas.
La notion de conscience de l'écrit peut être comprise en référence aux notions de réflexivité et de « language awareness ». Plutôt qu'une compétence générale de l'ordre de la méta-cognition, consistant pour le sujet à s'interroger sur son activité cognitive et à exercer ainsi une forme de régulation, sinon de contrôle sur son activité, nous visons des compétences de l'ordre de l'analyse linguistique. Comme le souligne M. Fayol à propos des compétences en production de texte, et de l'importance de la révision au cours du processus de production : « la détection, le diagnostic et la décision d'intervenir (de réviser, reprendre) ne suffisent pas. L'auteur doit disposer de moyens tactiques pour intervenir » (Fayol, 1997 : 135) – et ne serait-ce, pourrait-on ajouter, que pour la « détection ». Comme le souligne C. Beaudet (2002 : 9) : « le sens textuel est le résultat, entre autres, de diverses opérations de langage et de discours (...) sur lesquelles le rédacteur peut agir pour améliorer la clarté de son discours ».
C'est dans cette perspective de doter les apprenants d'outils stratégiques pour la lecture et y compris celle de leurs propres textes que nous proposons une didactique de l'écrit basée sur le principe de former des « apprentis-linguistes », au sens où il s'agit de favoriser la réflexion méta-linguistique des apprenants à partir d'observations de corpus, à la manière du linguiste.

[1] Unesco, the global literacy challenge http://unesdoc.unesco.org/images/0016/001631/163170e.pdf

Ce projet se traduit par exemple par des recherches sur la construction d'un point de vue dans un texte et la manière dont on peut guider les lecteurs pour qu'ils apprennent à repérer ce point de vue sur la base de phénomènes de modalisation et de discours rapporté (Rinck et Lebarbé, 2005). Des guillemets, aussi anodins qu'ils puissent paraître, sont porteurs de sens et peuvent donner lieu à un questionnement du lecteur interpellé par leur utilisation, comme du scripteur qui se demande s'il doit en mettre ou pas. Une démarche-test a été menée auprès d'étudiants de sciences du langage, dans le cadre d'un cours sur les approches énonciatives, qui visait à susciter leur réflexion sur les difficultés liées à l'interprétation des passages entre guillemets et sur les notions de dénomination, de modalisation, de discours rapporté, de dialogisme (Rinck, Grossmann et Boch, à par. ; Rinck et Tutin, 2007).
Sans aller jusque là auprès d'un public plus large, on peut penser que l'usage des guillemets dans la presse, ou l'usage abondant qu'en font les étudiants dans leurs écrits en font une entrée intéressante dans une formation à l'écrit, aussi bien dans une perspective de remédiation pour un public jugé faible que dans la perspective de former des correcteurs ou des rédacteurs professionnels.

**1.2. Des ressources didactiques basées sur les corpus**

Dans le cadre esquissé ci-dessus d'une formation de scripteurs et lecteurs attentifs à la matérialité des textes, les principes du recours aux corpus dans l'enseignement des langues (Granger et al., 2002 ; O'Keeffe et al., 2007 ; Sinclair, 2004 ; Tribble, 1997 ; Wichmann et al., 1997) nous servent de point d'appui : les hypothèses d'ordre cognitif sur lesquelles ils se fondent mettent en exergue l'intérêt d'une approche inductive des faits de langue sur la base d'une observation guidée et outillée, à même de susciter des prises de conscience sur les textes et sur les modes de construction du sens.
Deux composantes semblent alors essentielles dans la mise en place d'un scénario didactique de formation à l'écrit :
-le corpus sur lequel vont porter les observations,
-la description qui en est faite, dans la littérature scientifique, en amont pourrait-on dire, et, en aval, en guise d'aboutissement de l'intervention didactique, à travers notamment la phase que l'on peut qualifier d'institutionnalisation des savoirs : comme dans d'autres approches dites inductives, telle le débat en classe, il est recommandé d'aménager un temps didactique permettant de donner sa place au savoir enseigné (la « leçon » pourrait-on dire), et pas seulement au savoir en construction qui caractérise le processus d'apprentissage.
A partir de ces deux composantes, deux types de stratégies sont possibles qui vont jouer sur l'autonomie (vs l'hétéronomie) de l'apprenant dans sa démarche d'observation :
-en ce qui concerne le corpus, il peut être demandé aux apprenants eux-mêmes de travailler sur sa constitution, autrement dit sur la collecte du matériau et sur la structuration du corpus (par ex., a-t-on affaire à des genres de discours variés, voire à des séquences textuelles (narratif, descriptif...) variées ? A l'inverse, le scénario peut se baser sur un corpus mis à disposition par l'enseignant, en fonction de ses objectifs.
-en ce qui concerne la description, il peut s'agir pour les apprenants de décrire le corpus selon des critères d'analyse qui leur sont fournis ou de ne se baser d'abord que sur leurs propres observations pour mettre en évidence par eux-mêmes des critères d'analyse. On voit ici que la frontière entre une démarche inductive et déductive n'a rien d'évident : donner certains critères n'interdit pas de faire construire aux apprenants des catégories d'analyse et laisser les catégories émerger n'exclut pas de (voire appelle à) les commenter, les faire bouger, bref, re-catégoriser.
Si l'on aboutit donc à quatre possibilités de scénario, selon que le corpus et la description sont plutôt du côté de l'enseignant ou de l'apprenant (corpus de l'apprenant vs corpus de l'enseignant et description par l'apprenant vs description par l'enseignant), ce ne sont donc là que des points de repères dans un ensemble varié de stratégies possibles.

Les avancées de la linguistique de corpus permettent aujourd'hui de nouvelles possibilités dans le champ de l'enseignement et de l'apprentissage des langues (outils pour la classe, outils d'auto-formation). Sont concernées en premier lieu les approches dites communicatives, basées sur des textes et des situations de production et de compréhension authentiques. Cependant, le regain d'intérêt, dans le cadre de telles approches, pour des questions d'ordre grammatical a également à y gagner.

L'exemple le plus emblématique du « teaching and learning with corpora » est celui des concordanciers (par ex. Tribble et al., 1997). Ils fournissent sous forme de liste un ensemble d'occurrences prises dans leur contexte proche et sont utilisés pour le développement de compétences lexicales et phraséologiques. Nous pouvons prendre comme exemple le projet Scientext[2] piloté par le laboratoire LIDILEM (Laboratoire de Linguistique et de Didactique des Langues Etrangères et Maternelles) à l'Université Grenoble 3. Il a permis de constituer un vaste corpus d'écrits universitaires et scientifiques, en français et en anglais, dans différentes disciplines et dans différents genre (thèse, articles, actes de colloques, écrits d'étudiants). Ce corpus représente un outil de recherche pour analyser l'écriture scientifique, les modes de raisonnement et la manière dont le chercheur positionne sa recherche, ou encore les difficultés des étudiants dans ce type d'écrits, mais il se veut aussi un outil didactique, pour les formateurs ou pour les étudiants en auto-formation. Le corpus est « annoté au niveau de l'en-tête (méta-données sur l'annotation et l'origine du corpus), au niveau de la structure (parties de l'article) et au plan morphologique et syntaxique à l'aide du logiciel Syntex développé par Didier Bourigault, et en respectant les recommandations XML et TEI en vigueur »[3]. Ces annotations permettent d'effectuer des requêtes sur tout ou partie du corpus, notamment dans une perspective didactique, pour travailler par exemple sur le terme d' « hypothèse » et voir avec quels verbes et adjectifs il s'associe préférentiellement. Les concordances du terme peuvent ainsi être envisagées comme un outil d'aide à la rédaction.

Scientext peut également donner lieu à des approches à l'échelle du texte (par ex. les introductions d'articles dans telle discipline). Le domaine des annotations textuelles et discursives permet aujourd'hui de compléter les approches didactiques basées sur les concordanciers par des approches basées sur des textes (par ex. Tribble, 2003). C'est le principe de la navigation textuelle comme outil didactique, principe au fondement du logiciel Navilire.

### 1.3. Navilire et la navigation didactique dans les textes

Le logiciel Navilire[4] (Lundquist, Minel et Couto, 2006) et la démarche didactique de navigation dans les textes (Lundquist, 2008, 2013) ont pour objectif de mettre à profit les apports de la linguistique textuelle pour l'apprentissage des langues étrangères et pour l'aide à la production de textes, notamment au départ autour de la cohérence textuelle :

> « Si la cohérence textuelle constitue bien une « heuristique générale » d'interprétation textuelle (Charolles, 1981), il est possible de l'exploiter pour montrer, et enseigner, comment la cohérence se manifeste linguistiquement dans les textes en général, et dans des types de textes différents en particulier, de même que dans des langues différentes. Il nous semble donc que l'enseignement de la linguistique textuelle contribue, dans de multiples contextes, à aiguiser l'attention des apprenants vers la réalisation de la « bonne formation » des textes, et à stimuler leur propre production de textes bien formés. » (Couto, Lundquist, Minel, 2005a : 2).

A partir de la « plate-forme informatique de Navigation Textuelle » développée dans le cadre de sa thèse par J. Couto (2006), le logiciel Navilire a été conçu comme un « outil qui permette à l'élève/l'étudiant de voir et de naviguer dans le texte entre des unités textuelles assurant la cohérence » (Couto, Lundquist, Minel, 2005b : 2). Le principe du logiciel est le suivant : des textes sont annotés par l'enseignant linguiste et servent de ressource pour l'apprenant qui doit identifier les

---

[2] Voir le site http://scientext.msh-alpes.fr

[3] Voir le résumé du projet sur le site indiqué dans la note précédente.

