

Chapitre 12
Approches cognitives
de la spatialisation du langage.

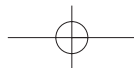
De la modélisation de structures spatiolinguistiques
des textes à l'expérimentation psycholinguistique :
le cas d'un objet textuel, l'énumération

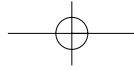
Jacques Virbel, Sabine Schmid, Laetitia Carrio, Corinne
Dominguez, Marie-Paule Pery-Woodley, Christian Jacquemin,
Mustapha Mojahid, Thierry Baccino, Claudine Garcia-Debanc

ABSTRACT

The page, whether on screen or on paper, constitutes a well circumscribed space constructed and regulated by a set of processing instructions (the so-called spatio-linguistic text structures) inscribed in the text in the form of lexical-syntactic phrases, typographic marks and dispositional devices. It is thus a specific, spatially artefactualized modality of language, and by the way, it is a particularly appropriate object for the study of spatial cognition. Our main objectives have been a) to throw light on the lexical-syntactic component of text and its visual realisation on the page (linguistic modelling), b) to construct computer-based tools for the extraction of information and the control of text images (computational formalisation), c) to study the impact of the visual structure of text on comprehension and recall in children and adults, and the specific role of visual properties on reading strategies (psycholinguistic experimentation). The corpus analysis has led to the construction of descriptive models for the text objects under study : enumerations (and some aspects of titles). A particular model of text structures has been referred to a common background for the three approaches : the Model of Text Architecture (rooted in the notion of speech acts and a view of text layout as a trace of textual metalanguage), linked with the Rhetorical Structure Theory (which do not take into account visual aspects). As a result, number of the aspects and the relations of the structure of enumerations have been evidenced, including: the syntactic nature of the introductory expression, the syntactic and rhetorical constraints between items and layout features signalling these items.

Virbel, J., Schmid, S., Carrio, L., Dominguez, C., Pery-Woodley, M. P., Jacquemin, C., et al. (2005). Approches cognitives de la spatialisation du langage: le cas de l'énumération. In C. Thinus-Blanc & J. Bullier (Eds.), *Agir dans l'espace* (pp. 233-254). Paris: Editions de la MSH.



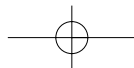


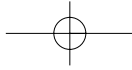
Les sciences du langage comme, dans une moindre mesure, celles de l'information sont restées longtemps «aveugles» au rôle des propriétés visuelles du langage inscrit, alors que d'autres recherches (anthropologie, histoire des textes, psychologie cognitive et expérimentale) avaient signalé l'importance fondamentale de ces aspects du point de vue cognitif (Ong, 1982 ; Goody, 1986 ; Olson, 1994 ; Hartley, 1980). Pour surmonter cette situation, nous pensons qu'il fallait que se réalise la conjonction des disciplines citées dans le cadre des recherches sur la cognition spatiale.

La face visuelle des textes (*i. e.* : le mode de présentation de l'information) comporte au moins quatre composantes principales : la contribution au sens du texte ; l'efficacité communicationnelle ; l'impact de valeurs culturelles, esthétiques ou expressives ; la prise en compte de contraintes (technique, comme les propriétés des supports, ou économique, comme la longueur physique imposée par un format de collection par exemple). Ces composantes sont profondément imbriquées, et chacune d'entre elles est multidimensionnelle : les caractères ne peuvent pas ne pas avoir des propriétés morphologiques qui n'ont pas de correspondance avec l'oral (corps, graisses, casses, inclinaisons, etc.), et par ailleurs, l'écrit implique des organisations spatiales globales et locales qui ont très peu à voir avec par exemple des phénomènes dits supra-segmentaux. Ces phénomènes n'interviennent pas en aval d'un message déjà constitué : de même qu'il n'est pas possible en génération de textes de déterminer d'abord le contenu d'un message, puis son expression linguistique (Danlos, 1985), il n'est pas possible non plus de déterminer d'abord le message (contenu + expression), puis sa présentation visuelle (Pascual, 1991). Cela s'explique par le fait que la composante signifiante des propriétés visuelles du texte est en correspondance avec des éléments discursifs précis à portée métalinguistique et qui appartiennent au message proprement dit, comme on le perçoit sur ce simple exemple : «L'introduction de ce premier chapitre est consacrée aux structures textuelles.» *vs* la réalisation suivante, qui comporte les traces de retraitements discursifs (syntaxiques, lexicaux, ponctuationnels), et morphologiques (typographiques et dispositionnels) :

CHAPITRE I

Introduction Les structures textuelles

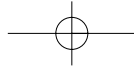




Les recherches sur la cognition spatiale permettent d'appréhender les mécanismes impliqués dans le traitement de l'information visuelle, ainsi que la catégorisation des marqueurs linguistiques de l'espace, les classes d'entités spatiales que la langue distingue, et les structures ainsi définies (relations partie-tout, dépendances fonctionnelles, etc.). Or la page (ou la double-page), qu'elle soit d'écran ou de papier, constitue un espace circonscrit, construit et régi par des conventions, d'une nature au demeurant très différente de celles d'une syntaxe. Ici, l'écriture se situe d'une manière originale : si de nombreux artefacts ont été conçus pour représenter, stocker, exploiter des données spatiales (cartes, plans, etc.), avec l'écriture, des objets spécifiques (tablettes d'argiles, rouleaux, feuilles, écrans, fenêtres) ont été directement conçus pour circonscrire et exercer la cognition spatiale à des fins de notation et d'interprétation du langage. Un tel déplacement n'a pu intervenir qu'au prix de, et grâce à la construction de concepts originaux caractérisant la maîtrise de l'activité d'écriture comme de lecture (la *literacy*) : page, double-page, marge, colonne, où la composante spatiale est immédiatement sensible ; tout comme des types d'objets spécifiquement textuels tels que titre, paragraphe, énumération, liste, note marginale ou infrapaginale, citations, etc. qui ne peuvent que reposer sur une interprétation spatiale des supports et/ou grâce à des propriétés morphologiques contrastives des caractères (gras, capitales, etc.). Avec la modalité du support électronique du texte et l'extension d'Internet, intervient une autre interprétation de deux aspects importants de la cognition spatiale : le rôle encore plus fondamentalement constitutif de l'espace dans la construction même des objets (ici : textuels), et la sélection de l'information environnementale pour l'orientation et la planification de déplacements (la « navigation »).

