

MASTER 2 HANDI

Nouvelles Technologies et Handicaps



LOGICIEL D'AIDE À LA LECTURE POUR LES PERSONNES DYSLEXIQUES

RAPPORT DE STAGE

AMDJAD BACAR

Responsables

Entreprise : Catherine PERIES

Université : J. LOPEZ KRAHE

DU 01 MARS AU 31 JUILLET 2006

Remerciements

Ce présent rapport serait incomplet sans un mot de remerciement pour tous ceux qui ont, de près ou de loin, contribué à l'élaboration de ce projet. Je voulais remercier :

- Tout particulièrement, Mme Catherine PERIES, responsable du département Troubles du langage de l'entreprise CECIAA aussi responsable de mon stage, de m'avoir fait bénéficier de son expérience dans le domaine des troubles du langage et m'accorder de son temps,*
- L'ensemble du personnel de l'entreprise CECIAA pour leur accueil et leur sympathique collaboration,*
- Mes collègues Sarah Schiller et Mballo Seck pour leur soutien pendant mon stage,*
- Et l'ensemble de mes camarades du Master.*

Sommaire

Introduction.....	4
I. Présentation de l'entreprise CECIAA.....	5
1. Historique.....	5
2. Organisation.....	5
3. Le secteur d'activité.....	6
4. Le département Trouble du langage.....	6
II. Les troubles spécifiques du langage.....	9
1. Généralité sur les troubles du langage.....	9
1.1. L'évolution du langage chez l'enfant.....	9
1.2. Les troubles de l'évolution du langage.....	10
a. Les troubles de l'apprentissage du langage oral.....	10
b. Les troubles de l'apprentissage du langage écrit.....	11
III. La Dyslexie.....	12
1. Définition.....	12
1.1. Les caractéristiques spéciales de la dyslexie.....	12
1.2. Les méthodes de traitement de la dyslexie.....	13
2. L'apport de l'informatique.....	14
IV. Les objectifs du projet DysBloc.....	15
1. Les adaptations des documents textes.....	15
1.1. Idée générale.....	15
1.2. Les recommandations sur l'accessibilité.....	16
2. Les fonctionnalités de DysBloc.....	19
2.1. Les fonctionnalités d'aide au déchiffrement.....	19
a. Les fonctions de paramétrage de l'accessibilité.....	19
b. Les fonctions de mise en évidence.....	20
2.2. Les fonctionnalités d'aide à la compréhension.....	20
V. Les solutions techniques.....	21
1. Les techniques de mise en page des fichiers textes.....	21
1.1. Les scripts et macros.....	21
1.2. Utilisations des API adaptées.....	21
1.3. Mise en page avec XML/XSL.....	22
2. Le XML et ses applications.....	22
2.1. Généralités sur le XML.....	22
2.2. Définition des balises XML.....	24
a. Le DTD et le XSD.....	24
2.3. Affichage des données XML.....	24
a. XML et CSS.....	25
b. XML et XSL.....	25

3. Le standard OpenDocument.....	27
3.1. Structure d'un fichier OpenOffice.org.....	28
3.2. Stockage du contenu.....	30
3.3. Stockage des informations de mise en page.....	32
3.4. Les extensions des fichiers associés.....	33
4. Comparaison avec WordML de Microsoft Office.....	34
VI. La structure de mise en page de DysBloc.....	36
1. Les solutions de mise en évidence choisies.....	36
2. Gestion des informations de mise en évidence.....	36
3. Analyse et traitement de texte.....	37
4. Les règles d'adaptation.....	39
4.1. Définitions.....	39
a. Une règle d'adaptation	39
b. Une transformation	39
4.2. Structure d'une règle d'adaptation.....	39
a. Les règles statiques.....	39
b. Les règles génériques.....	39
4.3. Gestion des règles d'adaptation.....	40
5. Interface de communication JAVA/XML.....	40
6. Analyse du contenu.....	41
VII. L'interface de DysBloc.....	42
Conclusion.....	45
Références.....	46
Annexes.....	47

Introduction

Aujourd'hui 8 % à 10 % des élèves scolarisés en France connaissent des troubles plus ou moins graves dus à la dyslexie (incapacité de lire, de reconnaître ou de reproduire correctement le langage écrit).

L'ordinateur est un outil précieux pour faciliter l'expression et la communication des personnes atteintes d'un handicap. L'usage de l'ordinateur avec des périphériques et des logiciels adaptés facilite énormément la scolarisation de ces personnes et peut dans un grand nombre de cas leur apporter une autonomie très importante.

Cependant, l'outil informatique ne doit pas être vu comme le remède miracle mais plutôt un complément.

Le projet DysBloc a été mis en place dans ce sens de complémentarité et d'assistance pour ce type de population. Il se présente comme un pont entre la détection des troubles qui nuisent à l'accès de l'information écrite et les méthodes qui visent à les éviter. DysBloc met l'accent sur la présentation de l'information écrite pour faciliter l'accessibilité. Cette procédure constitue une thérapie dont l'objectif est d'aider le lecteur à récupérer ou préserver ses capacités langagières.

I. Présentation de l'entreprise CECIAA

1. Historique

La Société CECIAA a été créée en 1990. A cette époque les services de la jeune société étaient entièrement consacrés aux personnes à déficience visuelle (mal voyantes et non voyantes) et à leur intégration dans la vie professionnelle.

Les années 90 sont marquées par l'évolution dynamique des nouvelles technologies et le boom de l'ordinateur personnel. Des technologies avancées qui étaient jusqu'ici réservées à l'usage professionnel deviennent accessibles au grand public. En effet des technologies telles que la synthèse vocale, la reconnaissance de la voix sont à la base de nombreuses solutions informatiques. Cette nouvelle génération de logiciels intégrant ces technologies va permettre aux personnes à déficience visuelle d'accéder à la technologie de l'information.

Parallèlement, le braille fait une évolution considérable. Le code braille s'adapte à l'informatique. Ce dernier va révolutionner le braille avec l'apparition des matériels spécialisés tels que les terminaux ou bloc-notes braille et embosseuses. Des machines d'écriture et lecture braille plus légères sont capables de restituer des informations jusqu'ici inaccessibles aux personnes non-voyantes.

Aux solutions techniques proposées aux non-voyants sont venus s'ajouter des produits d'accessibilité à la lecture (sur écran ou sur document imprimé) pour les mal voyants, basés sur des technologies de traitement de caractères, de grossissement et de la synthèse vocale.

CECIAA occupe bien le marché de la distribution de logiciels et matériels électroniques adaptés aux personnes souffrant de handicap visuel.

La société propose aussi des formations adaptées aux besoins de ses clients, réalisées sur site par des utilisateurs experts. CECIAA participe ainsi à l'élargissement des champs d'activité des personnes non ou mal voyantes en proposant des produits et services qui leur facilitent l'accès à l'information.

En 1999, CECIAA va élargir son domaine d'activité toujours consacré à l'handicap visuel. Une nouvelle politique basée sur le concept de l' « accessibilité » va permettre la création d'un département dédié à ce nouveau challenge.

Le département accessibilité est spécialisée dans le conseil et l'expertise auprès des collectivités locales, des particuliers, des associations pour l'adaptation de l'urbanisme aux personnes handicapées visuelles.

Ses conseils en aménagement favorisent l'apparition des feux sonores, de mobilier urbain adapté mais aussi le défi de rendre accessible les établissements recevant du public et les transports en commun.

2. Organisation

CECIAA emploie aujourd'hui une quarantaine de personnes dont le quart sont des personnes présentant une déficience visuelle. L'entreprise s'implique aussi aux

activités des nouvelles sociétés et acquiert des compétences et une ouverture sans réserve dans le domaine de l'handicap en général.

En effet CECIAA a étendu son activité avec la création des nouveaux départements tels que le handicap moteur, les troubles du langage, l'adaptation de poste ou l'accessibilité urbaine.

3. Le secteur d'activité

L'assistance aux personnes handicapées visuelles par les logiciels et matériels informatiques est un secteur qui a connu un développement considérable avec l'explosion des nouvelles technologies.

Des gammes de matériels et logiciels spécialisés dans l'accessibilité des personnes à déficience visuelle sont de plus en plus abondantes et compétitives.