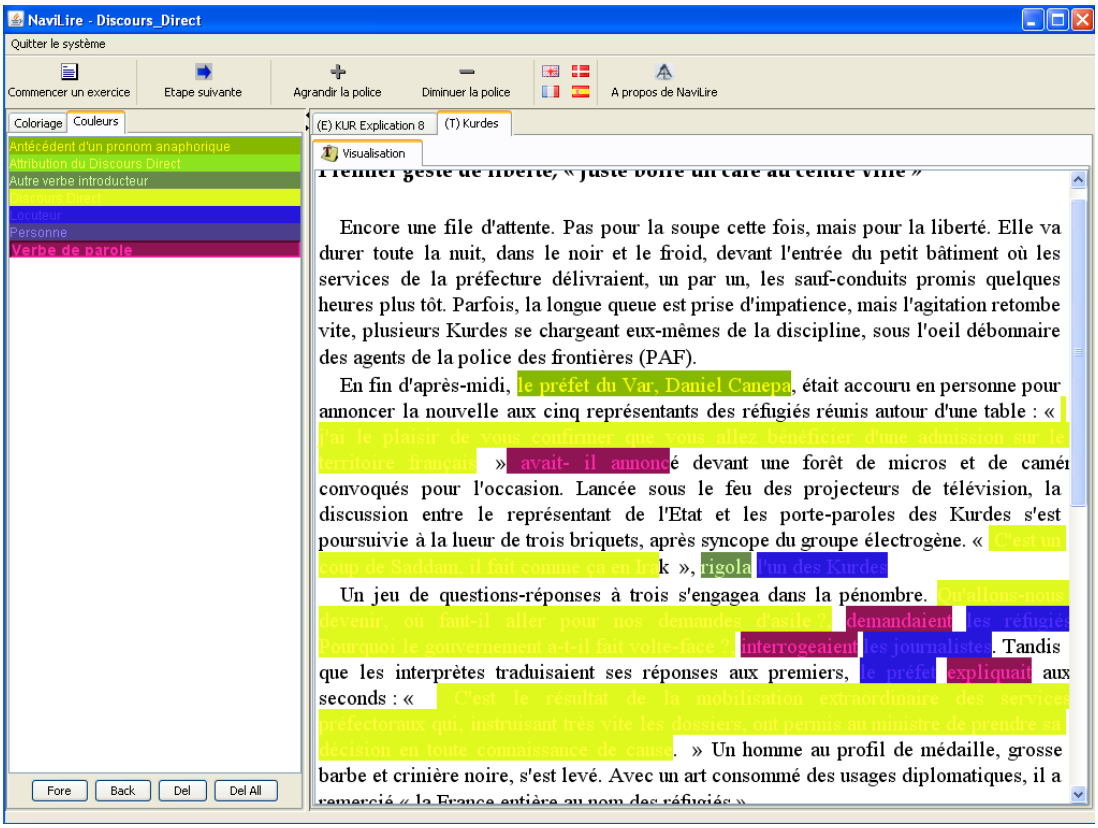
[4] Le logiciel (incluant des exercices) est téléchargeable sur http://www.jeanlucminel.fr/navitexte.html, ainsi que le logiciel NoteTool pour les annotations.

marques linguistiques et les relations de cohérence. Un scénario didactique guide l'observation des apprenants : ils sont amenés à travailler sur divers types de textes et sur les caractéristiques qui ont fait l'objet d'annotations (par ex., les anaphores en « le + nom » et « ce + nom »). Un système de pointage du texte permet à l'apprenant d'identifier les caractéristiques qu'on lui demande ; le segment correctement identifié apparaît en surbrillance, ce qui offre à l'apprenant un feed-back. Le feed-back demande aussi à l'apprenant d'approfondir son analyse : par ex., dans le cas des anaphores, un message apparaît et demande à l'apprenant de vérifier, par pointage sur le texte, les anaphores précédentes, identifier le passage que reprend l'anaphore, et l'éventuel déclencheur d'une anaphore, comme ci-dessous, dans l'un des exercices développé sous Navilire par L. Lundquist :

> « Dire qu'une préposition introduit un groupe prépositionnel semble une lapalissade. Il existe pourtant un certain nombre de cas pour lesquels cette évidence doit être mise en doute »

Dans cet extrait, « cette évidence » reprend « dire que... » et où le terme « évidence » est à mettre en relation avec le terme « lapalissade » considéré de fait comme un déclencheur.

L'exercice prend sens à l'échelle du texte : comme le montre la vue donnée ci-dessous, la fonction de visualisation fait apparaître grâce à un jeu de couleurs dans l'ensemble du texte les différentes annotations qui ont fait l'objet des différentes étapes de l'exercice. L'apprenant peut également manipuler cette représentation du texte : choisir par un simple clic de faire apparaître ou disparaître telle ou telle catégorie d'annotations. Un travail sur les anaphores permet par exemple, à l'échelle textuelle, de repérer des chaînes thématiques.




Vue du logiciel Navilire (exercice sur le discours rapporté développé par F. Sitri et C. Mellet, voir Desoutter et Mellet, à par.).

La cohérence textuelle n'est pas le seul aspect concerné. D'autres caractéristiques comme le discours rapporté par exemple (comme dans la vue donnée ci-dessus) font l'objet de réflexions didactiques et de développements d'exercices avec le logiciel Navilire (Desoutter C. et Mellet C., eds, à par.). L'entrée retenue peut également être de l'ordre de l'argumentation ou de la structure rhétorique

plutôt que de l'ordre d'une grammaire textuelle, par exemple en ce qui concerne l'écriture d'une introduction d'article de recherche et la manière dont le chercheur doit « créer une niche » dans l'introduction (Swales, 1990).
Les annotations de corpus qui servent de fondement à une démarche didactique de navigation textuelle se basent sur une approche à la fois locale des unités linguistiques, constitutives de structures de plus haut niveau, et une approche globale du texte comme unité. Elles impliquent de disposer de descriptions linguistiques sur les unités, sur leurs relations et sur la textualité d'une part et qui soient investies dans des genres et des textes en particulier d'autre part.

***

Dans ce qui suit, nous présentons une démarche basée sur la description linguistique des structures énumératives et qui vise à développer un scénario didactique dont l'objectif est le suivant : faire prendre conscience aux apprenants, à travers l'observation de corpus, des caractéristiques de ces structures, selon les principes de la navigation textuelle en didactique. Nous avons fait le choix de ces structures pour leur intérêt en termes de matérialité de l'écrit, de cohérence textuelle et de lisibilité et intelligibilité au niveau pragmatique.

## 2. Les structures énumératives

La démarche que nous adoptons dans une perspective de formation à l'écrit pose l'importance de la description linguistique comme préalable à l'élaboration de ressources didactiques. Nous avons choisi de nous intéresser aux structures énumératives qui sont relativement bien décrites d'un point de vue linguistique et psycholinguistique en particulier depuis les travaux menés aux laboratoires CLLE-ERSS et IRIT à Toulouse, sur lesquels nous allons revenir. Ces structures nous semblent représenter une entrée intéressante pour la formation à l'écrit car il s'agit d'une compétence transversale aux genres et d'une caractéristique omniprésente dans les textes, à même de susciter des prises de conscience quant à l'ordre scriptural, de la matérialité de l'écrit à ses enjeux pragmatiques.

### 2.1. Intérêt des structures énumératives dans l'ordre scriptural et dans une formation à l'écrit

Comme l'ont mis en évidence les travaux de J. Goody sur la « Raison graphique », la liste est aux fondements de nos pratiques de l'écrit et elle est emblématique de la rationalité propre à l'ordre scriptural. Compter, inventorier, classer sont les premières opérations par lesquelles nos sociétés sont entrées dans l'écrit. La liste s'oppose « à la continuité, la fluidité, la connexité propres aux formes ordinaires du langage parlé », elle en disjoint les éléments et les traite abstraitement (1979 : 105). Citant ces propos de J. Goody, J.F. Jeandillou (2008) souligne que la liste fait texte tout en montrant en quoi elle questionne les modèles d'analyse de la textualité et en particulier les principes de cohérence des textes.
L'importance de la dimension visuelle des textes d'un point de vue cognitif est bien mise en évidence (Goody, 1979, 1986, Olson, 1994, Ong, 1988) et a fait l'objet de recherches théoriques et empiriques en linguistique de l'écrit, domaine basé sur une conception du scriptural comme ordre langagier spécifique[5]. Le visuel entre dans le sens des textes et représente une compétence mise en œuvre par les lecteurs, comme le montrent par ex. les données de J. Virbel et al. (2005) sur des lecteurs experts et sur des enfants de 11 ans en cours d'acquisition de cette compétence.
Ce que nous désignons par « structures énumératives » combine des propriétés visuelles liées à la liste au sens de disposition spatiale des éléments inscrits sur une page et des propriétés linguistiques liées à l'énumération, de l'ordre du lexique, de la syntaxe, de la linguistique de texte (par ex. les connecteurs) et des relations sémantiques entre les constituants. Les deux dimensions du visuel et du verbal vont de pair, au point qu'on peut en discuter les frontières (par exemple entre la

[5] Voir par ex. les notions de visuo-texte chez J. Anis (1997) ou celle d'aire scripturale chez J. Peytard (1975).

ponctuation d'une part et les puces, numéros et tirets)[6]. Par ailleurs, ces structures exploitent aussi bien la linéarité de l'écrit, du fait de la co-énumérabilité des items qui les constituent (la liste est une suite) que l'organisation hiérarchique et logique du texte.
Un ensemble de marqueurs diversement désignés (marqueurs d'intégration linéaires (Turco et Coltier, 1988), marqueurs configurationnels (Charolles, 1994), prédicats architecturaux ou marques de signalisation de l'architecture textuelle (Péry-Woodley, 2001)) servent d'instruction dans la construction du sens et dans l'action, d'où l'affinité des structures énumératives avec les textes dits procéduraux (Garcia-Debanc, eds, 2001) (ou instructionnels-prescriptifs, chez J.M. Adam, 2001) et avec les critères de lisibilité, d'intelligibilité et de clarté, mis en exergue à propos plus largement des écrits techniques ou des écrits de travail dans le domaine de la rédactologie ou science de la rédaction experte (Beaudet, 2000, 2002 ; Labasse, 1999).
Dans le cadre d'une approche anthropologique attentive aux écrits ordinaires, on soulignera que ce type de structures est omniprésent dans tout type de texte. La liste de courses, l'annuaire, les trois atouts de tel produit vantés par telle affiche publicitaire en sont des exemples. Le critère de co-énumérabilité (Virbel et al., 2005) pose la question de l'extension de la notion : peut-on parler de structure énumérative dans « luxe, calme et volupté » ? Toute structure visuelle telle qu'un menu déroulant ou une arborescence, comme sur le web, peut-elle s'assimiler à une structure énumérative ?
Suite, série, inventaire, catalogue, l'énumération d'items relève de l'accumulation verbale et peut être traitée comme structure en tant qu'espace textuel, autonome ou intégré à un texte englobant. Pour des apprenants en formation à l'écrit, l'objectif n'est pas tant d'aboutir à une catégorie linguistique clairement circonscrite que de susciter ce type de questionnement pour montrer en quoi on a affaire à une structure aux fondements des pratiques ordinaires de l'écrit, mobilisée sous des formes variées dans une grande diversité de textes et de genres.
Dans le cadre d'une didactique visant le développement d'une conscience de l'écrit, les structures énumératives permettent d'attirer l'attention sur la matérialité des textes dans ses aspects micro liés au choix de tel ou tel signe (une marque de ponctuation par ex.) et macro (le genre du dictionnaire par ex.). La diversité des formes que prennent les structures énumératives est intéressante pour travailler sur la textualité de manière transversale aux genres. L'observation peut viser la construction d'une typologie ou, du moins, de critères d'analyse et il s'agit également d'envisager les caractéristiques des structures énumératives en regard de la production et de l'interprétation des textes, comme des ressources pour le scripteur et le lecteur.