L'objectif principal de l'étude conduite dans le cadre du projet financé par l'ACI Cognitive a été d'étudier de manière plus précise les modalités d'organisation spatiale de certains objets textuels des points de vues psycholinguistique et logico-informatique (modélisation). Il s'agit aussi de mettre en relation et d'articuler organisations spatiales et structures linguistiques, en nous intéressant à certains objets textuels particuliers. Le choix des objets textuels étudiés, énumérations et titres, a été déterminé par la conjonction de considérations théoriques et applicatives : ces objets textuels *a)* sont des structures textuelles dont l'identité tient beaucoup à l'organisation spatiale et *b)* permettent de tester la validité de méthodes d'extraction automatique d'informations lexicales. Par ailleurs, ces objets sont :

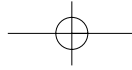
1) très fréquents, notamment dans les textes techniques et scientifiques,



2) peu étudiés et donc peu formalisés (il n'existe pas de règles de bonne formation attestées),

et 3) souvent à la limite de la frontière qui signale les ressources spécifiques de la spatialisation (*i. e.* ils appartiennent à la fois à la structure visuelle et à la structure discursive du texte).

Ce vaste ensemble de questions n'a été abordé que dans le cadre d'un spectre plus étroit présenté ci-dessous. Des raisons évidentes de faisabilité, mais aussi de méthode, ou encore relatives aux compétences réunies en relation avec les dimensions à prendre en compte, ont conduit à s'intéresser principalement à un seul modèle de structures visuelles, le modèle d'architectures textuelles (MAT) et à un seul type d'unité textuelle, l'énumération, objet textuel révélateur du rôle exact des propriétés visuelles (1^{re} partie, sous-division 1). MAT, modèle générique de représentation, a été exploité, et par le fait mis à l'épreuve et enrichi, dans deux contextes : d'une part celui de l'analyse automatique de textes sur des bases de linguistique computationnelle et de traitement de l'information, en vue du repérage des structures énumératives (1^{re} partie, sous-division 3), d'autre part celui de l'acquisition des relations sémantiques encadrées par ces structures (1^{re} partie, sous-division 5). Ces deux contributions s'appuient sur l'existence de la modalité numérique du texte, et de son mode d'accès sur Internet sous forme de pages Web, qui jouent ici la double fonction d'autoriser des recherches inatteignables «à la main», et de constituer le domaine même d'expérimentation et d'exploitation de ces recherches. D'autre part, les contributions psycholinguistiques permettent d'évaluer la plausibilité cognitive du rôle des facteurs visuels dans la constitution d'une composante spatiale dans les représentations mentales des textes (2^e partie), et dans l'impact d'une telle composante dans la mémorisation/compréhension de textes chez des enfants (1^{re} partie, sous-division 2). Des épreuves classiques de rappel, comme le suivi oculaire en activité de lecture, ou encore des jugements de «bonne formation» de structures énumératives révèlent l'existence chez les lecteurs adultes, et l'acquisition en développement, chez des enfants de 11 ans, d'une compétence visuelle en compréhension de textes, ignorée par les modèles cognitifs sémantiques classiques.



L'ÉNUMÉRATION : UN OBJET TEXTUEL INTÉRESSANT
POUR ANALYSER LES STRUCTURES VISUELLES DES TEXTES

Un modèle générique de structures visuelles des textes :
MAT (modèle d'architectures de textes)

Dans le modèle d'architectures de textes (MAT) (Virbel, 1989 ; Pascual, 1991 ; Virbel et Luc, 2001), est identifié un sous-ensemble des propriétés visuelles d'un texte qui jouent un rôle sémantique (sa *mise en forme matérielle*) et caractérisent des *objets textuels* qui forment l'*architecture* de ce texte. La MFM est le produit de la réduction avec traces typodispositionnelles de contreparties discursives performatives associées à un texte en son sein. MAT formalise ces analyses par un ensemble de schémas métaphrastiques de forme prédicative dont les arguments sont structurés. Le métatexte associé à un texte donné est constitué par l'instantiation de telles métaphrases de MAT ; sa bonne formation est contrôlée par deux relations interphrastiques de dépendance (obligation et précédence). Environ 600 verbes et expressions verbales, représentant les actes élémentaires d'architecturation (résumer, traduire, diviser, énumérer, titrer, etc.) ont fourni une cinquantaine de classes d'objets textuels au sein du modèle.

L'exemple suivant illustre certains des phénomènes architecturaux étudiés dans le cadre de ce projet¹ :

·
·
·
·
M _____
_____ .

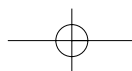
Est considéré comme « lecture savante », du point de vue fonctionnel, une pratique de lecture répondant aux critères suivants :

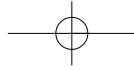
- c'est une lecture qualifiée,
- qui se développe sur le temps long de la recherche scientifique,
- dans un parcours forcément individualisé,
- où l'écriture se combine à la lecture, souvent dans une perspective de publication.

M _____

_____ .

1. Ce projet a été mené en collaboration avec Christophe Luc.





Dans cet exemple, l'énumération est formée de quatre *items*², et la *structure énumérative*, de l'*amorçe* (la phrase introductrice précédant l'énumération) et de l'*énumération*. Il présente un cas où les items sont en relation de dépendance les uns par rapport aux autres : le deuxième item dépend syntaxiquement du premier, le troisième dépend du second et le quatrième du troisième. Le non-parallélisme vient du fait que les items ne sont pas fonctionnellement équivalents au sein de l'énumération.

Étude d'énumérations non parallèles

Pour étudier ces énumérations non parallèles, nous avons constitué un recueil de 75 énumérations présentant divers aspects de non-parallélisme. Les trois catégories établies sont représentatives de plusieurs centaines (voire milliers) de cas rencontrés. En effet, les énumérations sont le plus souvent appréhendées (il n'existe pas de définition attestée) comme un moyen de mettre en relief de façon identique des objets ou entités de même importance. Lorsque les constituants de l'énumération (*i. e.* les items) sont fonctionnellement et textuellement équivalents (*i. e.* les items sont sur le même plan structural) au sein de cet objet, on parle d'énumération parallèle. Cette vision ne tient pas compte des possibilités offertes par l'écriture (et largement utilisées) pour réaliser des énumérations présentant des formes variées.

À partir du corpus, nous avons identifié 3 catégories de problèmes liés aux énumérations :

- 1) des énumérations dont les items entretiennent des relations diverses entre eux,
- 2) des énumérations dont les items ne sont pas structurellement équivalents,
- et enfin 3) des énumérations dont un ou des constituants entretiennent des relations avec un ou des objets textuels extérieurs à la structure énumérative.