Aujourd'hui ce secteur apporte plusieurs possibilités combinant essentiellement le braille, le grossissement de caractères, la synthèse vocale et la reconnaissance vocale.

En France, la distribution de matériels spécialisés pour déficients visuels est dominée spécialement par des sociétés telles que :

- EuroBraille, Baum France, TechniBraille, AlphaBraille.. pour le matériel braille et la synthèse vocale ;
- Visiole, Laboratoires Thomas Sinclair, Etex France, Association Valentin HAUY, Handialog.. pour les matériels et services pour le public malvoyant et non-voyant.

Il faut aussi noter le salon AUTONOMIC pour la santé et l'autonomie des personnes handicapées et des personnes âgées.

CECIAA occupe une place prépondérante dans le domaine de la distribution des matériels et logiciels spécialisés mais aussi l'expertise et le conseil en accessibilité.

4. Le département Trouble du langage

Dans le cadre de l'élargissement de son activité, l'entreprise CECIAA a porté une considération particulière à l'évolution du langage chez l'enfant et les troubles liés à son développement.

Le langage est essentiel à notre vie en société. Son développement est un processus compliqué, qui demande plusieurs années. Des premiers balbutiements à la formation de phrases élaborées, l'acquisition du langage est un processus long et complexe. Et pour de nombreux enfants, le chemin est parfois semé d'embûches. Bégaiement, dyslexie, dysphasie sont des troubles fréquents qu'il convient de repérer au plus tôt pour éviter l'échec scolaire.

C'est dans ce cadre que CECIAA a mis en place le département trouble du langage dont la mission est de mettre à la disposition des personnes souffrantes de troubles langagières :

- Des conseils et une expertise sur les choix des solutions techniques existantes. Ces solutions se présentent sous formes d'aides techniques et logiciels centrée sur l'autonomie des personnes dyslexiques, essentiellement faces aux difficultés de lecture et d'écriture.
- Des formations et un suivi pour une meilleure prise en main de ces outils. Directement aux utilisateurs finaux ou aux professionnels tels que les éducateurs spécialisés.

Dans ce cadre, la politique de CECIAA s'est structurée sur trois axes :

L'aide à la lecture

Les solutions proposées sont centrées sur les méthodes informatisées d'aide à la lecture chez les enfants et les adultes. On y trouve principalement des logiciels qui mettent en avant des techniques réduisant les efforts de lecture tels que:

- La vision : elle est soutenue par un système de grossissement de caractères mais aussi par la synthèse vocale.
- La compréhension : facilitée par des modèles de déchiffrage rapide du texte qui permet d'extraire directement le sens de celui-ci.
- Le suivi : Lors de la lecture la synthèse vocale permet d'éviter les sauts de ligne et de se perdre. Certaines solutions proposées par la gamme des produits CECIAA proposent en plus de la synthèse vocale un système de surbrillance permettant de mettre en évidence les parties du texte en cours de lecture.

L'aide à l'écriture

La production de l'écrit est un défi que le département trouble du langage mène depuis sa création. Il propose de soutenir l'action de l'écrit à l'aide de trois techniques :

- La reconnaissance vocale : Elle permet de léguer l'action de l'écrit aux logiciels équipés de cette technologie qui s'avère être d'une utilité incontestable.
- La correction de l'écrit : Elle peut être efficace pour corriger les erreurs d'orthographe et dans certaines limites proposer une présentation syntaxique acceptable.
- La prédiction de mots : Propose des mots bien orthographiés et accélère l'action de l'écrit.

L'aide à la planification et la structuration des idées

Ce type de solution combine à la fois du texte, des graphiques et des concepts pour permettre une vision globale d'une situation quelconque. L'accent est mis sur le développement des habilités métacognitives des sujets. Il s'agit principalement de :

- La capacité de structurer des informations dans la mémoire :
Classer, catégoriser des idées, mettre des informations en réseau, faire des relations, schématiser, dégager l'essentiel des détails...
- La capacité de planifier le travail :
Se donner des objectifs, faire un plan et le suivre, ajuster son rythme et ses stratégies...
- La capacité de gérer le travail et les ressources :
Gérer ses efforts, se donner confiance dans ses choix et se motiver...

II. Les troubles spécifiques du langage

1. Généralité sur les troubles du langage

Les troubles de l'apprentissage du langage concernent 4 à 5% des enfants d'une tranche d'âge et, parmi eux, 1% sont atteints de troubles sévères (INPES). Concrètement, il y'aurait en moyenne un enfant atteint de troubles plus ou moins sévères par classe en France.

Ainsi, on préconise en France l'adoption de certaines mesures visant à améliorer le repérage, le dépistage et la prise en charge des enfants porteurs de troubles du langage. Depuis une trentaine d'années, l'institution scolaire et les professionnels de la santé s'efforcent de prendre la mesure des difficultés rencontrées par certains enfants dans l'accès au langage oral et écrit, et parfois des troubles qui affectent son développement [Marc DeLahaie]. On peut citer par exemple le Plan d'action gouvernemental rendu public en mars 2001 par les Ministres de la santé et de l'Education Nationale. Ce plan a l'ambition d'apporter des réponses graduées, pédagogiques ou médicales, adaptées à l'importance du trouble de chaque enfant, et de développer un travail en équipe des différents professionnels de la santé et de l'enseignement, en complémentarité éducative avec les familles.

1.1. L'évolution du langage chez l'enfant

Chez l'homme, les conditions de vie et probablement les prédispositions innées de son évolution ont, en général, favorisé le développement du langage oral chez l'enfant et cela avec une régularité remarquable.

On constate, à neuf mois environ, que l'enfant comprend des consignes simples ; vers un an, il prononce ses « premiers mots » ; vers dix-huit mois, il organise ses premières combinaisons ; vers trois ans, il utilise les principaux modèles de phrases et les principaux outils grammaticaux (déterminants, prépositions, conjonctions..). Vers l'âge de cinq ans et demi, six ans, l'enfant a atteint un degré suffisant de maîtrise de la langue orale dans ses différentes composantes pour aborder l'apprentissage de la lecture.

Une fois ce niveau atteint, le langage écrit évolue rapidement. Contrairement au langage oral, l'apprentissage de l'écrit passe par l'acquisition des principaux savoirs et savoir-faire de la lecture et de l'écriture. Cet intervalle de temps correspond en France au CP et au CE1.

Si la majorité des enfants prononce leurs premiers mots vers l'âge de 10-12 mois, certains sont plus précoces, d'autres plus tardifs.

Dans certaines limites ces différences interindividuelles ne sont pas pathologiques mais elles peuvent retarder considérablement l'apprentissage du langage oral voire l'écrit.

1.2. Les troubles de l'évolution du langage

L'Organisation Mondiale de la Santé (OMS) retient les troubles spécifiques de la parole et du langage comme des troubles dans lesquels " les modalités normales " d'acquisition du langage sont altérées dès les premiers stades du développement.

Ces troubles ne sont pas directement attribuables à des anomalies neurologiques, des anomalies anatomiques de l'appareil phonatoire, des altérations sensorielles, un retard mental ou des facteurs de l'environnement.

Les troubles spécifiques du développement de la parole et du langage s'accompagnent souvent de problèmes associés tels des difficultés de lecture et d'orthographe, une perturbation des relations interpersonnelles, des troubles émotionnels et des troubles du comportement.

(Réf. OMS Classification Statistique Internationale des maladies et des problèmes de santé connexes - 10ème révision, CIM 10 - citée dans le rapport remis aux Ministres en février 2000 par Monsieur Jean-Charles Ringard)

a. Les troubles de l'apprentissage du langage oral

Dans son livre intitulé "L'évolution du langage chez l'enfant", Marc Delahaie définit par trouble spécifique du développement du langage toute apparition retardée et tout développement ralenti du langage qui ne peuvent être mis en relation avec un déficit sensoriel, avec des troubles moteurs des organes de la parole, avec une déficience mentale, avec des troubles psychopathologiques, avec des carences socio-affectives évidente.

On trouve dans cette catégorie des troubles du langage tels que :

- **Le retard de parole**

C'est une modification de la structure phonétique des mots avec des erreurs dans la combinaison de phonèmes juxtaposés par influence réciproque des différents phonèmes voisins, alors que les mêmes phonèmes sont correctement prononcés lorsqu'ils sont isolés. Ce sont d'abord la persistance à 3-4 ans, d'un parler "bébé", ceci sur des mois ou quelques années. C'est plus tard la non-acquisition de certaines consonnes alors que le langage lui-même (choix et ordre des mots) est normal.