### 2.2. Description des structures énumératives

Nous nous appuyons sur des travaux menés notamment aux laboratoires CLLE-ERSS et IRIT à Toulouse pour disposer de critères permettant de systématiser l'observation de structures énumératives et de rendre compte de leur diversité.
La première caractéristique qui fait la spécificité des structures énumératives d'un point de vue langagier est qu'elles appellent à une conception autonome de « l'ordre scriptural » (Peytard, 1970) et à une linguistique propre à l'écrit (Anis, 1988) : d'une part de nombreux aspects de la parole ne sont pas représentés à l'écrit (rythme, mélodie, accent), ou de manière ad hoc, et seule une partie de la typologie expressive (Jubert, 2007) peut être envisagée en termes d'équivalence avec la réalité sonore de la langue ; d'autre part, à l'écrit, le support, la page, et les propriétés morphologiques des caractères (taille, italiques etc.) fondent en propre l'écrit dans la cognition spatiale.
Par ailleurs, du point de vue des genres de discours et des activités de lecture et d'écriture, les structures énumératives sont caractéristiques des textes « à consigne » (Bouraoui, 2001) comme on l'a dit et leur usage semble se renforcer avec le renouvellement apporté par la modalité numérisée du texte (Virbel et al., 2005). Les recherches menées sur la réalité psychologique de ces structures

[6] Comme c'est le cas aussi avec d'autres objets graphiques tels que les tags et autres graffitis, ou dans les écrits numériques avec l'exemple des émoticônes.

mettent en évidence le rôle de la mise en forme matérielle (Pascual, 1991) dans l'élaboration de la représentation mentale des textes (Schmid, 2001). M. Péry-Woodley (2001) parle à ce propos de prédicats architecturaux en tant que modes d'organisation et de signalisation dans les textes procéduraux : la mise en forme matérielle apporte des informations utiles au lecteur et détermine les stratégies de lecture (Baccino et Pynte, 1998). Les blancs graphiques y compris représentent des indices sur lesquels prennent appui les lecteurs (études menées sur les mouvements oculaires) et qui facilitent la mémorisation (tâches de rappel) ou la recherche d'information. J.-L. Nespoulous et J. Virbel (2007) ont ainsi montré que l'univocité et l'équivalence fonctionnelle des items sont nécessaires dans certaines conditions pour les textes à consignes et peuvent influencer l'efficacité et déterminer une utilisation optimale (par exemple en terme de rappel et de recherche d'information).

Pascual (1991) a défini l'énumération de la manière suivante : « énumérer c'est attribuer un niveau égal d'importance aux entités et classer ces entités selon des critères variables ». Cette définition est certes conforme à de larges cas d'énumérations : les items sont équivalents fonctionnellement et sont réalisés à l'aide des propriétés de mise en forme matérielle équivalents, ce qui fonde la co-énumérabilité. Cependant, Virbel (1999), en constituant son corpus, se heurte à de nombreux cas contredisant cette définition qui se focalise sur le parallélisme entre fonction et formatage dans les énumérations.
Par exemple dans :

> « est considérée comme lecture savante (...) une pratique de lecture répondant aux critères suivants :
> - c'est une lecture qualifiée,
> - qui se développe sur le temps (...),
> - dans un parcours forcément individualisé (...) »

Les items sont identifiables comme tels d'un point de vue visuel, mais ils sont en relation de dépendance : on a affaire à un non-parallélisme au sens où ils ne sont pas fonctionnellement équivalents (exemple cité dans Virbel et al., 2005).
Ces énumérations ont amené J. Virbel (1999) à concevoir une nouvelle définition de la structure énumérative : l'acte du discours impliqué dans l'énumération est réalisé par un ensemble de propriétés linguistiques et visuelles qui ne peuvent être dissociées. Autrement dit, comme le souligne M. Péry-Woodley dans la suite de ces travaux sur l'énumération, les marques aussi bien lexico-syntaxiques que visuelles sont à intégrer à un modèle linguistique des énumérations. Elles remplissent la même fonction «  d'actes de discours spécifiques de la construction textuelle, faisant appel à des performatifs métalinguistiques dont la performativité est dirigée vers le texte lui-même » (Péry-Woodley, 2001 : 30).
Les marqueurs relèvent ainsi d'un métalangage textuel et le modèle de l'architecture textuelle développé dans les travaux de J. Virbel et E. Pascual à propos notamment des définitions et des textes à consigne (Pascual et Péry-Woodley, 1999 ;  Garcia-Debanc et Grandaty, 2001 ; Luc, Mojahid et Virbel 2002 ; Virbel et al., 2005) permet de montrer qu'on a affaire à un continuum entre des « formulations discursivement développées » et des « formulations réduites syntaxiquement à des réalisations spécifiques ». Pour les définitions par exemple, on pourra trouver des marqueurs tels que « je définis x comme », ou des procédés consistant simplement à souligner ou encadrer ce qui sera de fait perçu par le lecteur comme une définition, sans que cette dernière soit explicitement nommée comme telle.
Ces propriétés de l'énumération sont donc comprises comme les contreparties de l'acte mental consistant à inventorier des entités qui peuvent être énumérées. L'équivalence entre les formulations plus lexicales ou plus visuelles est postulée en termes de fonctionnalité dans les textes mais demande à être évaluée en regard de l'équivalence cognitive. Dans cette perspective, les approches expérimentales testent les effets de la mise en forme matérielle de l'architecture textuelle sur les processus de compréhension et de mémorisation, à travers différentes tâches (recherche d'informations, rappel etc.) (Virbel et al., 2005, Garcia-Debanc et Grandaty, 2001 ; Etcheverry et al. 2012).

Entre les marqueurs locaux et la structure globale de l'énumération, les composantes des structures énumératives sont alors définies selon le schéma suivant (Ho Dac et al., 2010) :
- l'énumération (une chaîne d'au moins deux items),
- est précédée dans la plupart des cas d'une « amorce » (segment qui annonce une énumération ; il peut inclure un énumérathème, ou élément prospectif spécifiant le critère de co-énumérabilité des items de l'énumération),
- peut contenir une « clôture » (segment qui conclut l'énumération ; il peut comporter un énumérathème, ou élément qui condense le contenu des items de l'énumération).
La structure énumérative conjugue des marqueurs lexico-syntaxiques (par ex. « premièrement », « les trois cas suivants », etc.), ponctuationnels (par ex. les virgules pour séparer les items, les parenthèses pour les hiérarchiser), typo-dispositionnels (par ex. les puces et numéros, la disposition horizontale ou verticale, la mise en gras[7]).
Une première distinction peut être faite en effet entre les structures énumératives horizontales, sans marquage visuel (cependant les séparateurs ponctuationnels jouent un rôle de l'ordre de la visibilité/lisibilité), qui se terminent fréquemment par « etc. » ou équivalent et les structures énumératives verticales, très présentes dans les articles scientifiques et dans les textes procéduraux comme le souligne C. Beaudet (2002), qui exploitent la dimension dispositionnelle sur laquelle prennent appui les lecteurs (études menées sur les mouvements oculaires) et qui facilite la mémorisation (tâches de rappel) (Virbel et al., 2005).

En ce qui concerne l'énumération d'items, deux catégories principales ont été isolées (Luc et al. 2000 ; Luc, 2001) :
**- Cas 1.** Les énumérations dans lesquelles au moins un **item est plus petit qu'une phrase** textuelle. Dans ce cas, les items pourraient correspondre à une proposition. Dans de telles énumérations, toutes les combinaisons peuvent être rencontrées. Par exemple l'amorce et le premier item forment une proposition et les autres items sont des propositions ou des phrases isolées. Nous illustrons ce premier cas avec deux exemples.
Exemple 1[8] :

> Les transports concernés sont : les transports vers un lieu distant de plus de 150 km, les transports en série (le nombre de transports prescrit pour un même traitement est au moins égal à quatre au cours d'une période de deux mois et chaque transport est effectué vers un lieu distant de plus de 50 km), les transports en avion ou bateau de ligne *(art. R. 322-10-4 du Code de la sécurité sociale).*

Dans cet exemple, l'énumération est formée par une seule phrase et les marqueurs sont essentiellement typographiques avec l'emploi de la ponctuation. Les deux points servent à marquer la fin de l'amorce et les virgules sont utilisées pour séparer les trois items. On peut noter également la régularité de la reprise des entités énumérées (« les transports ») en début de chaque item, qui sont par ailleurs annoncées dans l'amorce (« les transports concernés sont »). Cependant il est fréquent de rencontrer des emplois irréguliers.

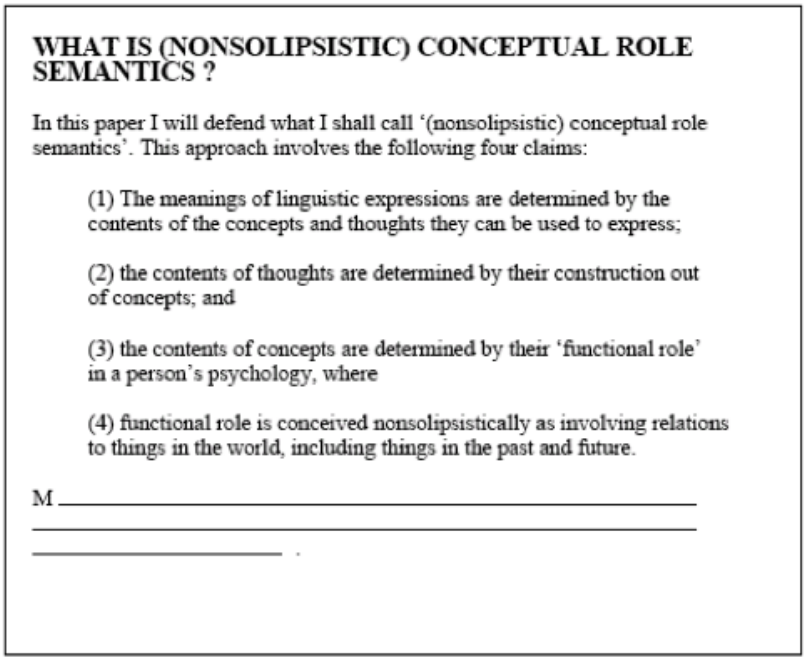
L'exemple 2 illustre l'utilisation du gras comme procédé typographique pour marquer les items de l'énumération :

[7] Comme on le voit dans notre propre énumération « cas 1 »/ « cas2 ».
[8] Les exemples 1 et 2 sont tirés du document « Demande d'accord préalable/Prescription médicale de transport – Notice », N° 50743#03.