Voici la liste des types des 3 catégories :

Catégorie 1 : *syntagmatique – paradigmaticque – hybride.*

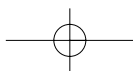
Catégorie 2 : *homogène – hétérogène.*

Catégorie 3 : *liée – isolée.*

Nous pouvons poser les définitions suivantes :

– *énumération parallèle* : énumération paradigmaticque, homogène et isolée ;

2. Les mots en italique donnent lieu à des définitions précises conformément au modèle (Pascual, 1991 ; Virbel, 1989 ; Luc, 2000).



– *énumération non parallèle* : énumération qui n'est pas paradigmatique, ou qui est homogène ou liée.

L'exemple cité plus haut relatif à la définition de la lecture savante présente une énumération syntagmatique, homogène et isolée.

Les types d'énumérations dans chaque catégorie sont exclusifs, mais une énumération relève des trois types, chacun des types appartenant à une catégorie différente. Ces types sont enfin définis de façon formelle à l'aide des modèles de la RST (*Rhetorical Structure Theory*) et de l'architecture textuelle.

Caractérisation de structures énumératives « tout-venant » dans un corpus diversifié

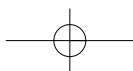
À la suite de l'étude des structures énumératives non parallèles, il a semblé nécessaire de prendre la question par l'autre bout, et d'examiner les structures énumératives dans un corpus construit en termes de genres discursifs et de domaines (voir Biber, 1993 ; Péry-Woodley, 2001). L'objectif est ici de tester le modèle descriptif sur des occurrences tout-venant, de l'enrichir si besoin, et d'examiner les variations liées au genre discursif et/ou au domaine. Par ailleurs, comme certaines formes, en particulier les énumérations syntagmatiques, sont, dans le corpus initial, fréquentes en français et très rares en anglais, nous avons décidé de constituer un corpus bilingue anglais-français.

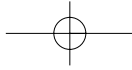
Le tableau 1 présente la composition du corpus :

| Catégorie | Langue | Taille (ko) | Nombre d'unités |
|--|------------------|-------------|-------------------------------|
| Articles scientifiques | Anglais | 886 | 15 |
| Textes procéduraux | Anglais/français | 946 | Anglais : 23 Français : 28 |
| Pages de présentation de laboratoires de recherche | Anglais/français | 609 | Anglais : 18 Français : 20 |
| Pages d'accueil de sites Web | Anglais/français | 1100 | Anglais : 30 Français : 28 |

Tableau 1 – composition du corpus.

Cette composition a été en partie déterminée par un questionnaire manifesté à plusieurs reprises dans notre projet : « vérifier les dimensions



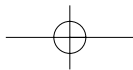


du renouvellement qu'apporte la modalité numérisée du texte, et ses formes de présentation et d'accessibilité (par exemple sur Internet)». Il s'agit donc d'examiner la spécificité des pages en ligne sur le plan des structures spatio-linguistiques. Les catégories du corpus forment un continuum allant du moins caractéristique au plus caractéristique du Web – depuis les articles scientifiques disponibles sur Internet mais rédigés et mis en forme pour la lecture sur support papier, jusqu'aux pages d'accueil de sites Web, conçues spécifiquement pour un usage en ligne. Le recueil de pages Web vise également à explorer l'utilisabilité du balisage HTML pour un repérage automatique des structures énumératives (sachant que la souplesse même d'HTML fait que le balisage est loin d'être univoque).

L'analyse prend comme point de départ les marques identifiées dans le modèle descriptif initial, en distinguant celles qui sont régulièrement associées à l'amorce, et celles qui organisent les items de l'énumération. L'objectif de cette phase du travail est de construire des patrons d'analyse qui permettront de procéder à un repérage automatique des structures énumératives. Ces patrons conjuguent marqueurs ponctuationnels, typodispositionnels et lexico-syntaxiques. Pour ces derniers sont identifiées des classes lexicales définies en extension : organisateurs, noms dénotant la méronymie, verbes dénotant la segmentation ou le rassemblement. Ces ensembles lexicaux sont en fait d'ordre différent : là où les listes de verbes ont été constituées en langue, et visent l'exhaustivité, il est clair que la relation d'hyponymie ou de méronymie qui est construite par l'acte d'énumérer ne peut faire l'objet d'une description lexicale exhaustive. En ce qui concerne les marqueurs typodispositionnels, sont prises en compte non seulement les balises explicitement et intentionnellement (dans la description du langage) associées aux énumérations – les balises de listes , , –, mais aussi les balises de structures tabulaires, et en conjonction avec d'autres marqueurs, des balises «généralistes» comme le changement de paragraphe.

Résultats

L'application des patrons a permis d'identifier dans le corpus 650 structures énumératives. Une validation manuelle a conduit à corriger certains silences, en particulier par des ajouts aux patrons lexicaux. On a constaté que les patrons les plus productifs étaient les introducteurs (surtout les déictiques *suivant, following*), les verbes de rassemblement et de segmentation, les annonceurs (surtout numériques), les organisateurs, etc. et équivalents. L'utilisation de la ponctuation de fin d'amorce permet de limiter grandement le bruit et crée peu de silence. Les résultats sont mitigés en



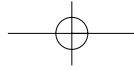
ce qui concerne les balises HTML généralistes, qui nécessitent d'être associés à des patrons plus restrictifs (lexico-syntaxiques) pour être utilisables.

Les structures énumératives horizontales, sans marquage visuel et se terminant le plus souvent par *etc.* ou équivalent, se retrouvent de façon homogène dans les trois sous-corpus (le corpus de pages Web sera traité à part). Les structures énumératives verticales sont systématiquement présentes dans les articles scientifiques, très présentes dans les présentations de laboratoires de recherches (souvent organisées de façon taxinomique), mais leur utilisation n'est pas du tout systématique dans les textes procéduraux.

Il ressort de l'analyse de ces 650 structures énumératives «tout-venant» que les structures paradigmatiques parallèles dominent largement (608 sur 650). La relation de co-énumérabilité établie par la structure peut être de type hyperonymique ou méronymique. Selon les termes de la RST, la relation entre amorce et énumération est une relation d'élaboration.