- **Le retard de langage**

Chez certains enfants, le décalage est purement chronologique. Les acquisitions se font dans un ordre normal mais, de manière étalée, avec un décalage inhabituel qui peut aller de quelques mois voire une à deux années.

Le plus souvent, le trouble est à la fois chronologique, qualitatif et quantitatif. On peut observer, outre les éléments témoignant d'un retard de parole, des difficultés à structurer le langage. Les difficultés sont essentiellement expressives :

- Limitation quantitative du stock lexical ;
- Difficultés à acquérir de nouveaux mots ;
- Syntaxe approximative, simplifiée, non-utilisation de certains temps ;
- Confusions touchant certains mots de liaison...

Les difficultés portent sur la compréhension des phrases plus que sur celle des mots isolés.

- **Les dysphasies**

Contrairement aux retards de parole et du langage, les dysphasies constituent des troubles significatifs, sévères et durables de l'évolution du langage oral.

La dysphasie constitue un handicap sévère sur le plan de la communication. Ce constat est valable quelle que soit la forme clinique envisagée. Il s'agit d'un trouble grave et durable du développement de la structure du langage qui se manifeste par trois types de dissociations :

- entre les compétences langagières et les autres compétences intellectuelles
- entre les capacités de compréhension et celles d'expression
- entre le lexique et la syntaxe

En effet, les troubles relatifs à la dysphasie peuvent toucher :

- le plan cognitif (fonctions symboliques non linguistiques, construction d'image mentale, mémoire, structure spatio-temporelle)
- le plan perceptif (mauvaise discrimination auditive pour certaines fréquences conversationnelles)
- le plan psychomoteur (latéralisation, praxies : les praxies sont des fonctions cognitives élaborées qui permettent la gestion de tous les gestes volontaires, finalisés).
- le plan comportemental (attention, hyperactivité, troubles des relations affectives, organisation de la personnalité).

Ces troubles ne sont pas nécessairement tous présents et systématiques.

b. Les troubles de l'apprentissage du langage écrit

Dans son livre intitulé « l'évolution du langage écrit chez l'enfant », Marc Delahaie définit la notion de « troubles spécifiques d'apprentissage du langage écrit » comme étant un concept médical. Cela suppose par conséquent l'existence d'un lien fonctionnel et spécifique entre une difficulté de lecture et un processus pathologique du développement. Ainsi, il préconise à l'instar de la dysphasie, un trouble durable, sévère et d'origine structurelle de l'apprentissage du langage écrit.

Ces troubles sont regroupés sous le terme de dyslexie ou encore dyslexie développementale. La dyslexie représente un trouble de l'apprentissage du langage écrit, une incapacité à reconnaître les mots, et donc à les prononcer.

Cette difficulté qui porte spécifiquement sur l'apprentissage des lettres et des mots n'est pas explicable par le retard mental, ni par un défaut de scolarisation. En effet, un enfant dyslexique n'a pas de difficulté de fonctionnement intellectuel. Lorsque le niveau d'acquisition de la lecture est très inférieur à l'âge réel de l'enfant, on peut parler de dyslexie. La dyslexie gêne tout autant les activités d'acquisition scolaire que les activités extérieures.

La dyslexie s'accompagne généralement de plusieurs caractéristiques comme des difficultés dans le langage parlé et dans le développement du langage, des difficultés d'écriture, ainsi que des difficultés d'attention et de concentration. Les enfants dyslexiques sont des enfants qui peuvent être agités, instables voire une vie scolaire affectée.

La dyslexie survient chez l'enfant en âge d'apprendre à lire et à écrire (généralement vers 7 ans). Chez certains enfants, les symptômes de la dyslexie sont faiblement visibles dès le début, et cela peut se révéler brusquement vers 8-10 ans quand l'enfant se trouve à un stade plus ou moins avancé dans sa période d'apprentissage scolaire.

III. La Dyslexie

1. Définition

La dyslexie est une difficulté durable d'apprentissage de la lecture et d'acquisition de son automatisme, chez des enfants intelligents, normalement scolarisés, indemnes de troubles sensoriels et de troubles psychologiques préexistants (APADYS France).

Un enfant peut être atteint de dyslexie à un degré de gravité plus ou moins sévère. Cette irrégularité rend souvent le dépistage de la dyslexie difficile.

Nous nous limiterons aux troubles qui peuvent gêner les activités de lecture, d'écriture et de compréhension chez l'enfant.

1.1. Les caractéristiques spéciales de la dyslexie

Aujourd'hui 8 à 10 % des élèves scolarisés en France connaissent des troubles plus ou moins graves dus à la dyslexie. Ce chiffre représente environ 1,5 enfant par classe (Source Apedys-France).

Ces difficultés se présentent en général par une incapacité de lecture, de reconnaissance ou de reproduction correcte du langage écrit. On peut citer parmi les caractéristiques de la dyslexie chez l'enfant :

- **La confusion**

Elle peut être :

Visuelle ; par exemple entre les graphèmes de formes proches (p/b, q/p, ou/on...).

Auditive ; Dans la plupart des cas, surtout entre les prononciations phonétiquement proches (p/b, f/v, s/c...).

- **L'omission**

Elle peut concerner la suppression de consonne voire de syllabe entière (réduction syllabique):

Admire -> amire, grain -> gain
Connaitre -> contre

- **L'addition**

Elle est due, le plus souvent, par la ressemblance phonétique des mots :

Ajout de consonnes : paquet -> parquet
Ajout de syllabes : bateau -> balateau

- **L'inversion**

Là aussi, l'erreur concerne les mots ou phrases qui sont phonétiquement proches :
Ortie -> rôtie, bel -> blé

- **La variation du sens de la lecture**

Lecture de droite à gauche au lieu de gauche à droite, cela induit une transposition :
Visa -> avis, venu-> neveu, faucheur -> chauffeur

- **La contamination par persévération**

Tendance à substituer certaines lettres qui se répètent par exemple :

Le « r » persévère : rorure -> dorure

Le « p » persévère : papier -> palier

- **La confusion entre petits mots**

Et -> de, sur -> sous, il -> le

- **La confusion entre mots ressemblants**

Elle peut conduire à :

Une suppression : merveilleux -> meilleur

Une substitution : pomme -> pompe

Un ajout : Ecueil -> écureuil

- **Dysrythmie, dysprosodie**

Lenteur en lecture, débit de lecture variable voire rigide (lecture syllabique).
Erreurs prosodiques (intonation non assortie avec le sens ou la ponctuation).

- **Faible niveau de compréhension**

Cela peut entraîner une précipitation dans le processus de lecture et pousse souvent le jeune à suppléer la compréhension par une intuition sémantique.

1.2. Les méthodes de traitement de la dyslexie

L'origine exacte de ces troubles reste en très grande partie inconnue. Cependant la dyslexie peut se repérer lorsque l'enfant rencontre des difficultés d'apprentissage de la lecture dès ses débuts scolaires:

- Par exemple quand il confond des lettres ou quand il ne parvient pas à lire seul un texte composé de sons dits simples.
- Si l'enfant est excessivement lent dans des travaux écrits, la possibilité d'une dyslexie doit également être envisagée.

Il est important de repérer ce trouble au plus tôt, car un enfant qui est en proie avec des difficultés, qui a l'impression de faire des efforts mais n'en retire aucun résultat, peut avoir tendance à se replier sur lui-même ou peut devenir très tendu et agressif. Cependant, il faut être attentif, car tout enfant présentant des difficultés d'apprentissage de la lecture ne souffre pas systématiquement de dyslexie.

Cette particularité non systématique qui caractérise la dyslexie est à l'image de nombreux traitements qui sont basés sur des objectifs diverses.

Il existe plusieurs types de traitements de la dyslexie. Certains préconisent la rééducation sensorielle comme celle de l'audition, de l'oral ou de la vision. D'autres envisagent aussi une rééducation motrice. Si certaines de ces méthodes s'avèrent être bénéfiques, il reste difficile de vérifier si cela est dû à une amélioration sensorielle ou plutôt à une acquisition d'automatismes résultant des entraînements qui font partis de leur programme.