Si votre patient peut se déplacer seul, sans assistance particulière, il convient de lui prescrire l'utilisation d'un **moyen de transport individuel** (son véhicule personnel, par exemple) ou les **transports en commun**.

**- Cas 2.** Les énumérations dans lesquelles les **items sont plus grands qu'une phrase textuelle**. Ces phénomènes proviennent dans ce cas de l'interaction entre la structure énumérative et les autres structures textuelles. C'est l'exemple d'un item lui même constituant une énumération imbriquée (horizontale, voire verticale). En effet, pour représenter ces cas de figures, un modèle non hiérarchique est nécessaire. L'exemple 3, qu'on détaillera ci-dessous illustre ce phénomène. Le premier item (1) de l'énumération principale possède en effet lui même une structure énumérative « horizontale » présentant les items (2) et (3).

Exemple 3[9] :



**WHAT IS (NONSOLIPSISTIC) CONCEPTUAL ROLE SEMANTICS ?**

In this paper I will defend what I shall call '(nonsolipsistic) conceptual role semantics'. This approach involves the following four claims:

(1) The meanings of linguistic expressions are determined by the contents of the concepts and thoughts they can be used to express;

(2) the contents of thoughts are determined by their construction out of concepts; and

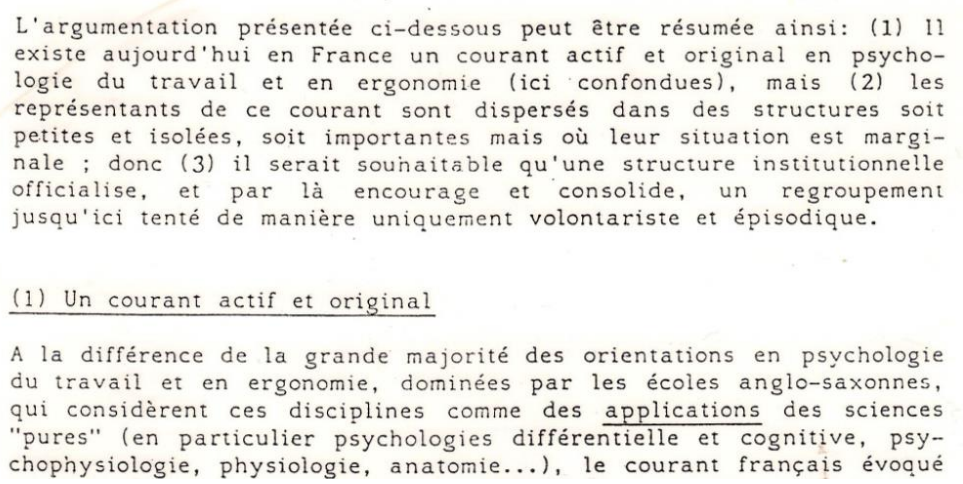
(3) the contents of concepts are determined by their 'functional role' in a person's psychology, where

(4) functional role is conceived nonsolipsistically as involving relations to things in the world, including things in the past and future.

M

L'exemple 4 est très fréquent. Il représente les cas où l'on commence un segment de discours en introduisant ce que l'on va présenter sous forme d'une première énumération marquée avec des indices typographiques et/ou dispositionnels, et où les items (ou leurs reformulations) sont repris un à un dans une deuxième énumération.
Exemple 4 :

[9] Les exemples 3 et 4 sont tirés du corpus (Virbel, 1999).



L'argumentation présentée ci-dessous peut être résumée ainsi: (1) Il existe aujourd'hui en France un courant actif et original en psychologie du travail et en ergonomie (ici confondues), mais (2) les représentants de ce courant sont dispersés dans des structures soit petites et isolées, soit importantes mais où leur situation est marginale ; donc (3) il serait souhaitable qu'une structure institutionnelle officialise, et par là encourage et consolide, un regroupement jusqu'ici tenté de manière uniquement volontariste et épisodique.

(1) Un courant actif et original

A la différence de la grande majorité des orientations en psychologie du travail et en ergonomie, dominées par les écoles anglo-saxonnes, qui considèrent ces disciplines comme des applications des sciences "pures" (en particulier psychologies différentielle et cognitive, psychophysiologie, physiologie, anatomie...), le courant français évoqué

Ce sont ces observations qui ont suscité la recherche de modèles ou de théories susceptibles de décrire ces phénomènes langagiers manifestés par des relations de nature visuelle, syntaxique et sémantiques entre les items et entre l'amorce et le premier item. Les deux modèles qui ont été retenus et combinés sont le Modèle d'Architecture Textuelle (MAT) (Pascual, 1990) et celui de la Théorie des Structures Rhétoriques (RST) (Mann & Thomson, 1998), comme exposé dans (Luc et al. 1999, 2000) et (Luc, 2001).
Nous illustrons ci-dessous quelques propriétés des énumérations à travers un exemple tiré du web, l'exemple 3 donné ci-dessus.
A première vue, cet exemple apparaît régulier : il commence par une amorce signalée par des marques lexicales et typographiques (« following four claims », « : »). De plus, les items sont présentés avec des procédés de mise en forme matérielle identiques : numérotation, diacritiques, espaces sur les dimensions verticales et horizontales. Cependant, si on regarde la structure syntaxique des items, nous remarquons qu'elles ne sont pas équivalentes. En effet, le dernier item est une proposition subordonnée du troisième item (« where »). En outre, le premier item est une sorte d'introduction aux deux items suivants : il introduit les notions de « contents of thought » et de « contents of concepts » qui distinguent le second et troisième item. Contrairement à une énumération classique, les items de cette énumération ne peuvent pas être interchangés.
Dans le but de décrire ces différences dans les structures internes des énumérations, nous faisons appel aux notions classiques en linguistique de paradigmatique vs syntagmatique. À travers ces relations, nous sommes capables de spécifier le parallélisme et le non-parallélisme entre ces items sans connotation sémantique. Cela peut-être comparé à la distinction de Noyau/Satellite dans la RST. Dans notre exemple, nous avons une relation syntagmatique entre l'item 3 et 4 (« where »), et une relation paradigmatique entre l'item 2 et 3 («and » coordonne les deux propositions introduites par un GN « the content of [thought/concepts] »).

Une modélisation des énumérations ne peut ignorer la diversité des cas de figure et le fait qu'on n'ait pas affaire, dans les textes, à des règles de bonne formation parfaitement immuables. Si les cas de non-parallélisme, ou d'hétérogénéité des modes de relations de dépendance entre items ne sont pas très fréquents dans les textes, ils n'en représentent pas moins des procédés attestés, qui jouent sur le sens.
Ainsi l'analyse de corpus de structures énumératives permet de formuler les constats suivants (Virbel, 1999 ; Luc, 2001) :

- Les items d'une énumération peuvent appartenir à différentes classes de constituants textuels ou syntaxiques.
- Des factorisations irrégulières (par exemple de propositions) et des coordinations sont fréquentes.

- Les organisateurs dans une même énumération sont hétérogènes.
- La borne finale d'une énumération n'est souvent pas claire et le marquage ponctuationnel est celui qui fonctionne le mieux.
- Des dépendances complexes peuvent exister entre des items d'une même énumération (ou cas d'énumération imbriquée ou adjacente voire distante).

### 2.3. La description des « Images de Page » en linguistique appliquée

Les travaux sur lesquels nous nous sommes basés ci-dessus ont été conduits dans la perspective notamment d'un repérage automatique des structures énumératives pour des enjeux de l'ordre de l'extraction d'information. Les éléments de description linguistique qui en résultent, ou disons sémiographique, compte tenu de l'importance du visuel, peuvent être mis à profit dans différents domaines d'application qui montrent l'intérêt d'une attention portée à la matérialité textuelle dans le cadre des dites sociétés de la connaissance.
C'est précisément dans cette perspective que nous nous situons avec l'idée d'une formation de scripteurs et de lecteurs attentifs à la matérialité des textes. Nous évoquerons cependant ici un autre exemple d'application qui concerne un cas bien particulier de lecteurs : les non-voyants. Il ne s'agit pas d'une simple digression dans notre propos : la littéracie, a fortiori à l'ère du numérique, n'est plus le fait seul d'une minorité constituée en élite, mais invite à n'exclure aucun profil de scripteur ou de lecteur et à envisager chacun en fonction de ses ressources propres et des contraintes générales pesant sur les activités de production et d'interprétation des textes.