Le cas des pages d'accueil de sites Web

On ne peut qu'être frappé dans les pages d'accueil par la prédominance de structures visuelles qui «ont l'air» de structures énumératives : mais la disposition verticale d'items souvent précédés d'une puce suffit-elle à faire une structure énumérative ? Dans un document hypertextuel (ou hypermédia), la distinction entre énumération et sommaire s'estompe : la référence à la pagination, typique du sommaire, disparaît, chaque item étant un lien cliquable directement. Dans cette fonction de sommaire, on ne trouve pas d'amorce, le critère de co-énumérabilité étant présupposé par la fonction même (parties du document). Mais les éléments du sommaire sont en même temps des titres qui fonctionnent comme des amorces minimales, suivis d'énumérations horizontales ou verticales. On trouvera ainsi à la suite des titres qui fournissent le sommaire de la page, des items eux-mêmes cliquables rassemblés par le critère de co-énumérabilité explicité par le titre (Services → Librairie, Emploi, Forums). En fait, même les parties plus «profondes» des pages, éloignées du sommaire initial, sont absolument dominées par ces structures d'énumération d'items regroupés sous un titre. C'est donc la relation entre structure énumérative et sommaire, entre amorce et titre, qui est à repenser dans le cadre de documents non linéaires, où la page n'est plus un espace clos.

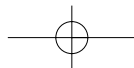


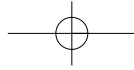
Analyse automatique des énumérations et applications

Le MAT offre un cadre générique d'analyse et de représentation de la structure spatiale d'un document et son explicitation dans le cadre d'une sémantique textuelle (Virbel, 1989; Pascual, 1991; Luc, 2000). Une exploitation informatique de ce modèle peut se percevoir sous plusieurs angles : utiliser le métatexte associé aux structures spatiales des textes pour en faire une interprétation automatique qui intègre le métatexte associé à la mise en forme spatiale (Pascual et Péry-Woodley, 1999), aider à la génération de mises en formes textuelles en explicitant le rôle interprétatif de la structure et aboutir ainsi à un formatage intelligent des documents (Luc, 2000) ou, dans une approche de linguistique de corpus (Biber, 1993; Habert 1998), utiliser des masses de documents pour y analyser un paradigme de formatage et en extraire les connaissances associées.

C'est cette dernière approche que nous avons choisie, puisqu'elle nous semble à la fois raisonnable dans ses objectifs pour un projet court et prometteuse, dans la mesure où elle permettait une observation large d'un paradigme précis. Nous avons choisi d'analyser des contextes définitoires et, plus particulièrement, les énumérations dont la structure est généralement bien marquée et facilement identifiable (Péry-Woodley, 1998). L'interprétation de la structure énumérative offre deux facettes intéressantes pour l'acquisition automatique de relations sémantiques : d'une part une relation d'hyponymie/hyperonymie s'instancie entre l'amorce d'une énumération et ses items et, d'autre part, le parallélisme entre les items permet des rapprochements ou des complétions de rôles entre les items d'une énumération. Seule la première relation a été utilisée afin d'acquérir automatiquement des liens de généralité sur des gros volumes de données textuelles, afin de constituer des bases de connaissance parcellaires reposant sur des observables en corpus.

Ainsi, l'exemple ci-contre comporte une amorce constituée d'un marqueur discursif d'énumération (*following, such as* ou *list of*) et d'un hyperonyme typant les éléments attendus dans l'énumération (*international organization, political party, actress, journalist, car maker, river, city...*). Lorsque son analyse réussit, cette énumération produit donc une liste de trois *organisations internationales* avec, dans cet exemple, leur forme abrégée et leur forme étendue :





The following international organizations are collaborating on the Project :

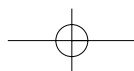
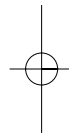
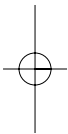
- International Commission on Non-Ionizing Radiation Protection (ICNIRP)
- International Agency for Research on Cancer (IARC)
- United Nations Environment Programme (UNEP)

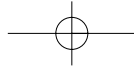
Résultats

Le travail effectué dans le cadre du projet Cognitique consiste, sur son versant linguistique, à approfondir les travaux antérieurs de description de la structure des énumérations. Nos travaux en linguistique étant orientés vers l'analyse automatique et, ayant remarqué que l'identification automatique d'une énumération repose sur une reconnaissance précise de son amorce et sur la détection de la fin de l'énumération, nous avons fait un effort particulier pour décrire précisément la structure des amorces d'énumération (Bush, 2002).

La reconnaissance de l'énumération proprement dite repose sur l'identification d'une structure énumérative (du type de celle de l'exemple précédent avec puces, retours à la ligne et retraits) et sur l'analyse syntaxique parallèle des items. Pour simplifier la reconnaissance de la structure typodispositionnelle de l'énumération, nous avons choisi de travailler sur un corpus de pages HTML (des pages tirées de site Internet). Dans ces pages, les structures sont marquées soit directement par les balises `` et ``, soit indirectement par des retours à la ligne (`
`) et des marqueurs typographiques (`-`, `–`, `—`, `*` ou `•`), voire par une table dont chaque ligne est un item de l'énumération.

Le système de collecte et d'analyse automatique des énumérations présenté dans Jacquemin et Bush (2000) est donc composé d'un collecteur de pages Web susceptibles de contenir des énumérations via des requêtes sur un moteur de recherche (un *moissonneur*), d'un analyseur d'énumérations au sein des pages ainsi collectées, et de filtres produisant en sortie des listes d'entités nommées classées par types. Un corpus de 3 Go de code HTML a ainsi été collecté à partir de requêtes construites sur 34 hyperonymes et les trois marqueurs d'énumération précédents. L'analyse automatique de ces pages HTML pour y repérer les amorces et les items des énumérations associées a fourni 31 759 entités nommées candidates, chacune étant associée à un hyperonyme particulier. Une analyse manuelle de 504 d'entre elles a montré que la précision de l'acquisition de ces entités varie entre 63 % et 74 % selon les critères de correction choisis.





VALIDATION DU MODÈLE DE MAT
PAR DES EXPÉRIMENTATIONS PSYCHOLINGUISTIQUES

La place de la structure spatiale en psychologie cognitive

En psychologie cognitive, la structure spatiale du texte n'est pas prise en compte dans les modèles de compréhension (Gernsbacher, 1989 ; Kintsch, 1998) qui postulent l'élaboration d'une représentation mentale du contenu du texte en cours de lecture. Pourtant, le codage spatial évoque l'importance du repérage de la position des mots dans une page en fonction de points de référence tels que les limites de la page ou du texte (Baccino et Pynte, 1998).