Les objectifs de ces méthodes ne sont pas communs. Des démarches basées sur la rééducation cherchent, le plus souvent, à corriger les erreurs visibles telles que les erreurs de lecture ou écriture (voir paragraphe précédent). D'autres qui s'appuient sur des fondements médicaux ou psychopathologiques encouragent le suivi de traitements thérapeutiques de plus en plus rares.

La diversité parfois contradictoire des stratégies de traitement proposées n'exclue pas cependant la pauvreté voire l'absence d'évaluation rigoureuse de ces traitements, permettant d'en établir l'efficacité et la validité, sur un mode autre que celui de l'empirisme [Marc DeLahaie].

2. L'apport de l'informatique

Le développement de l'informatique favorise indiscutablement l'amélioration de la vie des personnes handicapées.

L'ordinateur est un outil précieux pour faciliter l'expression et la communication des personnes atteintes d'un handicap. L'usage de l'ordinateur avec les périphériques et des logiciels adaptés facilite énormément la scolarisation de ces personnes et peut dans un grand nombre de cas leur apporter une autonomie très importante.

Pour certains types de handicap il est aujourd'hui facile de trouver une solution informatique pour réduire voire compenser les désavantages liés. Grâce à la prise de conscience des éditeurs sur les difficultés des utilisateurs handicapés, des nouveaux concepts ont vu le jour et ont pour but de façonner les outils physiques ou virtuels pour les rendre plus accessibles.

On peut citer à, ce sujet, quelques exemples de technologies de l'informatique les plus couramment utilisées dans ce domaine.

Handicap moteur :

- Les outils de saisie vocale : pour les personnes ne pouvant pas se servir du clavier.
- Les claviers virtuels : manipulable à l'aide d'un dispositif de pointage (souris, joystick, systèmes de commande par la respiration, le regard...).
- Les filtres de clavier : compensent en partie les mouvements erratiques, les tremblements...
- Les systèmes de communication alternative ou assistée (interfaces parlantes...)

Handicap visuel :

- Logiciels de grossissement de caractères.
- Lecteurs d'écran : convertissent l'information affichée à l'écran en parole synthétique ou en information tactile sur un afficheur braille.

Handicap auditif :

- La reconnaissance vocale : transcrit à l'écrit les paroles de plusieurs interlocuteurs. L'information est consultable en temps réel ou en différé.
- Filtre de son : pour rendre audible un enregistrement bruité.
- Système de boucle magnétique
- Interfaces de communication alternative (espaces de dialogue et d'information accessibles par Internet : exemple Web Sourd, pour les personnes sourdes qui signent...)

Déficiência cognitive ou linguistique :

- Pour ce type de trouble, des solutions informatiques combinant des technologies comme celles illustrées précédemment sont de plus en plus efficaces : les solutions de dictée vocale permettent aux personnes présentant des troubles du langage écrit de produire des textes sans être pénalisés ou gênés par l'orthographe des mots. Quant à l'emploi des synthèses vocales, elles permettent aux personnes souffrant de troubles du langage oral et/ou écrit d'accéder plus aisément et plus rapidement aux informations textuelles contenues dans un document papier et éditable.

IV. Les objectifs du projet DysBloc

L'objectif central du projet DysBloc est de réfléchir à un outil d'adaptation de document texte pour les personnes dyslexiques. En effet la présentation d'un texte brut (non formaté) est différente de celle d'un texte mise en page. Cette dernière fonctionnalité joue un rôle non négligeable au niveau de l'accessibilité de l'information écrite. Nous allons voir dans cette partie comment les différents styles de mises en forme peuvent faciliter l'accès à l'information et surtout attirer l'attention du lecteur sur les particularités syntaxiques, phonétiques, grammaticales... du texte en cours de lecture.

1. Les adaptations des documents textes

1.1. Idée générale

Les fonctionnalités de DysBloc ont pour objet de faciliter la lecture des documents texte et par conséquent, sa compréhension. Ces fonctionnalités s'appuient sur différents styles de mise en page proposés tout en respectant les normes d'accessibilité facilitant, ainsi l'accès aux documents texte numérisés ou imprimés. Il s'agit de mettre en évidence les parties sensibles du texte pour attirer l'attention du lecteur dyslexique lors de la lecture. Le but est d'éviter les erreurs de déchiffrement que nous allons voir dans le prochain paragraphe et d'entraîner le lecteur à ne pas les répéter.

Il s'agit là, non pas d'une solution aux troubles liés à la dyslexie, mais d'un outil visant à soutenir les activités de lecture et d'écriture.

Nous allons voir dans les paragraphes suivants, les styles de mise en page choisis ainsi que les recommandations sur leurs utilisations.


1.2. Les recommandations sur l'accessibilité


Une personne peut être affectée par la dyslexie de façon et à de niveaux différents. Les personnes dyslexiques ont beaucoup du mal à lire un texte chargé. En général, les recommandations sur d'accessibilité visent à alléger la présentation de l'information et à assurer une restitution claire et nette de l'écrit.

Sont présentées ici des techniques de mise en page proposées par l'Université de Manchester (intranet.man.ac.uk/accessibility). Ces techniques s'inspirent notamment des recommandations du WAI (Web Accessibility Initiative/w3.org).

- **Largeur des colonnes**

Certaines personnes dyslexiques préfèrent lire les textes présentées en colonnes. Il est conseillé d'utiliser au maximum deux colonnes. Cette méthode facilite la distinction des différents paragraphes et améliore ainsi l'évolution de la lecture.

Texte en une colonne	
<p>Au TP, mes copains ont toujours fini de lire l'énoncé avant moi. Et ils n'arrêtent pas de trépigner à côté en me demandant si j'ai fini. J'essaye d'accélérer ma lecture, ça me stresse. Je lis en diagonale. Une fois fini je n'ai rien compris car j'ai lu des mots mais je n'ai pas essayé de comprendre ce qu'il disait. De toute façon dans l'urgence je suis incapable de lire.</p> <p>Extraits de témoignages d'élèves en difficulté : www.grenoble-universites.fr</p>	

Texte en deux colonnes			
<table> <tr> <td> <p>Au TP, mes copains ont toujours fini de lire l'énoncé avant moi. Et ils n'arrêtent pas de trépigner à côté en me demandant si j'ai fini. J'essaye d'accélérer ma lecture, ça me stresse. Je lis en diagonale.</p> </td> <td> <p>Une fois fini je n'ai rien compris car j'ai lu des mots mais je n'ai pas essayé de comprendre ce qu'il disait. De toute façon dans l'urgence je suis incapable de lire.</p> </td> </tr> </table>	<p>Au TP, mes copains ont toujours fini de lire l'énoncé avant moi. Et ils n'arrêtent pas de trépigner à côté en me demandant si j'ai fini. J'essaye d'accélérer ma lecture, ça me stresse. Je lis en diagonale.</p>	<p>Une fois fini je n'ai rien compris car j'ai lu des mots mais je n'ai pas essayé de comprendre ce qu'il disait. De toute façon dans l'urgence je suis incapable de lire.</p>	
<p>Au TP, mes copains ont toujours fini de lire l'énoncé avant moi. Et ils n'arrêtent pas de trépigner à côté en me demandant si j'ai fini. J'essaye d'accélérer ma lecture, ça me stresse. Je lis en diagonale.</p>	<p>Une fois fini je n'ai rien compris car j'ai lu des mots mais je n'ai pas essayé de comprendre ce qu'il disait. De toute façon dans l'urgence je suis incapable de lire.</p>		

- **Texte non justifié ou justifié à gauche**

Le fait de justifier le texte assure une fin de ligne unique pour tout le texte. Cependant, cette technique ajoute au contenu des espacements entre les mots qui peuvent être de longueurs variables. Les personnes dyslexiques peuvent éprouver des difficultés à lire un texte justifié. Il est vivement conseillé de ne pas justifier le texte

ou alors d'appliquer un alignement à gauche. Ce qui assure un espacement de longueur fixe entre les mots.

- Espace entre les mots : il doit être de taille unique
- L'interligne : un et demi ou un double interligne

• Couleur de fond

Un contraste net entre le fond et la couleur de police permet aux personnes dyslexiques de mieux distinguer le contenu et éviter l'éblouissement. Le mariage de certaines couleurs qui, séparément sont nettement différentes, peut induire un effet aveuglant de même que l'usage d'une image de fond.