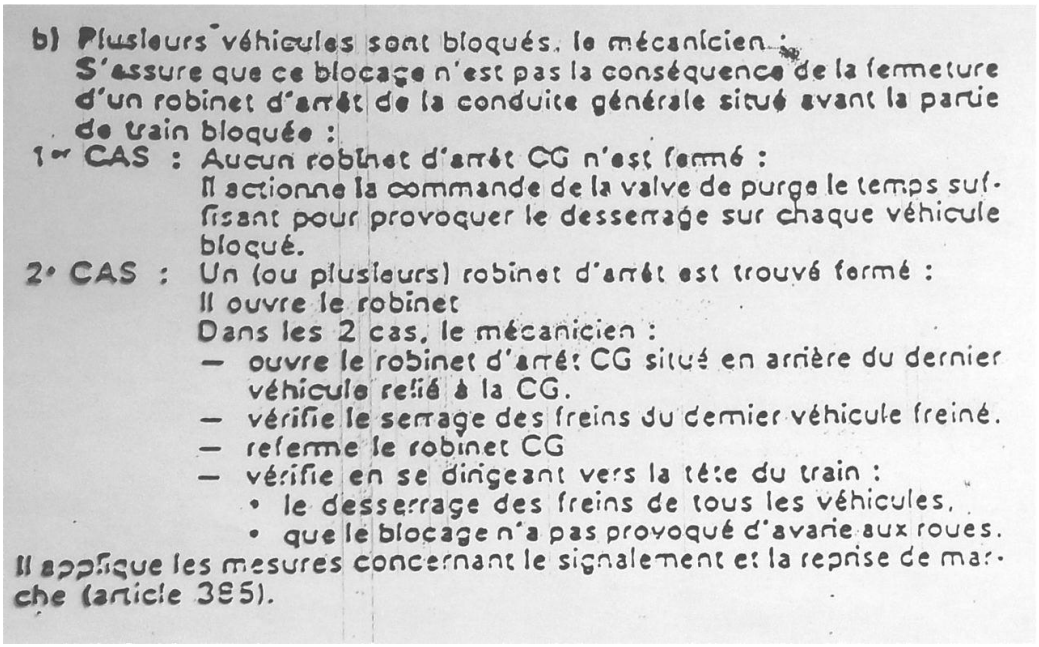
Un des grands problèmes rencontrés par les non-voyants est celui de la non accessibilité à la structure globale d'un contenu d'une page d'un document (« first glance » ou « first touch »). En effet l'outil utilisé par un non voyant (un lecteur d'écran) permet d'oraliser le contenu textuel (verbal) de manière séquentielle. De fait, de nombreuses indications visuelles (typographique et spatiale) utilisées pour la structuration ne sont donc pas accessibles.
Dans (Mojahid 2011 ; Mojahid et al. 2011) un modèle appelé « Images de Pages » (IdP) a été élaboré permettant une description sur plusieurs niveaux de granularité et autorisant la représentation du contenu des pages web. L'étude ayant porté sur des sites web de type « office de tourisme », les exemples d'objets concernés étaient 1) les regroupements visuels (par exemple les menus et les rubriques), 2) les relations visuelles (fournies par des propriétés typographiques de ressemblance/différence comme la couleur du texte et du fond, le style de police…) et 3) les relations contextuelles (comme par exemple « pour en savoir plus » et « voir aussi »). Le cas des objets appelés « regroupement visuels » concerne directement les structures énumératives avec, et cela est lié à la spécificité du corpus, une composante visuelle plus importante que celle représentée par les indices discursifs. Le menu comporte bien la structure d'une liste d'items qui peuvent eux-mêmes contenir des sous items. Un langage notationnel d'images de pages permet de hiérarchiser dans plusieurs niveaux d'images les marques de mise en forme matérielle, typo-dispositionnelles, lexico-syntaxiques et rhétoriques. En superposant les niveaux (comme des transparents), on cumule et on superpose les marques pour arriver à une image finale du texte. On peut ainsi jouer sur les indices à mettre de côté ou à faire apparaître dans la structure globale pour l'exhiber dans une perspective d'accessibilité des textes aux non-voyants.
Une fois les images de pages modélisées et imprimées, elles sont passées dans un four thermogonflable pour obtenir une reproduction tactile de la page. Ces représentations tactiles ont été testées auprès de personnes non-voyantes et ont montré leur intérêt dans les tâches de recherche d'information, de lecture en diagonale et de mémorisation (Bou Issa et al., 2010). Une représentation sous forme d'images de pages équivalentes aux énumérations dans, par exemple, des textes à consigne (recettes de cuisines, prescriptions médicales, itinéraires…) est ainsi susceptible d'apporter une aide importante et d'améliorer l'accessibilité des non voyants dans l'appréhension de la structure globale et dans les différents modes de lecture.

Sans aller jusqu'à la dimension tactile des textes, la notion d' « images de pages » a sa pertinence pour un public ne présentant pas de pathologie visuelle. On a vu en effet que la dimension visuelle entre dans la construction du sens et fait partie intégrante de l'activité de lecture. Un dispositif de superposition des images (par types de marques) entrant dans l'image globale du texte peut permettre de faire travailler un apprenant sur les marques, leurs combinaisons et leurs effets dans les textes. A ce titre, un logiciel comme Navilire représente un support intéressant car il se base sur ce même principe de visualisation par types de marques et de visualisation finale intégrant l'ensemble des marques pour fournir une image globale du texte. On peut envisager alors d'un point de vue expérimental de tester l'effet de combinaison des différentes marques, à différents niveaux de structuration, sur les différents modes de lecture et de navigation de l'apprenant.
En termes didactiques, l'objectif est de rendre les lecteurs conscients des opérations mises en œuvre lors de l'activité de lecture et des caractéristiques des textes sur lesquelles s'appuie leur interprétation, avec l'hypothèse que cela peut les aider, en tant que scripteurs, à identifier les ressources linguistiques à leur disposition et à mieux prendre en charge ce qui est traité en termes de cohérence des textes ou de guidage du lecteur. L'enjeu général de notre démarche est donc de développer des descriptions linguistiques suffisamment fines pour rendre compte de la textualité, qui soient en prise avec l'activité des sujets dans la production et la compréhension de textes et qui puissent être mises à profit dans une perspective didactique.

**3. Un scénario didactique à possibilités multiples**

Du fait de leur omniprésence, les structures énumératives peuvent concerner différents publics en difficulté avec l'écrit. Le cas des non-voyants évoqué plus haut en est un exemple. Le champ des écrits de travail ou écrits professionnels est directement intéressé par ces structures caractéristiques des textes procéduraux. L'approche ergonomique, en termes d'efficacité communicationnelle par exemple, peut être relayée par une approche esthétique, avec les jeux littéraires liés à la liste, comme chez Perec qui en est un illustre représentant. Il va de soi que la scénarisation didactique dépend étroitement du public ciblé et des compétences visées.
Sur la base de la description donnée supra des structures énumératives, nous proposons ici à grands traits un scénario à destination d'étudiants et d'adultes. L'objectif du scénario ne se situe pas au niveau seulement des caractéristiques d'une structure énumérative ; l'enjeu est de susciter, à partir de la matérialité textuelle, un questionnement sur l'écrit dans ses dimensions visuelle et verbale, à travers la diversité des écrits ordinaires et la variété des genres concernés, et en prise avec les activités de production et de compréhension des textes.

**3. 1. Point de départ : une situation-problème**
Nous prenons comme point de départ une situation-problème basée sur le document suivant, intégré au *Rapport de la Commission d'étude de l'accident de la Gare de Lyon* (Ministère des transports et de la mer, Paris, septembre 1988, 34 pages). Ce document a été étudié par J. André (1989) qui montre qu'une analyse de la structure énumérative par des éditeurs structurés comme LaTeX et SGML aurait aidé à repérer l'ambigüité du discours.



b) Plusieurs véhicules sont bloqués, le mécanicien :
S'assure que ce blocage n'est pas la conséquence de la fermeture d'un robinet d'arrêt de la conduite générale situé avant la partie de train bloquée :
1er CAS : Aucun robinet d'arrêt CG n'est fermé :
Il actionne la commande de la valve de purge le temps suffisant pour provoquer le desserrage sur chaque véhicule bloqué.
2e CAS : Un (ou plusieurs) robinet d'arrêt est trouvé fermé :
Il ouvre le robinet
Dans les 2 cas, le mécanicien :
— ouvre le robinet d'arrêt CG situé en arrière du dernier véhicule relié à la CG.
— vérifie le serrage des freins du dernier véhicule freiné.
— referme le robinet CG
— vérifie en se dirigeant vers la tête du train :
• le desserrage des freins de tous les véhicules,
• que le blocage n'a pas provoqué d'avarie aux roues.
Il applique les mesures concernant le signalement et la reprise de marche (article 385).

Comme l'indique le rapport de commission d'où est tiré ce document :

> « la commission a remarqué que, face à une situation concrète délicate, la liaison que doivent faire les agents entre les articles de règlements divers, comportant eux-mêmes des dispositions complémentaires et des renvois en bas de page, les prescriptions des consignes d'application et les procédures précisées par les guides spécialisés — guide du dépannage par exemple — n'est pas toujours aisée... La commission a relevé également — dans le cas de l'espèce — une présentation typographique défectueuse du PCM qui peut conduire à une interprétation erronée ».

Le passage qui pose problème fait état d'un écart en regard de bonnes pratiques ou de règles de bonne formation – dont on sait qu'elles résistent peu à la réalité des pratiques langagières : simple question d'écart par rapport aux marges, la mention « Dans les deux cas » se trouve alignée aux précisions réunies dans la rubrique « 2è cas » alors que d'un point de vue logique, la hiérarchie des constituants fait qu'elle subsume le 1er et le 2nd cas de figure auxquels le mécanicien peut avoir à faire face. Or, si le mécanicien estime avoir affaire au 1er cas, il n'a aucune raison de lire ce qu'il y a dans la rubrique « 2ème cas », du moins pas au-delà de la première ligne, qui se présente en symétrique avec celle du premier cas (« Aucun robinet d'arrêt CG n'est fermé »/« Un (ou plusieurs) robinet d'arrêt est trouvé fermé »).

La commission d'enquête signale dans son rapport final que c'est un des facteurs, mais pas le seul, de l'accident ; ainsi, ce problème d'instructions procédurales nous semble exemplaire pour montrer, dans une perspective didactique, l'importance de la matérialité de l'écrit et du langage envisagé dans une dimension praxéologique et le fait que les « bonnes pratiques » de l'écrit ne se limitent pas à une bonne orthographe et ne représentent pas uniquement une norme arbitraire et désuète pour académiciens et enseignants.

**3.2. Cheminement dans le corpus**

La seconde étape du scénario se base sur un corpus qui fait l'objet d'observations. Ces dernières sont guidées par des catégories d'analyse issues des travaux de description des structures énumératives. Nous proposons une analyse progressive de ces structures à l'aide de critères qui permettent de commencer par une approche générale d'un corpus ouvert et l'observation de la diversité des structures pour aller vers le détail de leurs composantes avec un corpus plus restreint. Le modèle des Images de Pages nous servira de support, dans la perspective d'une implémentation du scénario dans un logiciel de navigation comme Navilire.

3.2.1. Diversité des structures énumératives

Objectifs : repérer la présence de structures énumératives sur des supports et dans des genres variés. Quels éléments peuvent être compris comme des structures énumératives ? Cette étape sert de brainstorming sur les critères qui vont pouvoir servir de critères d'analyse. Elle porte sur un corpus ouvert (le corpus d'observation peut être complété par d'autres exemples). On peut utiliser des exemples attestant de la diversité des écrits se présentant sous la forme de listes (liste de courses, playlist, annuaires, inventaires, archives...) et intégrer éventuellement des exemples tirés d'autres systèmes graphiques (ex. écritures sémitiques) pour montrer que le lecteur même sans connaître la langue peut repérer des indices.
Il s'agit ici de pointer la double composante verbale et visuelle des structures, de faire prendre conscience du continuum existant entre des formulations discursivement développées et des formulations réduites, en prise avec la description des structures énumératives exposée dans la section 2.2. On pourra également faire émerger les premières notions comme celle de suite et d'ordre (ou non) des items, en confrontant des genres différents (par ex. une liste de Perec/ une liste dans un écrit procédural) pour faire ressortir la productivité de ce type de structures.