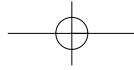
Selon le modèle MAT, la mise en forme matérielle MFM (Pascual, 1991 ; Virbel, 1989) apporterait des informations métatextuelles utiles au lecteur pour organiser l'intégration des informations. Elle offre des repères visuels qui modifient leur stratégie de lecture de textes procéduraux (Schmid et Baccino, 2001). Les marques visuelles de MFM jouent un rôle sémantique et, à ce titre, sont impliquées dans l'élaboration de la représentation mentale des textes (Schmid, 2001). Les résultats de plusieurs recherches mettent en évidence l'élaboration d'une *représentation spatiale*, dont l'influence varie selon la nature du texte (narration, description, procédure).

L'étude des structures énumératives pour évaluer les effets cognitifs des éléments typodispositionnels en référence au modèle MAT (Pascual, 1991) est rendue pertinente par les travaux antérieurs de définition et de caractérisation de ce type d'unité textuelle (Luc, 2000 ; Bush, 2000). En effet à partir de ces travaux, la manipulation des caractéristiques linguistiques et visuo-spatiales fondamentales des énumérations permet des recherches expérimentales en psychologie cognitive.

Les questions posées portent sur la discrimination d'une énumération dans un texte par les lecteurs experts. Sur quelles composantes d'une structure énumérative les lecteurs se fondent-ils ? La définition d'une « énumération canonique » établie en linguistique a-t-elle une réalité psychologique, et quels types de déviations par rapport à cette structure initiale perturbent particulièrement les lecteurs ?

Une étude expérimentale auprès de lecteurs experts :
tâche de jugement et étude du comportement oculomoteur

Selon la modélisation linguistique exposée plus haut (1^{re} partie, sous-divisions 1 et 2), une structure énumérative canonique minimale



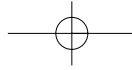
comporte une phrase amorce ou introductive, suivie de plusieurs items indépendants et équivalents du point de vue lexico-syntaxique ou de mise en forme (dispositionnelle ou typographique). Cependant, les énumérations non canoniques dites non parallèles sont fréquentes dans tous types de textes. Des irrégularités apparaissent portant sur l'absence d'amorce ou sur la non-équivalence des items entre eux, qu'elle soit de nature lexico-syntaxique ou de mise en forme.

Nous posons l'hypothèse que les lecteurs se fondent sur des indices de mise en forme pour discriminer une énumération.

Afin de tester cette hypothèse, 93 étudiants de l'université de Nice Sophia Antipolis ont participé à une expérience, dont la tâche consistait en l'évaluation de la bonne formulation d'énumérations sur une échelle en cinq points. Le matériel expérimental était composé de 25 énumérations non parallèles issues du corpus de l'IRIT (Virbel, 1999) présentant des variations par rapport à la forme canonique d'un point de vue *dispositionnel*, *typographique* ou *lexico-syntaxique*, et de 10 énumérations parallèles issues de la littérature. Une analyse de régression multiple a été menée visant à discriminer les trois types de marques et à mettre en évidence celles qui prédisent les jugements d'acceptabilité des énumérations réalisés par les sujets. Les résultats de cette analyse mettent en évidence que les marques typographiques ($B\hat{e}ta = 0.25$) et lexico-syntaxiques ($B\hat{e}ta = 0.07$) ne prédisent pas de manière significative les jugements. À l'inverse, *les marques dispositionnelles* ($B\hat{e}ta = 0.57$) prédisent significativement le jugement $t(27) = 3.73$ $p < .01$. Les sujets se fondent majoritairement sur des indices dispositionnels pour juger l'acceptabilité d'une énumération. Lorsque l'unité textuelle à évaluer ne présente pas d'indices dispositionnels adéquats, les lecteurs jugent que l'énumération est mal formulée. Les indices lexico-syntaxiques et typographiques ne sont pas utilisés par les lecteurs dans cette tâche d'évaluation. Les indices dispositionnels pourraient donc être utiles aux lecteurs pour repérer les énumérations dans un texte et pour anticiper la gestion de l'encodage de telles unités textuelles.

Une seconde expérience utilisant l'enregistrement des mouvements oculaires lors de la lecture d'énumérations parallèles, présentées sans aucun indice de mise en forme *vs* avec des marques dispositionnelles, permet de valider cette hypothèse (Closson, 2001).

Les mouvements oculaires sont enregistrés sur l'ensemble des textes composés d'une amorce et d'une liste d'items énumérés. La moitié des textes est présentée avec une mise en forme (retour à la ligne et tiret devant chaque item énuméré) et l'autre moitié sans mise en forme (tous



les items sur une même ligne séparés par des virgules). Quarante textes expérimentaux ont été présentés à 27 sujets.

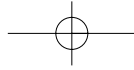
Une analyse de variance menée sur l'ensemble des résultats révèle un effet significatif de la mise en forme ($F(1,26) = 9,25$ $p < .01$). Les lecteurs fixent plus longtemps les items énumérés présentant une marque dispositionnelle (retour à la ligne et indentation) que ceux qui ne présentent aucune marque. Ce type d'indices induit une lecture plus attentive de l'énumération. Ce temps de fixation supplémentaire doit permettre un encodage immédiat des informations énumérées. Cette stratégie d'encodage a déjà été observée en présence d'indices de structuration des textes (Lorch *et al.*, 2001 ; Schmid et Baccino, 1998, 2001). La lecture de l'amorce n'est pas affectée par la présence d'indices dispositionnels au niveau des items énumérés. Cette absence de différence s'explique par la présence d'un saut de ligne entre l'amorce et les items énumérés dans les deux conditions expérimentales. L'effet de fin de paragraphe ne peut donc être mis en évidence sur l'amorce.

La lecture d'un texte n'est pas linéaire, l'élaboration d'une représentation mentale cohérente du contenu du texte en cours de lecture nécessite la mise en correspondance des informations issues du texte et de la base de connaissances du lecteur. Le lecteur organise sa prise d'information en utilisant les indices mis en place par le rédacteur pour faciliter ce travail intégratif. Ces indices de nature linguistique, dispositionnelle ou typographique provoquent des changements de rythme dans la lecture qui révèlent leur importance pour le lecteur. Dans le cas particulier des énumérations, les marques dispositionnelles sont utilisées par le lecteur pour repérer l'unité textuelle et pour adapter sa stratégie d'encodage. L'importance de l'organisation de la prise d'information en cours de lecture à partir de la structure dispositionnelle d'un texte est confirmée par ces deux expériences. Les indices de MFM permettent ainsi au lecteur de distinguer les titres, les alinéas, etc., et l'étude du comportement des lecteurs sur chacune de ces unités textuelles mettrait sûrement en évidence des stratégies de lecture adaptées à leur rôle dans le texte, comme dans le cas des énumérations.