Certaines personnes dyslexiques trouvent que l'usage d'un fond blanc (plâtre) est trop lumineux et inconfortable lors de la lecture. Dans ce cas de figure il est préférable d'utiliser plutôt un blanc cassé ou une autre couleur éclaircie.

• Type de Police

Les fontes avec serif, étant composées de caractères avec empattements (exemple : Times New Roman) sont difficilement lisibles pour la plupart des personnes dyslexiques. Cela rajoute une décoration parfois trompeuse lors de la lecture. Les fontes sans serif sont beaucoup plus nettes et faciles à lire. Il est donc conseillé d'utiliser des polices du type Arial, Verdana, Trébuchet MS, Courier...

Type de police		Exemples	
Serif	TimeNews Roman	Pouvez-vous lire cette phrase ?	✗
	Georgia	Pouvez-vous lire cette phrase ?	✗
	Bodoni	Pouvez-vous lire cette phrase ?	✗
Sans serif	Arial	Pouvez-vous lire cette phrase ?	✓
	Verdana	Pouvez-vous lire cette phrase ?	✓
	Trébuchet MS	Pouvez-vous lire cette phrase ?	✓

• Taille de la police

L'usage des caractères de petite taille pose des problèmes aux personnes dyslexiques comme aux personnes non dyslexiques. La taille du texte doit être suffisamment grande pour permettre une bonne lisibilité. 12 est une taille minimale. En général, il est conseillé d'utiliser la taille 16.

Taille du texte	Exemple (fonte sans serif : Arial)	
8	Pouvez-vous lire cette phrase ?	✗
10	Pouvez-vous lire cette phrase ?	✗
12	Pouvez-vous lire cette phrase ?	✗
16	Pouvez-vous lire cette phrase ?	✓

• **Style italique et style gras**

Pour faciliter la lisibilité, évitez l'écriture en style italique. Par contre le style gras est très apprécié et permet de mettre en évidence certaines parties du texte. Une application du style gras en plus des guillemets permet de distinguer nettement les parties du texte à mettre en évidence du reste.

Texte en italique	✘	Texte en gras	✔
La question posée est : <i>"Pouvez-vous lire cette phrase ?"</i>		La question posée est : "Pouvez-vous lire cette phrase ?"	

• **Puces et numéros**

Privilégiez l'utilisation des numéros plutôt que les puces beaucoup plus difficiles à suivre.

Numéros	✔	Puces	✘
Je ne sais jamais comment mettre les accents sur le e alors j'ai deux solutions : 2. Je n'en mets pas, ou alors 3. Je les mets verticalement		Je ne sais jamais comment mettre les accents sur le e alors j'ai deux solutions : • Je n'en mets pas, ou alors • Je les mets verticalement	

• **Images et illustrations**

L'utilisation d'images ou d'illustrations peut s'avérer intéressante pour les personnes dyslexiques (notamment les élèves) car elles permettent d'illustrer une idée ou de donner une information supplémentaire. Dans certains cas, l'emploi d'images peut se révéler plus judicieux que le texte. Cependant il faut privilégier des images appropriées et si possible sans trop d'éléments superflus.

• **La longueur des paragraphes et des phrases**

La longueur des phrases et des paragraphes doit être aussi courte que possible. L'utilisation des paragraphes est beaucoup plus recommandée, ils aident à isoler chaque idée du texte dans un paragraphe composé de phrases courtes et concises. La technique de mise en page des journaux est un bon exemple, ils résument une histoire avec des phrases simples et courtes.

• **Présentation générale du contenu**

La mise en page du document numérisé doit assurer une lisibilité claire de son contenu. Il convient d'utiliser une mise en page uniforme tout au long du document. Les critères de mise en page doivent servir à faciliter le déchiffrement du contenu et surtout le suivi lors de la lecture.

- Pour les documents imprimables, éviter le plus possible les notes de bas de pages et les notes de renvoi à d'autres pages/documents.
- Certaines personnes dyslexiques utilisent des lecteurs de documents pour lire notamment le contenu textuel des pages numérisées. Assurez-

vous que la mise en forme des pages respecte les normes d'accessibilité permettant l'utilisation de lecteur de documents et navigateur vocal.

- Fournir un glossaire pour aider à comprendre les termes spéciaux utilisés dans le document.
- Expliquer les abréviations au moins lors de leur première apparition.
- Fournir des conseils pour permettre aux utilisateurs de paramétrer leurs matériels et logiciels de manière optimale.

2. Les fonctionnalités de DysBloc

L'idée est de proposer des fonctionnalités de mise en évidence ou de mise en page paramétrables et donc adaptables aux besoins spécifiques de chaque utilisateur dyslexique. On peut ainsi envisager qu'un enseignant ajuste les paramètres du logiciel afin que ceux-ci conviennent aux troubles des élèves dyslexiques. Pour cela, nous avons voulu définir les différentes méthodes qui peuvent être utilisées afin de pallier, dans une certaine mesure les difficultés de lecture et de compréhension présentées par ces utilisateurs.

Cette partie du projet est importante car avant d'implémenter pour la première fois ces fonctions, il convient de faire un certain nombre d'études sur leur efficacité.

Nous avons eu pour cela l'idée de travailler en collaboration avec des enseignants spécialisés afin de valider et tester cet outil.

Pour des raisons techniques, cette collaboration n'a pas pu être mise en place dans le temps imparti. Nous nous sommes donc basés sur les informations fournies par divers enseignants spécialisés et sur les demandes exprimées par ceux-ci lorsque le projet leur a été présenté. Nous leur présenterons dès que possible notre maquette afin de valider ou d'invalider les différentes options retenues lors du développement.

Pour revenir aux fonctionnalités, celles-ci se classent en deux groupes selon leur niveau d'intervention.

2.1. Les fonctionnalités d'aide au déchiffrement

Il s'agit des fonctions qui ont pour but de faciliter l'accès à l'information écrite et éviter les erreurs de déchiffrement. On distingue :

a. Les fonctions de paramétrage de l'accessibilité

- arrière-plan
- police
- espacement de mots
- interligne
- Etc.

b. Les fonctions de mise en évidence

- Mise en évidence des graphèmes complexes (pl/gr/pr...)
- Mise en évidence des syllabes complexes (ou/oi/eu/..., oui/oei/uei/...)
- Mise en évidence de la ponctuation
- Mise en évidence des lettres muettes
- Etc.

Il est à noter qu'étant donné les variations interindividuelles des personnes dyslexiques, nous devons prévoir la possibilité de créer des nouvelles fonctions, ou de modifier pour adapter celles qui existent ou supprimer celles qui ne sont plus utilisées.

Les styles de mise en évidence aussi ne doivent pas être fixes. On doit laisser la possibilité aux utilisateurs de choisir les couleurs, le type de police et la taille.

2.2. Les fonctionnalités d'aide à la compréhension

On trouve principalement la fonction d'illustration des mots par des images mais aussi de structuration des groupes de mots par idée (exemple le texte en colonne).

V. Les solutions techniques

Plusieurs solutions sont possibles pour parvenir à nos besoins. Cependant, pour des raisons de maintenance, nous privilégions trois critères de sélection :

- **Portabilité** : La solution qu'on aura choisie doit satisfaire la contrainte de portabilité sur toutes les plateformes et outils de traitement de texte en priorité Word et OpenOffice.
- **Facilité de maintenance** : En effet, les techniques choisies doivent être ouvertes à toutes modifications visant à les faire évoluer sans avoir besoin à recoder les fonctionnalités.
- **Libre d'utilisation.**

1. Les techniques de mise en page des fichiers textes

Pour pouvoir automatiser les fonctionnalités de mise en évidence, l'application doit être capable d'analyser le texte avant d'appliquer les règles de mise en pages qui conviennent. L'application des styles de mise en page se fera indépendamment de l'utilisateur. Pour ce faire, nous présentons brièvement trois moyens de parvenir à manipuler des fichiers textes formatés.

1.1. Les scripts et macros

Une macro-commande est un script ou un ensemble de scripts qui permet d'enregistrer une série d'actions qu'on effectue souvent, de façon à pouvoir les reproduire très rapidement après coup.