3.2.2. Notion de structure
Objectifs : analyser les structures énumératives en références aux notions (1) d'énumération horizontale et verticale, (2) d'amorce, liste et clôture et (3) d'énumérathème (notamment les déictiques textuels comme « ci-dessous », « suivant »). Cette étape porte sur le corpus de travail. Il ne s'agit pas de le décrire systématiquement mais, sur la base de la diversité des structures, de rendre compte de composantes macro de leur facture.

3.2.3. Ce qui fait liste
Objectifs : identifier la chaîne d'items et les marques lexico-syntaxiques, ponctuationnelles et typo-dispositionnelles qui servent à les séparer, à les aligner et éventuellement à les hiérarchiser. A cette étape, le corpus doit être restreint à quelques exemples pour en recenser l'ensemble des caractéristiques et montrer en quoi elles fonctionnent ensemble. Par exemple, pour une liste en trois points sur un diaporama : puces, tabulations, casse, syntaxe (items sous la forme d'un groupe nominal, d'un groupe verbal à l'infinitif...). D'autres structures tirées ou non du corpus peuvent alors servir de points de comparaison (par ex., comparer la syntaxe des items d'une diapositive à la pratique des titres d'un ouvrage par l'observation de sommaires, ou du menu d'un site web). Par ailleurs, on peut être amené en fonction du public à mettre l'accent sur les effets de non-parallélisme par exemple, ou sur la diversité des relations de dépendance entre items.

3.2.4. Pour approfondir
En fonction du public auquel on a affaire, on peut travailler plus en détail soit sur des analyses soit sur la manipulation de structures énumératives. Nous présentons ci-dessous deux exemples pour deux publics d'apprentis linguistes de niveaux radicalement différents :
-un public d'étudiants en linguistique et de type formation de formateurs à l'écrit, pour les faire travailler sur la notion d'énumérathème en regard des notions de deixis textuelles, de balises méta-textuelles assurant le guidage du lecteur, d'anaphore résomptive (une anaphore qui résume le contenu d'un passage plus ou moins long, typiquement avec un déterminant démonstratif et un nom abstrait). On peut faire lister et analyser les énumérathèmes d'amorce et de clôture dans le corpus ou une partie, ou dans un nouveau corpus, par exemple un corpus relevant d'un genre de texte spécifique.
-un public en difficulté avec l'écrit, pour les faire travailler sur la dimension dispositionnelle de l'écrit et les instructions d'ordre méta- qu'elle apporte pour la lecture. Sans originalité aucune, on pourra ici demander aux apprenants de présenter de manière descriptive le plan de tel ou tel article ou dissertation, ou, à l'inverse, de mettre sous forme de structure énumérative un paragraphe rédigé. Un genre ou un support spécifiques peuvent permettre de s'adapter à la spécificité du public. Les caractéristiques des structures énumératives doivent permettre de guider la production (y compris la révision et la réécriture).

**3.3. Bilan. Comment lit-on une énumération ?**

En fin de séance, les objectifs sont de s'appuyer sur les acquis des étapes précédentes pour interroger la mise en relation des composantes d'une structure énumérative dans la manière dont elle est lue. Cette étape représente la phase d'institutionnalisation des savoirs : elle doit permettre de reprendre les notions visées par le scénario. Elle peut prendre des formes variées : leçon, exercice d'application et évaluation. La visualisation finale des textes annotés que propose Navilire comme aboutissement du scénario didactique (voir section 1.3.) représente un outil particulièrement intéressant. On peut imaginer par ailleurs une activité de production textuelle basée sur un principe de mise en situation et demandant aux apprenants un travail de description du texte qu'ils ont produit, de manière à récapituler les notions abordées et à favoriser par hypothèse le transfert des compétences d'analyse textuelle entre lecture et écriture et entre divers genres et types de textes.

***

Le scénario à possibilités multiples proposé ici reste en l'état largement imprécis et peut être adapté à des publics divers et à des corpus variés : il lui manque donc des objectifs didactiques spécifiques, au-delà de l'objectif général d'une prise de conscience des formes et du fonctionnement des structures énumératives comme entrée intéressante pour travailler sur la conscience de l'écrit, la matérialité des textes et les modes de construction du sens.

On pourrait envisager un travail sur les listes dans un atelier d'écriture créative aussi bien que dans une formation « pro » au diaporama. Les propositions faites ici posent les jalons d'une démarche qui consiste à mener des analyses linguistiques (« description ») à partir d'un ensemble diversifié de textes ou d'extraits de textes (« corpus »). Restent à définir en fonction du public et des objectifs spécifiques les critères de constitution du corpus (par ex., en fonction du public, quelle(s) langue(s), quels genres, quels supports, quel degré de facilité de lecture...) et les critères d'analyse (en extension et en intension) aux fondements de la description qui en sera faite, in fine, en contexte d'enseignement ou de formation.

Quant à l'innovation pédagogique, comme enseigner et apprendre avec les corpus n'a au fond rien de nouveau, c'est sans doute sous l'angle des possibilités offertes par les technologies du numérique que l'on peut inventer de nouvelles modalités de travail.

**4. Navigation dans les textes avec Navilire et évolutions possibles**

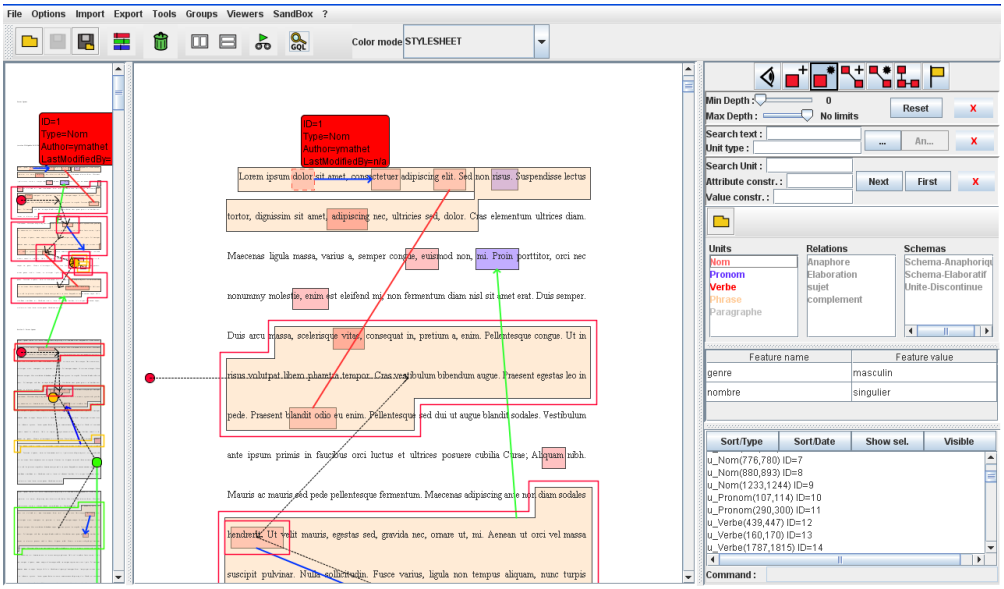
Dans cette partie, nous proposons d'intégrer fictivement le scénario exposé ci-dessus dans le logiciel Navilire pour voir quelles évolutions de Navilire pourrait appeler ce scénario. A travers le cas des structures énumératives, il s'agit d'envisager comment un logiciel à usage didactique basé comme Navilire sur des corpus textuels et leur description sous la forme d'annotations peut évoluer en regard des avancées de la linguistique de corpus et des contraintes de l'utilisation des technologies dans l'enseignement. N'étant pas nous-mêmes développeurs, nous ne formulons ici que quelques remarques qui puissent servir d'éléments de réflexion dans une perspective ingénierique.

La première série de remarques concerne la dimension visuelle de l'écrit. Le principe même du logiciel Navilire étant celui de la navigation dans les textes, il intègre cette composante en donnant accès sur une même fenêtre à un large extrait et à la possibilité de dérouler le texte dans son ensemble. Les jeux de couleurs liés au schéma d'annotation rendent visible la textualité (ainsi par ex. de telle chaîne de co-référence dans les mentions d'un personnage au fil d'un texte). Cependant, comme le montre l'exemple des structures énumératives, il pourrait être intéressant de pouvoir obtenir une vue authentique du texte, non pas uniquement comme une suite de caractères structurés en paragraphes, mais sur le support et avec la mise en forme sinon originales, du moins reproduits sous un format image. Dans les logiciels issus du champ de l'analyse des données

textuelles[10], cet accès au texte « original » a été pointé comme un élément important du travail d'interprétation dont ces outils sont des artefacts.
Par ailleurs, l'accès à une image du texte peut s'accompagner aussi d'un travail de description de la composante visuelle du texte. Il s'agirait alors de coupler les méthodes d'annotations des textes avec celles de l'image. On peut imaginer, a minima, un système de pointeur sur des objets ou espaces textuels (par ex. pointer des puces, pointer l'espace des notes de bas de page). Plus simplement, le recours à l'image ou à la vidéo peut être pertinent pour donner à voir telles ou telles composantes de l'énumération, par exemple dans la phase d'institutionnalisation des savoirs à la fin du scénario. Images et vidéo peuvent aussi servir comme point de départ, pour permettre aux apprenants d'explorer la diversité des textes et extraits de textes du corpus. A cette étape exploratoire cependant, il peut être intéressant de pouvoir avoir simultanément quelques textes sous les yeux ; on peut imaginer une vue multifenêtres avec des possibilités de navigation de l'une à l'autre.