Une étude expérimentale auprès d'enfants de 11 ans :

incidence de la mise en forme matérielle (MFM) des textes
sur la compréhension/mémorisation de règles de jeu et de textes
de vulgarisation scientifique, et sur le traitement cognitif des énumérations

L'objectif de cette étude, dans le prolongement de travaux antérieurs (Garcia-Debanc et Grandaty, 2001) est d'évaluer les effets de mises en



forme différentes sur l'espace de la page d'un contenu sémantique constant. Le protocole expérimental permet d'évaluer :

– d'une part, à un niveau global qu'est le texte, l'incidence des variations de la mise en forme matérielle d'un texte sur la compréhension et la mémorisation de celui-ci,

– et d'autre part, à un niveau local qu'est la structure énumérative, l'incidence de la mise en forme matérielle d'une énumération sur le traitement cognitif des informations apportées par celle-ci.

Les questions auxquelles nous nous sommes efforcés de répondre sont les suivantes :

– dans quelle mesure et à partir de quel âge la mise en forme matérielle d'un texte est-elle prise en compte par les sujets dans le traitement cognitif de ce texte et des objets textuels qui le constituent ?

– Parmi les diverses versions typodispositionnelles possibles pour une même architecture textuelle, laquelle est la plus efficace pour son traitement (compréhension/mémorisation/recherche sélective de renseignements ponctuels) ?

Les hypothèses de travail sont les suivantes :

– À la suite de Schmid (2001), Schmid et Baccino (2001), nous postulons un effet de la mise en forme matérielle (MFM) du texte sur sa compréhension/mémorisation. Notre première hypothèse de travail est donc que les variations de MFM portant sur la structure d'un texte affectent l'interprétation et le stockage en mémoire des informations apportées par ce texte.

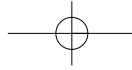
– À la suite de Luc (2000), nous nous sommes intéressés aux structures énumératives qui, relevant de la dimension linguistique et typodispositionnelle, constituent un objet textuel particulièrement intéressant pour une étude de l'effet de la MFM d'un objet textuel sur son traitement cognitif. Notre seconde hypothèse de travail est que les variations de MFM portant sur une structure énumérative présente dans un texte affectent le traitement des informations contenues dans cet objet textuel.

Le protocole a pour objectif de :

a) mesurer l'impact de chaque version de texte sur son traitement cognitif,

et b) déterminer la version la plus efficace pour la compréhension/mémorisation.

Le matériel expérimental est constitué de 4 versions d'un texte présentant une architecture textuelle constante, exprimée spatialement de manières différentes. Le modèle d'architecture textuelle permet le contrôle des différentes versions de texte lors de leur déclinaison (Pascual,



1991). Deux genres textuels différents ont été choisis, du fait de leur forte structuration textuelle : les textes de vulgarisation scientifique et les textes procéduraux (Garcia-Debanc, 2002). Les variables retenues ont été les suivantes :

- au niveau du texte :
segmentation en parties et sous-parties *vs* absence de segmentation,
détachement des titres avec ou sans réduction du discours *vs* absence
d'extraction des titres ;
- au niveau de la structure énumérative :
amorce syntaxiquement complète *vs* amorce syntaxiquement
incomplète,
amorce avec nombre *vs* amorce sans nombre.

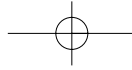
Les textes expérimentaux ont été élaborés à partir de textes écolologiques, publiés dans des revues ou dans des ouvrages documentaires pour enfants. La grille d'analyse mise au point par J. Virbel permet de faire une analyse de ces textes pour aboutir à une réécriture présentant les informations de manière systématique et homogène (Dominguez, 2001). À partir de la version de travail ainsi élaborée, est fabriquée ce qui est appelé une version « étendue » (ou « prototexte »), dans laquelle les fonctions de structuration textuelle sont développées discursivement, faisant ainsi apparaître l'architecture du texte (texte + métatexte).

À partir du prototexte, est rédigée la version développée du texte. Elle se caractérise par l'absence de toute segmentation et l'intégration des titres au corps du texte. L'expression de la structure énumérative se caractérise par la présence d'organiseurs d'intégration linéaire au niveau des items énumérés.

De la version développée est dérivée la version réduite syntaxiquement ou version typodispositionnelle. Elle se caractérise par l'exploitation de l'espace du texte (extraction du titre, blanc horizontal et vertical). L'expression de la structure énumérative se caractérise par l'itémisation des éléments énumérés : les organisateurs sont ici de nature typographique et spatiale.

Ces manipulations contrôlées aboutissent à la production des 4 versions ci-contre (fig. 1) pour un même contenu informatif (pour une synthèse des précautions méthodologiques observées lors du travail d'élaboration des textes expérimentaux, voir Dominguez, 2002).

Chaque sujet voit chaque texte sous une mise en forme différente. L'ordre d'apparition des textes est contrôlé afin d'éviter tout biais dû à leur ordre de présentation. Trois épreuves ont permis de tester la mémorisation : le rappel libre oral, le rappel libre écrit et le rappel indicé. En

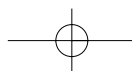


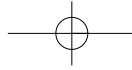
| Versions développées | |
|--|---|
| Amorce syntaxiquement incomplète | Amorce syntaxiquement complète |
| <i>Le matériel du jeu comprend d'une part 1 plateau de jeu percé de 49 trous, d'autre part 2 dés et enfin 48 pions constitués de 24 pions clairs et 24 pions foncés.</i> | <i>Le matériel du jeu comprend trois types d'éléments : d'une part 1 plateau de jeu percé de 49 trous, d'autre part 2 dés et enfin 48 pions constitués de 24 pions clairs et 24 pions foncés.</i> |
| Versions réduites | |
| Amorce syntaxiquement incomplète | Amorce syntaxiquement complète |
| Matériel Il comprend : 1 plateau de jeu percé de 49 trous, 2 dés, 48 pions (24 pions clairs et 24 pions foncés). | Matériel Il comprend 3 types d'éléments : 1 plateau de jeu percé de 49 trous, 2 dés, 48 pions (24 pions clairs et 24 pions foncés). |

Figure 1 – Exemple : extrait de la règle du jeu Avalam

dernier lieu, pour vérifier la compréhension des règles de jeu, l'expérimentateur joue avec l'enfant (Dominguez, 2001) ou, pour les textes de vulgarisation scientifique, lui fait choisir le schéma pertinent parmi plusieurs schémas.