Inconvénients

- Différence de codage : Par exemple les macros sous Word sont écrites en VisualBasic ce qui n'est pas le cas sous OpenOffice.
- Lourdeur dans la mise en pratique : Les macros sont difficiles à mettre en œuvre et connaissent des limites pour les manipulations avancées.

1.2. Utilisations des API adaptées

Les interfaces de programmation nous permettent d'exploiter directement les objets types utilisés dans une application et ce dans un langage de programmation donné. Etant donné le choix fait sur le langage de programmation qui est Java, nous avons donc besoin d'API écrite dans ce langage.

Nous présenterons les raisons de ce choix (Java) dans les prochains paragraphes.

Cette solution consiste à mettre en place les outils appropriés pour la manipulation des fichiers dans leur structure native. C'est sans doute la solution la plus correcte mais nous allons voir qu'elle n'est pas facile à mettre en œuvre.

Inconvénients

- Gestion des fichiers Word : Le projet POI est la seule API Java qui donne accès à la manipulation des fichiers au format de type Microsoft OLE 2 Compound document. En résumé, il est possible de lire et écrire sur des fichiers Excel ou

Word en utilisant Java. Le projet POI est composé de plusieurs composants dont HWPF pour la manipulation des fichiers Word. Malheureusement, cette partie du projet est restée à l'état de prématuré. Cette API n'est pas tout à fait accomplie pour pouvoir satisfaire des manipulations avancées.

- Programmation parallèle : Opter pour cette solution, implique la réalisation d'un programme pour chaque type de fichier étant donnée la différence de formats.

1.3. Mise en page avec XML/XSL

A l'image des pages Web, plusieurs types de fichier utilisent la technologie XML pour proposer une structure de codage lisible et ouverte. Cette technique est aussi utilisée sous la forme de filtre XSL pour générer des fichiers tels que PDF, RTF...

C'est le cas aussi pour tous les fichiers supportés par la suite bureautique libre d'utilisation OpenOffice.org et de Microsoft Word depuis sa version de 2003.

Cette solution est très intéressante dans la mesure où il suffirait de pouvoir manipuler des fichiers XML pour accéder à la modification des fichiers finaux. Cela veut dire que le choix du langage de développement est indépendant du type des fichiers. En effet, le XML est devenu un format standard et normalisé. Il existe plusieurs interfaces de programmation pour la plupart des langages de programmation, conçus pour la manipulation des fichiers XML.

Nous allons voir, dans le paragraphe suivant :

- Un rappel sur le format XML
- La mise en page avec le XSL
- Les structures XML des fichiers textes d'OpenOffice et de WorldML

2. Le XML et ses applications

2.1. Généralités sur le XML

Le XML pour eXtensible Markup Language est un langage de balises comme le HTML mais il est extensible, évolutif. En XML, les balises ne sont pas prédéfinies. C'est au créateur de définir ses balises et d'en donner un sens.

A l'image du HTML, le XML est un langage de balise qui a été créé pour structurer des informations. Mais le XML est de loin plus abstrait et donc plus complexe que le HTML. Ce dernier a été conçu pour afficher de l'information, le XML sert à structurer l'information comme le fait par exemple une base de données. Mais le XML, en lui-même ne fait rien.

Voici un exemple de fichier XML :

```

<?xml version="1.0" encoding="iso-8859-1"?>
<cv>
  <nom>
    Deluxe
  </nom>
  <prénom>
    Alpha
  </prénom>
  <age>
    25
  </age>
  <formation>
    <baccalaureat>
      <serie>S
      </serie>
      <année>2000
      </année>
      <mention>Assez-bien
      </mention>
    </baccalaureat>
    <diplomeuniversitaire>
      <diplôme1>
        DEUG Economie
      </diplôme1>
      <diplôme2>
        Master Gestion Financiere
      </diplôme2>
      <diplôme3 />
      </diplôme3>
    </diplomeuniversitaire>
  </formation>
  <experienceprofessionnelle>
    <stages />
    </stages>
    <emplois />
    </emplois>
  </experienceprofessionnelle>
</cv>

```

cv.xml

Sur un navigateur, cela donne le résultat suivant :

```

<?xml version="1.0" encoding="iso-8859-1" ?>
- <cv>
  <nom>Deluxe</nom>
  <prénom>Alpha</prénom>
  <age>25</age>
- <formation>
  - <baccalaureat>
    <serie>S</serie>
    <année>2000</année>
    <mention>Assez-bien</mention>
  </baccalaureat>
  - <diplomeuniversitaire>
    <diplôme1>DEUG Economie</diplôme1>
    <diplôme2>Master Gestion Financiere</diplôme2>
    <diplôme3 />
  </diplôme3>
  </diplomeuniversitaire>
  </formation>
- <experienceprofessionnelle>
  <stages />
  </stages>
  <emplois />
  </emplois>
</experienceprofessionnelle>
</cv>

```

Nous remarquons que sur le plan esthétique il n'y a presque aucun changement. Le XML n'est que de l'information encodée entre des balises. En effet les navigateurs ne

« comprennent » pas le sens des balises XML et notamment comment afficher l'information. Il faudrait d'autres éléments, comme par exemple un fichier XSL, pour que le navigateur puisse « comprendre » les balises XML et afficher les informations sous une forme plus conviviale.

Le XML est un métalangage soit un langage pour écrire d'autres langages. Mais le XML est une véritable révolution dans le panorama des langages de publication sur Internet. Il apparaît comme incontournable car il est déjà à la base de toute une série de nouveaux langages qui sont ou qui seront utilisés dans la conception des pages Internet comme le XHTML, le successeur désigné du HTML, le WML pour le Wap des téléphones mobiles, le MathML pour les mathématiques et bien d'autres encore.

2.2. Définition des balises XML

Nous avons vu plus haut que le XML en lui-même n'est « rien ». Il est cependant tout à fait utilisable seul à condition de savoir ce que renferme chaque balise. On peut, de cette façon utiliser le XML pour par exemple stocker des données.

Mais la façon la plus élégante d'utiliser le XML est de définir le document XML et donc le sens des balises qui le composent.

a. Le DTD et le XSD

Le DTD (Document Type Declaration) ou encore XSD (eXtended Schema Document) est l'ensemble des règles et des propriétés que doit suivre le document XML. Ces règles définissent généralement le nom et le contenu de chaque balise et le contexte dans lequel elles doivent exister. Cette formalisation des éléments est particulièrement utile lorsqu'on utilise de façon récurrente des balises dans un document XML.

Le XSD est une évolution naturelle des DTD. Contrairement au DTD, le XSD est un langage XML, il permet de décrire formellement les types, les espaces de noms, les contraintes sur les données, etc.... il est de ce fait plus adapté au développement d'application.

L'étude détaillée du DTD et du XSD dépasse de loin le cadre de ce projet mais leur existence prouve la rigueur du XML mais en même temps sa souplesse. Ce qu'il faut retenir est que le DTD ou le XSD servent à vérifier la conformité puis la validité d'un document XML et ce par rapport à un schéma prédéfini.

2.3. Affichage des données XML

Après avoir vu comment on définit un document XML, nous allons à présent voir comment on peut afficher les données d'un document XML.

Pour afficher les balises XML, on peut faire appel à ce qu'on appelle les feuilles de style. Il existe deux types de feuilles de style : le CSS et le XSL

A l'image du DTD et du XSD le XSL est l'évolution du CSS car il est lui-même un document XML ce qui n'est pas le cas du CSS.

a. XML et CSS

Devenu classique dans le paysage HTML, le CSS permet d'afficher et donc d'appliquer une mise en page pour chaque donnée encapsulée dans une balise. A chaque balise dans le fichier XML ou HTML, on va définir un élément de style que le navigateur pourra alors afficher.