La seconde série de remarques concerne la question de la description du corpus et les annotations qui en résultent. Le développement du domaine des annotations textuelles et discursives est prometteur dans la perspective de l'utilisation didactique de la navigation textuelle. Nous évoquerons à ce titre la plateforme d'annotation discursive Glozz[11] (Widlöcher et Mathet, 2009, 2012) conçue au départ dans le cadre du projet ANNODIS[12] (Péry-Woodley et al., 2009). Ce projet avait pour objectif le développement d'un corpus de référence annoté discursivement et d'outils d'annotation et d'exploitation de corpus annotés.
La plateforme Glozz se base peu ou prou sur les mêmes principes que le logiciel Navilire que nous avons présenté en section 1.3. : elle permet à partir d'un texte ou d'un ensemble de textes de procéder à des annotations sur ses composantes. Sa visée est de l'ordre de recherches en linguistiques, mais un tel dispositif pourrait se prêter semble-t-il à une utilisation dans une perspective didactique.
Les logiciels ou plateforme d'annotation présentent une interface dont l'ergonomie est relativement intuitive : à une fenêtre de texte est couplée une fenêtre dédiée à l'outil d'annotation. Comme on le voit sur la figure ci-dessous, dans Glozz, on a deux fenêtres pour le texte (à gauche, une vue globale d'affichage de plusieurs pages et au centre, la fenêtre sur laquelle vont se porter les annotations). Si la granularité des annotations peut être telle qu'on aboutit à une représentation textuelle relativement complexe (comme dans la vue donnée ci-dessous), on peut envisager des modèles d'annotations plus simples, adaptés aux objectifs didactiques.

[10] Voir par ex. le logiciel TXM et le site Textométrie http://textometrie.ens-lyon.fr/
[11] Voir le site http://www.glozz.org/
[12] Voir le site http://w3.erss.univ-tlse2.fr/textes/pagespersos/annodis/ANNODISfr.html



Vue de la plateforme Glozz

Par rapport aux possibilités actuelles de Navilire, l'intérêt de la plateforme Glozz se situe à plusieurs niveaux :
-La possibilité d'annoter non seulement des unités mais également des relations entre unités et des structures de plus haut niveau combinant unités, relations entre unités, voire relations entre relations.
-La possibilité de superpositions des annotations et d'annotations multiples (en fonction des phénomènes annotés/ en fonction de différents annotateurs).
-La possibilité de comparer les annotations faites par un apprenant avec les annotations expertes comme cela a été fait dans une perspective d'amélioration du modèle d'annotation et du guide donné aux annotateurs (Adam et Vergez-Couret, 2012) et qui pourrait servir à l'évaluation ou l'auto-évaluation de l'apprenant.
-Des annotations rendues visibles ou non aux apprenants en fonction du scénario (un même texte ou corpus annoté peut ainsi donner lieu à plusieurs « images de pages » ainsi qu'à plusieurs type de parcours selon le public).
-Une fonction de tracking de l'activité de l'annotateur (temps passé, segments annotés, annotations utilisées) qui fait de Glozz un bon support pour des expérimentations ; nous pensons à des études expérimentales sur l'écrit d'une part et en didactique de l'écrit d'autre part, pour analyser l'activité du lecteur via les annotations qu'il porte sur le texte, et en particulier l'activité de l'apprenant face aux consignes et aux textes qui lui sont soumis.
La plateforme Glozz ou les possibilités qu'elle offre pourraient ainsi ouvrir de nouvelles perspectives pour l'usage didactique de l'annotation de corpus, par les enseignants pour l'élaboration d'exercices et par les apprenants pour leur réalisation, selon les principes de la navigation textuelle pour la didactique des langues initiés avec le logiciel Navilire.

Notre dernière remarque concernera l'utilisation des technologies en contexte d'enseignement et d'apprentissage. Outre les hypothèses cognitives qui justifient l'intérêt d'un travail d'observation de corpus et de navigation textuelle en didactique des langues, le recours à ce type de ressources et outils est jugé pédagogiquement intéressant à la fois en présentiel et comme complément de cours en présentiel ou comme mode d'auto-formation, par exemple dans des laboratoires de langue. Il pourrait être intéressant de favoriser un accès on-line à ce type de ressources, pour éviter les contraintes liées à l'installation d'un logiciel.
Plus largement, on voit l'intérêt de la démarche de navigation textuelle en didactique de l'écrit et l'importance 1) d'envisager les questions de structuration textuelle en y intégrant la dimension visuelle des textes avec la notion d' « image de page », 2) de coupler les activités de l'ordre de la

navigation textuelle avec d'autres activités en corpus comme celles qui sont le plus souvent utilisées jusqu'ici, basées sur les concordanciers : par exemple, un scénario didactique sur les anaphores démonstratives « ce + N » dans le discours scientifique peut à la fois mobiliser un travail sur quelques textes et un travail sur les concordances extraites d'un plus large corpus pour faire prendre conscience du type de noms concernés (en l'occurrence, souvent, un lexique transdisciplinaire du type « cette évidence », « cette hypothèse », etc.).
Le développement d'une plateforme didactique pourrait ainsi être envisagé, qui permettrait aux apprenants de mener des activités d'observation, de réaliser des exercices d'analyse des textes et de mobiliser des outils divers de la formation à l'écrit ou de l'aide à la rédaction (par exemple, l'utilisation de dictionnaires, de corpus multilingues alignés, etc.). Les outils didactiques prendraient ainsi place dans la formation du scripteur-expert à côté des outils d'analyse automatique qui guident le scripteur au cours de son activité de production, comme le correcteur orthographique qui souligne en rouge les erreurs et en vert ce qu'il identifie comme étant possiblement erroné, ou comme les propositions faites dans le sens de la rédaction automatique (de résumés par ex.).
Enfin, le développement d'annotations de corpus et d'exercices pour des usages de l'ordre de la navigation didactique dans les textes pose la question de la mutualisation des ressources. En particulier, on se demandera, à l'instar du travail de thèse de M. Loiseau (2009), avec quels critères il faut identifier les ressources pour les indexer, de manière à ce qu'un formateur (voire un apprenant en auto-formation ?) puisse s'y retrouver, comme lorsqu'il cherche dans un ouvrage ou dans ses propres classeurs des supports pédagogiques et des séances ou exercices-types.

**Conclusion**

Le cas des structures énumératives, qui représentent des structures omniprésentes et fondamentales dans l'ordre scriptural nous a permis de rendre compte d'une démarche qui met l'accent sur l'importance de la description linguistique pour le développement de ressources en didactique de l'écrit. La description linguistique sert en effet de base à l'élaboration d'un scénario visant la prise de conscience par les apprenants de caractéristiques textuelles (ici celles des énumérations), qui s'instancient différemment selon les textes, les types de séquences textuelles et les genres et qui sont un appui pour le lecteur dans son activité de compréhension comme pour le scripteur dans son activité de production.
Le principe de visualisation des marques entrant dans la structure textuelle est au fondement de la notion d' « images de page » développée notamment dans une perspective d'accessibilité des textes aux non-voyants. Ce principe est également au coeur du logiciel Navilire qui peut permettre d'annoter les textes à plusieurs niveaux, selon des types de marques, et de superposer ces différents niveaux pour fournir une image globale du texte. Le scénario didactique doit permettre d'accompagner l'apprenant dans ses observations et de faire porter son attention, étape par étape, sur les différents types de marques et sur leurs combinaisons à l'échelle du texte.
Il s'agit de faire porter l'attention des apprenants sur la matérialité de l'écrit et sur ses effets pragmatiques. La navigation textuelle, en tant que principe de va-et-vient entre la dimension locale et la dimension globale des textes se présente comme une approche heuristique pour favoriser le développement chez les apprenants de compétences réflexives de l'ordre de l'analyse linguistique. Cet objectif de former des « apprentis-linguistes » peut tout aussi bien concerner la formation de futurs linguistes, d'enseignants et formateurs, de publics en difficulté avec l'écrit ou de rédacteurs professionnels.
Les apports actuels dans le champ des annotations textuelles et discursives en linguistique de corpus permettent d'envisager les évolutions d'un logiciel comme Navilire et des ressources développées selon le principe de la navigation didactique dans les textes. Les corpus annotés sont également un outil pour la recherche en linguistique et un logiciel comme Navilire ou une plateforme d'annotation et de navigation comme Glozz pourraient alors servir de support pour des approches expérimentales évaluant d'un point de vue cognitif les effets de la mise en forme matérielle des textes sur la lecture.

**Références bibliographiques**


Adam J.-M. (2001). « Types de textes ou genres de discours ? Comment classer les textes qui disent de et comment faire ? » *Langages*, 141, 10-27.

Adam C. & Vergez-Couret M. (2012). « Exploiting naive vs expert discourse annotations: an experiment using lexical cohesion to predict Elaboration / Entity-Elaboration confusions », in *Proceedings of the Sixth Linguistic Annotation Workshop, Association for Computational Linguistics*, 22-30

André J. (1989). « LaTeX ou SGML pouvaient-il éviter la catastrophe de la gare de Lyon ? », Cahiers GUTenberg, 1, 21-25, IRISA, Rennes.

Anis J. (1988). *L'écriture. Théories et descriptions*. En collaboration avec J.-L. Chiss et J. Puech, Paris, Bruxelles, De Boeck Université.

Anis J. (1997) « Lire des visuo-textes », *Cahiers du Français contemporain* n° 4, « Écritures et textes d'aujourd'hui », M. Marquilló (éd.), Crédif, ENS Éditions, Fontenay-St Cloud, 29-45.

Anis J. (2000). « Vers une sémiolinguistique de l'écrit ». *Linx*, 43. En ligne : http://linx.revues.org/1046

Baccino T. & Pynte J. (1998). « Spatial Encoding and Referential Processing during Reading ». *European Psychologist*, 3(1), 51-61.

Beaudet C. (2000). « Clarté, lisibilité, intelligibilité des textes : un état de la question et une proposition pédagogique », *Recherches en rédaction professionnelle*, 1(1), 1-19.

Beaudet C. (2002). « Lisibilité textuelle et configuration des énumérations dans un texte procédural », *Revue canadienne de linguistique appliquée*, 5(1-2), 7-23.

Bou Issa Y., Mojahid M., Oriola B. & Vigouroux N. (2010). « Analysis and Evaluation of the Accessibility to Visual Information in Web Pages ». *International Conference on Computers Helping People with Special Needs*, *Vienna, Austria*, Vol. 6179, Springer, Lecture Notes in Computer Science.

Bouraoui J.-L. (2000). *Les structures énumératives : caractérisation linguistique et reconnaissance automatique.* Rapport de DEA, Université Toulouse le Mirail.

Charolles M. (1981).« Coherence as a principle in the interpretation of discourse », *Text* 3, 71-99.