L'expérimentation, menée auprès d'un public d'enfants de 11 ans et d'adultes, a permis de confirmer une partie des hypothèses concernant une incidence des variations typodispositionnelles sur la mémorisation de textes (règles de jeu et textes de vulgarisation scientifique). Le nombre d'informations rappelées dans le rappel libre est plus important avec les versions réduites riches en marques typodispositionnelles (38,99 % pour les règles de jeux, 25 % pour les textes explicatifs) qu'avec les versions développées (34,02 % pour les règles de jeux, 22 % pour les textes explicatifs) (Carrio, 2002 ; Dominguez, 2002). Il est à noter, pour ce qui est de la règle de jeu, que les items relatifs aux différentes possibilités d'actions dans le jeu, placées au milieu de la règle du jeu, sont globalement moins bien rappelées que les items relatifs au matériel du jeu placés au début de la règle de jeu, et ce particulièrement pour les versions réduites typodispositionnelles (Garcia-Debanco et Grandaty, 2001).





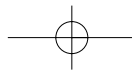
Les versions réduites favorisent le rappel des items des énumérations de rang 1 dans les textes explicatifs : 48 % d'éléments rappelés dans les textes explicatifs en version réduite contre 25 % pour les versions développées. Le traitement cognitif des énumérations contenues dans ces textes est facilité par les présentations riches en marques typodispositionnelles. Les versions réduites apparaissent donc comme étant les plus performantes du point de vue du nombre d'informations rappelées, au niveau textuel et au niveau de l'énumération.

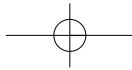
D'un point de vue plus qualitatif, l'influence de la mise en forme du texte est perceptible dans l'organisation des informations dans le rappel oral. En effet, les enfants ayant lu la version de texte réduite typodispositionnelle organisent, pour la plupart d'entre eux, leur rappel par blocs d'informations, qui correspondent environ aux unités constituées par les rubriques. La mise en forme typodispositionnelle semble favoriser ce mode de rappel à cause de l'extraction des titres, les titres de rubriques pouvant servir par ailleurs d'indices de récupération.

À un niveau plus local qu'est la structure énumérative, les données recueillies montrent que le type d'amorce semble également avoir une incidence sur la mémorisation de l'énumération. Cela se vérifie en particulier chez les enfants. Les versions caractérisées par une amorce syntaxiquement complète, avec nombre pour signaler les différents items, suscitent de meilleurs résultats pour le nombre d'informations rappelées (39 % d'informations rappelées, contre 34 % pour les versions avec amorce incomplète). Les informations erronées sont moins nombreuses avec les versions contenant des structures énumératives avec amorce complète. L'étude qualitative fait apparaître que l'amorce en tant que telle est rappelée aussi bien à l'oral qu'à l'écrit.

CONCLUSIONS ET PERSPECTIVES

Le modèle modèle d'architecture textuelle (MAT) a joué un rôle central dans les études conduites dans le cadre de ce projet. Les résultats obtenus par l'observation de corpus variés, papier ou pages Web, ainsi que les expérimentations psycholinguistiques mises en œuvre, ont permis de valider le modèle et de l'affiner. D'un point de vue logico-linguistique, la recherche conduite dans le cadre de l'ACI Cognitive Espace a permis de modéliser les structures énumératives, d'en proposer une typologie et de proposer un mode de composition de modèles de structuration des textes, MAT et RST (Luc, 2001) pour rendre compte des structures énumératives.





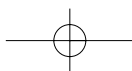
Du point de vue de la mise en œuvre de méthodes informatiques d'exploitation des structures énumératives dans les documents, les principaux résultats obtenus sont un analyseur de titres et d'énumérations, un analyseur des amorces, et des patrons de fouilles pour le recueil automatique des énumérations.

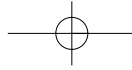
D'un point de vue psycholinguistique, les expériences menées au cours de ce projet ont fait avancer la connaissance des processus cognitifs liés au traitement des énumérations selon leurs représentations spatiales, chez les adultes et chez les enfants. Nous avons pu mettre en évidence l'importance de la représentation spatiale du texte pour la lecture et la mémorisation d'énumérations parallèles par des adultes et par des enfants. La MFM semble permettre un encodage particulier de ce type d'unité textuelle. La poursuite de ce travail est envisagée par des expériences portant sur la comparaison de la lecture et de la compréhension d'énumérations non parallèles et parallèles. Nous avons pu également montrer l'influence de la MFM sur la compréhension et la mémorisation d'informations contenues dans des règles de jeux et des textes explicatifs par des enfants de 11 ans.

Ces recherches, qui n'en sont qu'à leur début, peuvent connaître des développements applicatifs importants dans les domaines considérés, notamment dans le domaine de l'édition et de la présentation de pages Web. Elle se sont poursuivies sur d'autres objets textuels, notamment les titres et les définitions. Ces travaux pourraient aussi corroborer le jugement émis récemment (Parisse, 2003), selon lequel les recherches sur l'écrit mettent en cause la nature même du langage dont s'occupent habituellement les linguistes, ni vraiment écrit, ni non plus vraiment oral.

Références bibliographiques

- Baccino, T. et J. Pynte. 1998. «Spatial Encoding and Referential Processing during Reading». *European Psychologist*, 3/1 : 51-61.
- Biber, D. 1993. «Representativeness in Corpus Design». *Literary and Linguistic Computing*, 8 : 243-257.
- Bush, C. 2000. *Analyse des déclencheurs des énumérations d'entités nommées sur le Web*. Technical Report 2000-05, LIMSI-CNRS, mars.
- Bush, C. 2002. «Analysis of Named Entity Enumerations on the Web. Computer and the Humanities».
- Carrio, L. 2001. *Incidence des variations typodispositionnelles sur la compréhension/mémorisation de textes de vulgarisation scientifique et le traitement cognitif des énumérations*





contenues dans ces textes par des enfants de 9 à 11 ans : élaboration d'un protocole expérimental. Mémoire de DEA, sciences du langage, université Toulouse-Le Mirail.