Exemple de XML + CSS

<pre> <?xml version="1.0" encoding="iso-8859-1" ?> - <cv> - <identite> <nom>Deluxe</nom> <prénom>Alpha</prénom> <age>25</age> </identite> - <formation> - <baccalaureat> <serie>S</serie> <année>2000</année> <mention>Assez-bien</mention> </baccalaureat> </formation> </cv> </pre>	<pre> 1 /*CSS Document*/ 2 <style type="text/css"> 3 cv.identite{ 4 nom, prenom, age{ 5 display: block; width:250px; 6 font-size: 16pt; font-family: arial; 7 font-weight: bold; background-color: yellow; 8 color: blue; padding-left: 10px; 9 } 10 serie{ 11 display: block; font-size: 12pt; color: blue; 12 padding-left: 10px; 13 } 14 année, mention{ 15 display: block; font-size: 12pt; color: red; 16 font-weight: bold; padding-left: 10px; 17 } 18 </style> </pre>
---	---

exemple2.xml

exemple2.css

Après avoir déclaré la feuille de style « exemple2.css » sur le fichier XML en insérant le code suivant :

```
<? xml-stylesheet href="exemple2.css" type=text/css" ?>
```

Nous obtenons l'affichage suivant sur le navigateur Firefox :



Mais il existe encore un autre moyen, plus performant et aux possibilités plus étendues : afficher du XML avec le XSL soit le langage de feuilles de style extensible. La version XML du CSS.

b. XML et XSL

Etant donné que le XML n'utilise pas de balises prédéfinies (car on peut inventer ses propres balises), le navigateur ne « comprend » pas les balises du XML et ne sait pas comment afficher les données du document XML.

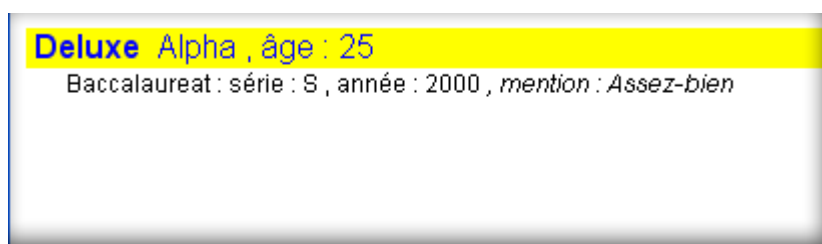
Ainsi il est nécessaire d'avoir un mécanisme pour décrire comment le document pourrait être affiché. Un de ces mécanismes est les feuilles de style classique du HTML, le CSS. Mais le XSL (eXtensible Stylecheet Language) est de loin un langage de feuille de styles plus adapté au XML et donc plus performant. En résumé, le XSL est un langage qui transforme le XML en HTML. Mais il fait bien plus !

En reprenant le même exemple que dans le cas du CSS et en remplaçant la feuille de style exemple2.css par une feuille de style de type XSL suivante :

```
<?xml version="1.0" encoding="iso-8859-1" ?>
- <xsl:stylesheet version="1.0" xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
- <xsl:template match="/">
- <html>
- <body style="font-family: Arial; font-size:12pt;">
- <xsl:for-each select="cv/identite">
- <div style="background-color: yellow; color: blue;">
- <span style="font-weight: bold; color: blue; padding:4px;">
  <xsl:value-of select="nom" />
</span>
  <xsl:value-of select="prenom" />
- <span>
  , âge :
  <xsl:value-of select="age" />
</span>
</div>
</xsl:for-each>
- <xsl:for-each select="cv/formation/baccalaureat">
- <div style="margin-left:20px; font-size: 12px;">
- <span>
  Baccalaureat : série :
  <xsl:value-of select="serie" />
</span>
- <span>
  , année :
  <xsl:value-of select="annee" />
</span>
- <span style="font-style:italic">
  , mention :
  <xsl:value-of select="mention" />
</span>
</div>
</xsl:for-each>
</body>
</html>
</xsl:template>
</xsl:stylesheet>
```

exemple2.xsl

On obtient, sur le navigateur Firefox, l'affichage suivant :



Le XSL ne permet pas uniquement l'affichage de données XML. Il permet entre autre de filtrer les éléments XML en fonction de certains critères, de rajouter du contenu, etc.

Le XML comporte en fait trois langages tous descendants du XML :

1. Le **XSLT** qui est le langage qui transforme un document XML en un autre format (exemple le Html).
2. Le **Xpath** qui permet de définir et d'adresser des parties du document XML.
3. Le **XML Formater** pour « formater » du XML de manière à ce qu'il puisse être rendu sur d'autres plateformes (téléphonie mobile, synthèse vocale, braille...).

Le but de ce rapport n'est pas d'étudier le XSL de manière approfondie mais de montrer les possibilités sur les fonctionnalités de mise en page du XSL.

En effet, il faut noter que cette technique de présentation de l'information très performante est utilisée sur la plupart des formats de document. C'est le cas d'OpenOffice et MS Office qui nous intéressent plus particulièrement ici.

OpenOffice utilise depuis son lancement un format XML public pour sauvegarder les documents. Depuis la version 2003, Word utilise également un format XML, dont les spécifications ont été presque entièrement rendues publiques.

3. Le standard OpenDocument

La suite bureautique OpenOffice.org est actuellement un des logiciels libres les plus utilisés dans le monde, et dont le développement est le plus actif : elle est devenue en effet une alternative crédible aux offres des acteurs dominants du marché. Un de ses avantages historiques mis régulièrement en avant, mais souvent mal connu ou compris, est son support natif du XML.

Nous allons étudier dans cette partie l'intérêt du XML pour le poste de travail, l'utilisation faite du XML dans OpenOffice.org, les opportunités qu'elle offre aux développeurs d'applications sur la manipulations directes des fichiers, et les différences entre l'utilisation du XML par Microsoft Office 2003 et celle d'OpenOffice.org.

OpenOffice.org produit nativement du XML et cette particularité est totalement transparente pour l'utilisateur. En effet, le format de stockage d'OpenOffice.org est un format XML tout à fait classique, qui permet de faire ainsi cohabiter simplicité d'utilisation, formats bureautiques (hétérogènes, souvent peu structurés) et fortes possibilités d'échanges. Le format XML d'OpenOffice.org 1.0 a d'ailleurs été pris

comme base de travail par l'OASIS¹ pour créer le standard "OpenDocument", qui a été adopté comme standard le 23 mai 2005. OpenOffice.org 2.0, supporte bien entendu le format OpenDocument de façon native et sert, en quelque sorte, d'implémentation de référence.

En plus de son format initial, OpenOffice.org est capable d'éditer tout type de format XML, tant qu'on lui fournit la DTD (définition du format) et les feuilles de transformations correspondantes. La distribution d'OpenOffice.org inclut d'ailleurs des définitions pour éditer nativement du XHTML, des fichiers Word 2003 XML et du DocBook. Ces possibilités lui confèrent un large potentiel :

- édition directe de documents XML,
- intégration du format OpenDocument à une chaîne de traitement XML complète
- génération de documents
- manipulation de documents dans des applications métiers
- génération de modèles de documents, etc.

3.1. Structure d'un fichier OpenOffice.org

OpenOffice.org reste l'implémentation de référence d'OpenDocument. Un fichier OpenOffice.org est en fait une ressource JAR (une archive ZIP au sens Java) contenant divers éléments dont les fichiers texte XML. Le format est donc totalement ouvert et pérenne puisque un simple éditeur de texte permet de consulter les informations des fichiers XML, l'algorithme de compression zip étant lui même libre et ouvert⁵ Lorsque l'on décompresse un fichier OpenOffice.org (à l'aide d'un simple décompresseur ZIP comme WinZip sous Windows ou unzip sous UNIX), on obtient entre autres les fichiers suivants :

- META-INF/manifest.xml : liste et description des fichiers contenu dans le JAR
- meta.xml : méta-données du document
- mimetype : type MIME du fichier
- settings.xml : paramètres d'édition, impression et visualisation d'un document
- **styles.xml** : description des styles du document
- **content.xml** : Contenu sémantique du document
- Pictures/ : répertoire contenant les images insérées dans le document
- Basic/ : répertoire contenant les macros contenues dans le document
- Thumbnails/ : répertoire contenant les vignettes du document pour l'intégration au bureau
- Configurations2/ : répertoire stockant les fichiers de paramétrages propre à l'interface graphique d'OpenOffice.org (comme certaines barres d'outils par exemple).

Ce qui nous intéresse c'est surtout les deux fichiers XML content.xml et styles.xml.