Charolles M. (1994). « Cohésion, cohérence et pertinence du discours », *Revue internationale de linguistique française*, 29, 125-151.

Couto J., Lundquist L. & Minel J.-L. (2005a). « Navigation interactive pour l'apprentissage en linguistique textuelle », *EIAH 2005*, Montpellier, En ligne sur http://hal.archives-ouvertes.fr/docs/00/03/18/27/PDF/3.pdf

Couto J., Lundquist L. & Minel J.L. (2005b). « Naviguer dans les textes pour apprendre », *TALN 2005*, Dourdan, 6-10 juin 200. En ligne sur http://halshs.archives-ouvertes.fr/docs/00/09/80/24/PDF/TALN_COUTO_LUNDQUIST_MINEL.PDF

Desoutter C. et Mellet C., eds (à paraître, 2013). *Le discours rapporté et ses marques, perspectives théoriques et didactiques*, *Actes des Journées d'étude organisées par le « NaviLireResearch Group » (NRG)*, Université de Bergame, Italie, 5 et 6 mai 2011.

Dewey, J. (1916/ 1966). *Democracy and education*. New York, Free Press.

Etcheverry I., Baccino T., Terrier P., Marquié J.-C., Mojahid M. (2012). « *Age differences in information finding tasks: performance and visual exploration strategy with different page layouts ». Computers in Human Behavior*, 28(5), 1670-1680.

Fayol M. (1997). *Des idées au texte, psychologie cognitive de la production verbale, orale et écrite*. Paris, PUF.

Freire, P. (1970). *Pedagogy of the Oppressed*. New York, Continuum.

Garcia-Debanc C. (eds) (2001). *Les discours procéduraux*, *Langages*, 141, Paris, Larousse.

Garcia-Debanc C. & Grandaty M. (2001). « Incidence des variations de la mise en forme textuelle sur la compréhension et la mémorisation de textes procéduraux (règles de jeux) par des enfants de 8 à 12 ans », *Langages*, 141, 92-104.

Goody J. (1979) *La Raison graphique. La domestication de la pensée sauvage,* Paris, Éditions de Minuit. Edition originale 1977.

Goody J. (1986). *La logique de l'écriture*. Paris, Armand Colin.
Granger S., Hung J. and Petch-Tyson S. (eds.) (2002). *Computer Learner Corpora, Second Language Acquisition, and Foreign Language Teaching*. Amsterdam, John Benjamins.
Ho-Dac, L. M., Péry-Woodley, M.-P. & Tanguy, L. (2010). « Anatomie des structures énumératives ». *Actes TALN 2010*, Montréal, Québec, 19-23 juillet 2010.
Jeandillou J.-F. (2008). *Effets de textes,* Limoges, Editions Lambert-Lucas.
Jubert R. (2007). *Typography and graphic design : from Antiquity to the present*, Paris, Flammarion.
Labasse B. (1999). « La lisibilité rédactionnelle : fondements et perspectives », *Communication et Langages*, 121, 86–103.
Loiseau M. (2009). *Élaboration d'un modèle pour une base de textes indexée pédagogiquement pour l'enseignement des langues*. Thèse de doctorat, Université Stendhal Grenoble 3. http://tel.archives-ouvertes.fr/tel-00440460/fr/
Luc C. (2001). « Une typologie des énumérations basée sur les structures rhétoriques et architecturales du texte ». *TALN*, Université de Tours.
Luc, C., Mojahid, M., Virbel, J., Garcia-Debanc, C. & Péry-Woodley, M-P. (1999). « A linguistic approach to some parameters of layout: A study of enumerations ». In R. Power & D. Scott (Eds.), *Using Layout for the Generation, Understanding or Retrieval of Documents*, AAAI 1999 Fall Symposium, 20-29. (aussi *AAAI Fall Symposium Technical Report* FS-99-04, pp. 35-45).
Luc, C., Mojahid, M., Péry-Woodley, M.P. & Virbel, J. (2000). « Les énumérations : structures visuelles, syntaxiques et rhétoriques ». *In Actes CIDE'2000 (Colloque International sur le Document Électronique)*, Lyon.
Luc, C., Mojahid, M., Virbel, J. (2002). « Le modèle d'architecture de texte », in J. Virbel (ed.) *L'Inscription Spatiale du Langage : structures et processus*, Toulouse, IRIT-CNRS.
Lundquist L. (2008). *Navigating in Foreign Language Texts*. Frederiksberg, Sam- fundslitteratur
Lundquist L. (2013). *Lire un texte académique en français*. Paris, Ophrys.
Lundquist L., Minel J. & Couto J. (2006). NaviLire, Teaching French by Navigating in Texts. Information Processing and Management of Uncertainty in Knowledge-based Systems, En ligne sur http://halshs.archives-ouvertes.fr/halshs-00097849/en
Mann W.C. & Thompson, S. A. (1988). « Rhetorical structure theory: toward a functional theory of textorganization », *Text*, 8 (3), 243-281.
Mojahid M. (à par.). « *The Foundations and Testing of Page Images Language (IdP) (short paper) ».* Actes de la conférence Linguistic and Psycholinguistic Approaches to Text Structuring (LPTS 2011), Louvain, Belgique,
Mojahid M., Bou Issa Y., Oriola B. & Vigouroux N. (2011). « Un modèle d'architecture de pages web pour une accessibilité augmentée destinée aux non-voyants », *Colloque International sur le Document Électronique,* Maroc.
Nespoulous J-L. & Virbel J. (2007). « From the study of language dysfunction and handicap to a better understanding of linguistic processing in normality », in M.J. Ball & J.S. Damico (Eds.) *Clinica lAphasiology. Future directions,* Hove & New York, PsychologyPress, Taylor & Francis Group, 107-124.
O'Keeffe A., McCarthy M. and Carter R. (2007). *From Corpus to Classroom: Language Use and Language Teaching*. Cambridge, Cambridge University Press.
Ong W.J. (1988). *Orality and literacy: the technologizing of the word,* Routledge. 1ère édition 1982.
Pascual E. (1991). *Représentation de l'architecture textuelle et génération de textes*. Thèse de doctorat, Université Paul Sabatier, Toulouse.
Pascual E. & Péry-Woodley M-P. (1999). « Définition et action dans les textes procéduraux ». In J. Virbel, J-M. Cellier & J-L. Nespoulous (Eds.), *Cognition, Discours Procédural, Action*, Vol. II. Toulouse, PRESCOT, 80-108.
Péry-Woodley M.-P. (2001). « Mode d'organisation et de signalisation dans des textes procéduraux », *Langages*, 141, 28–46.

Péry-Woodley, M.-P., Asher, N., Enjalbert, P., Benamara, F., Bras, M., Fabre, C., Ferrari, S., Ho-Dac, L.-M., Le Draoulec, A., Mathet, Y., Muller, P., Prévot, L., Rebeyrolle, J., Tanguy, L., Vergez-Couret, M., Vieu, L. & Widlöcher, A. (2009). « ANNODIS: une approche outillée de l'annotation de structures discursives ». In *Actes de la 16ème Conférence TALN*, session posters, Senlis, France.

Peytard J. (1970). « Oral et scriptural : deux ordres de situations et de descriptions linguistiques », *Langue Française*, 6(6), 35-47.

Peytard J. (1975). « Lecture(s) d'une « aire scripturale » : la page de journal ». *Langue Française*, 28, 39-59.

Rinck F. & Lebarbé T. (2005). « Constitution et exploitation de ressources linguistiques : enseignement et apprentissage du discours rapporté ». *Actes des 4èmes Journées linguistique de corpus*, 15-17 septembre 2005, Lorient. <http://web.univ-ubs.fr/corpus/jlc4.html>

Rinck F. & Tutin A. (2007). « Annoter la polyphonie dans les textes : le cas des passages entre guillemets », *Corpus*, 6, 79-100.

Rinck F., Grossmann F. & Boch F. (à paraître). « Observer le rôle des guillemets dans un corpus : une voie d'accès à l'énonciation ? » *Actes du colloque Enonciation et texte au cœur de la grammaire*, Université Toulouse le Mirail, 11-13 mars 2009, Presses Universitaires de Rennes.

Schmid S. (2001). *Représentation organisationnelle et compréhension : rôle de la mise en forme matérielle dans la lecture.* Thèse de doctorat, Université de Nice Sophia Antipolis.

Sinclair J. (ed.) (2004). *How to Use Corpora in Language Teaching*. Amsterdam, John Benjamins.

Swales J. (1990). *Genre analysis. English in academic and research settings*, Cambridge, Cambridge University Press.

Tribble C. and Jones G. (1997). *Concordances in the Classroom: Using Corpora in Language Education*. Houston TX, Athelstan.

Tribble C. (2003). 'The text, the whole text…or why large published corpora aren't much use to language learners and teachers' in B. Lewandowska-Tomaszczyk (ed.) *Practical Applications in Language and Computers*, Frankfurt, Peter Lang, 303-318.

Turco G. & Coltier D. (1988). « Des agents doubles de l'organisation textuelle, les marqueurs d'intégration linéaire », *Pratiques*, 57, 57-79.

Virbel J., Garcia-Debanc C., Baccino T., Carrio L., Dominguez C., Jacquemin C. Luc Ch., Mojahid M., Pery-Woodley M.-P. & Schmid S. (2005). « Approches cognitives de la spatialisation du langage. De la modélisation de structures spatiolinguistiques des textes à l'expérimentation psycholinguistique : le cas d'un objet textuel, l'énumération ». In Thinus-Blanc C., Bullier J. (eds.), *Agir dans l'espace*, Paris, Editions de MSH, 233-254.

Virbel J. (1989). The contribution of linguistic knowledge to the interpretation of text structure. In André, J., Quint, V. et Furuta, R. (eds), *Structured Documents*, 161–181. Cambridge UniversityPress.

Virbel, J. (1999). *Structures textuelles, planches fascicule 1 : Enumérations, Version 1*, Technical report, IRIT.

Wichmann A., Fligelstone S., McEnery A. and Knowles G. (eds.) (1997). *Teaching and Language Corpora*. London, Longman.

Widlöcher, A. & Mathet, Y. (2009). « La plate-forme Glozz : environnement d'annotation et d'exploration de corpus ». In *Actes de la 16e Conférence Traitement Automatique des Langues Naturelles (TALN)*, session posters, Senlis, France.

Widlöcher A. & Mathet Y. (2012). "The Glozz Platform: A Corpus Annotation and Mining Tool". In *Proceedings of the 2012 ACM Symposium on Document Engineering*, 171-180.