Carrio, L. 2002. « Incidence des variations typodispositionnelles sur la compréhension/mémorisation de textes de vulgarisation scientifique et le traitement cognitif des énumérations contenues dans ces textes par des enfants de 9 à 11 ans ». In Journées de PRESCOT à l'initiative et avec le soutien de l'ACI Cognitive, *Inscription spatiale du langage : structures et processus*, 29-30 Janvier 2002, IRIT Toulouse. Toulouse, PRESCOT : 207-220.

Closson, B. 2001. *Effets de la mise en forme matérielle et de la prototypie catégorielle sur la lecture des énumérations parallèles*. Mémoire de DEA, université de Nice Sophia Antipolis.

Danos, L. 1985. *Génération automatique de textes en langue naturelle*. Paris, Masson.

Dominguez, C. 2001. *Structures spatio-linguistiques du texte : élaboration de textes expérimentaux pour un protocole de psycholinguistique*. Mémoire de DEA, sciences du langage, université Toulouse-Le Mirail.

Dominguez, C. 2002. « Étude de l'influence de la mise en forme matérielle d'un texte et d'un objet textuel particulier, la structure énumérative, sur leur traitement cognitif (compréhension/mémorisation) ». In Journées de PRESCOT à l'initiative et avec le soutien de l'ACI Cognitive, *Inscription spatiale du langage : structures et processus*, 29-30 Janvier 2002, IRIT Toulouse. Toulouse, PRESCOT : 223-232.

Garcia-Debanc, C. 2002. « Les genres du discours procédural : invariants et variations ». *Pratiques*, 111-112 (*Les textes de consignes*) : 65-76.

Garcia-Debanc, C. et M. Grandaty. 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 ». In C. Garcia-Debanc (éd.), *Les discours procéduraux*, *Langages* 141, mars 2001 : 92-104.

Gernsbacher, M. A. 1989. « Mechanism that Improve Referential Access ». *Cognition*, 32 : 99-156

Goody, J. 1986. *La logique de l'écriture*. Paris, Armand Colin.

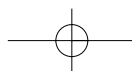
— 1993. *Entre l'oralité et l'écriture*. Paris, Presses universitaires de France.

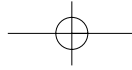
Hartley, J. 1980 *The Psychology of Written Communication*. Londres, Kogan Page.

Kintsch, W. 1998. *Comprehension a Paradigm for Cognition*. Cambridge MA, Cambridge University Press.

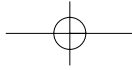
Jacquemin, C. et C. Bush. 2000a. « Fouille du Web pour la collecte d'entités nommées ». In E. Wehrli (éd.), *Actes, TALN'2000*, Lausanne, Suisse.

— 2000b. « Combining Lexical and Formatting Cues for Named Entity Acquisition from the Web ». In H. Schutze (éd.), *Actes, Joint Sigdat Conference On Empirical Methods In Natural Language Processing And Very Large Corpora (EMNLP/VLC-2000)*, Hong Kong, Chine.





- Lorch, R. F., E. P. Lorch, K. Ritchey, L. McGovern et D. Coleman. 2001. «Effects of Headings on Text Summarization». *Contemporary Educational Psychology*, 26 (2) : 171-191.
- Luc, C. 2000. *Représentation et composition des structures rhétoriques et visuelles du texte. Approche pour la génération de textes formatés*. Thèse de doctorat, université Paul Sabatier, Toulouse.
- Luc, C., M. Mojahid et J. Virbel. 2002. «Propriétés formelles du modèle d'architecture de texte». In J. Virbel et J.-L. Nespoulous (éds), *Inscription spatiale du langage : structures et processus*. Toulouse, PRESCOT [http://www.irit.fr/ACTIVITES/PRESCOT/Prescot.f.html].
- Mojahid, M. 1990. *Une architecture multi-agents en PAO pour le formatage de documents structurés*. Thèse de doctorat, université Paul Sabatier, Toulouse.
- Olson, D. 1994. *The World on Paper : the Conceptual and Cognitive Implications of Writing and Reading*. Cambridge MA, Cambridge University Press.
- Ong, W. 1982. *Orality and Literacy : the Technologizing of the Word*. Londres, Methuen.
- Parisse, C. 2003. «Le débat inné-acquis et le développement du langage à l'aube du XXI^e siècle». *Intellectica*, 35, 2002/2 : 269-286.
- Pascual, E. 1991. *Représentation de l'architecture textuelle et génération de texte*. Thèse de doctorat, université Paul Sabatier, Toulouse.
- Pascual, E. et M.-P. Péry-Woodley. 1999. «Définition et action dans les textes procéduraux». In J. Virbel, J.-M. Cellier et J.-L. Nespoulous (éds), *Cognition, discours procédural, action, vol. II*. Toulouse, PRESCOT : 80-108.
- Péry-Woodley, M.-P. 1998. «Signalling in Written Text : a Corpus-Based Approach». In M. Stede, L. Wanner et E. Hovy (éds.), *Actes, COLING 98 (Workshop on Discourse Relations and Discourse Markers)*. Montréal, ACL : 79-85.
- 2001. «Modes d'organisation et de signalisation dans des textes procéduraux». In C. Garcia-Debanco (éd), *Les discours procéduraux. Langages 141*, mars 2001 : 28-46.
- 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.
- Schmid, S. et T. Baccino. 1998. *Influence de la mise en forme des textes sur la compréhension*. Document électronique, Paris, Europia productions : 185-192.
- 2001. «Stratégies de lecture pour les textes à consigne». In C. Garcia-Debanco (éd), *Les discours procéduraux. Langages 141*, mars 2001 : 105-124.
- Virbel, J. 1989. «The Contribution of Linguistic Knowledge to the Interpretation of Text Structure». In J. André, V. Quint et R. Furuta (éds), *Structured Documents*, Cambridge MA, Cambridge University Press : 161-181.
- 1999. *Structures textuelles – planches, fascicule 1 : Énumérations*. Rapport IRIT, Toulouse.
- Virbel, J. et C. Luc. 2001. «Le modèle d'architecture textuelle : fondement et expérimentation». *Verbum*, XXIII, 1 : 103-123.



22 AGIR DANS L'ESPACE

Virbel, J. et J.-L. Nespoulous (éds). 2002. *Inscription spatiale du langage : structures et processus*. Toulouse, PRESCOT.

ATTENTION

**il reste une référence biblio manquante : Habert 1998 (cité p. 10)
une référence imprécise dans le texte : Jacquemin et Bush 2000a ou
2000b (cité p. 11) ?**