On constate ici encore la modularité du format OpenOffice.org :

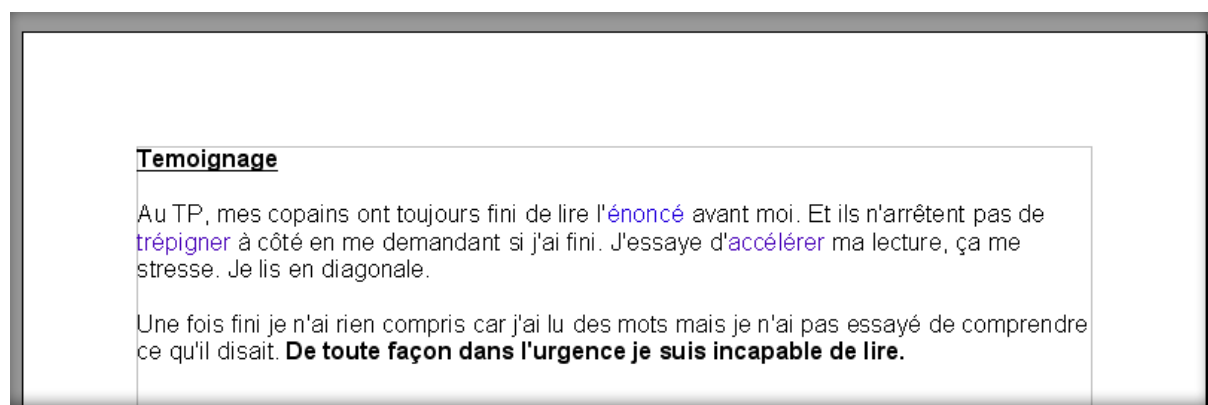
- un document XML pour les métadonnées
- un pour les styles (styles.xml)
- un pour le contenu du document (content.xml)

¹ Organization for the Advancement of Structured Information Standards

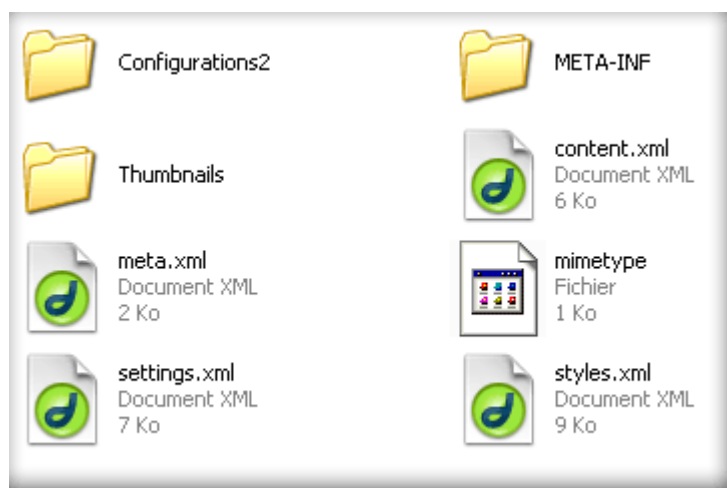
- un pour les paramètres d'impression, d'édition et d'affichage.
- Les images et les macros sont tout simplement stockées dans leur répertoire dédié.

Cette organisation facilite vraiment l'utilisation du format par d'autres logiciels. Ainsi, il est possible de ne travailler que sur le contenu d'un document sans être perturbé par les styles (ce qui est très utile lors de la génération de document ou de l'extraction de données, par exemple). De même, on peut appliquer aux données un autre mise en page ou tout simplement modifier les données aisément. Ce qui est le but de notre projet.

Prenons l'exemple d'un fichier OpenOffice « *temoignage.odt* » dont le contenu est :



En renommant ce fichier en « *temoignage.zip* » nous obtenons le contenu suivant après décompression du fichier zip :



3.2. Stockage du contenu

Voyons de près le fichier **content.xml** :

```
<?xml version="1.0" encoding="UTF-8" ?>
<office:document-content
  xmlns:office="urn:oasis:names:tc:opendocument:xmlns:office:1.0"
  xmlns:style="urn:oasis:names:tc:opendocument:xmlns:style:1.0"
  xmlns:text="urn:oasis:names:tc:opendocument:xmlns:text:1.0"
  xmlns:table="urn:oasis:names:tc:opendocument:xmlns:table:1.0"
  xmlns:draw="urn:oasis:names:tc:opendocument:xmlns:drawing:1.0"
  xmlns:fo="urn:oasis:names:tc:opendocument:xmlns:xsl-fo-compatible:1.0"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:meta="urn:oasis:names:tc:opendocument:xmlns:meta:1.0"
  xmlns:number="urn:oasis:names:tc:opendocument:xmlns:datastyle:1.0"
  xmlns:svg="urn:oasis:names:tc:opendocument:xmlns:svg-compatible:1.0"
  xmlns:chart="urn:oasis:names:tc:opendocument:xmlns:chart:1.0"
  xmlns:dr3d="urn:oasis:names:tc:opendocument:xmlns:dr3d:1.0"
  xmlns:math="http://www.w3.org/1998/Math/MathML"
  xmlns:form="urn:oasis:names:tc:opendocument:xmlns:form:1.0"
  xmlns:script="urn:oasis:names:tc:opendocument:xmlns:script:1.0"
  xmlns:ooo="http://openoffice.org/2004/office"
  xmlns:ooow="http://openoffice.org/2004/writer"
  xmlns:oooc="http://openoffice.org/2004/calc"
  xmlns:dom="http://www.w3.org/2001/xml-events"
  xmlns:xforms="http://www.w3.org/2002/xforms"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" office:version="1.0">
<office:scripts />
  <office:font-face-decls>
    <style:font-face style:name="Tahoma1" svg:font-family="Tahoma" />
    <style:font-face style:name="Arial Unicode MS" svg:font-family="Arial Unicode MS"
      style:font-pitch="variable" />
    <style:font-face style:name="MS Mincho" svg:font-family="MS Mincho" style:font-p
      itch="variable" />
    <style:font-face style:name="Tahoma" svg:font-family="Tahoma" style:font-p
      itch="variable" />
    <style:font-face style:name="Times New Roman" svg:font-family="Times New Roman"
      style:font-family-generic="roman" style:font-pitch="variable" />
    <style:font-face style:name="Arial" svg:font-family="Arial" style:font-family-
      generic="swiss" style:font-pitch="variable" />
  </office:font-face-decls>
  <office:automatic-styles>
    <style:style style:name="T1" style:family="text" style:parent-style-name="Emphasis">
    <style:text-properties fo:font-variant="normal" fo:text-transform="none" style:font-
      name="Arial" fo:font-size="12pt" fo:font-style="normal" style:text-underline-
      style="solid"
      style:text-underline-width="auto" style:text-underline-color="font-color" fo:font-
      weight="bold" style:font-weight-asian="bold" style:font-weight-complex="bold" />
    </style:style>
    <style:style style:name="T2" style:family="text" style:parent-style-name="Emphasis">
    <style:text-properties fo:font-variant="normal" fo:text-transform="none" style:font-
      name="Arial" fo:font-size="12pt" fo:font-style="normal" />
    </style:style>
    <style:style style:name="T3" style:family="text" style:parent-style-name="Emphasis">
    <style:text-properties fo:font-variant="normal" fo:text-transform="none"
      fo:color="#2300dc" style:font-name="Arial" fo:font-size="12pt" fo:font-style="normal"
    />
  </style:style>
</office:scripts />
```

```

<style:style style:name="T4" style:family="text" style:parent-style-name="Emphasis">
<style:text-properties fo:font-variant="normal" fo:text-transform="none"
fo:color="#4700b8" style:font-name="Arial" fo:font-size="12pt" fo:font-style="normal"
/>
</style:style>
<style:style style:name="T5" style:family="text" style:parent-style-name="Emphasis">
<style:text-properties fo:font-variant="normal" fo:text-transform="none" style:font-
name="Arial" fo:font-size="12pt" fo:font-style="normal" fo:font-weight="bold"
style:font-
weight-asian="bold" style:font-weight-complex="bold" />
</style:style>
</office:automatic-styles>
<office:body>
<office:text>
<office:forms form:automatic-focus="false" form:apply-design-mode="false" />
<text:sequence-decls>
<text:sequence-decl text:display-outline-level="0" text:name="Illustration" />
<text:sequence-decl text:display-outline-level="0" text:name="Table" />
<text:sequence-decl text:display-outline-level="0" text:name="Text" />
<text:sequence-decl text:display-outline-level="0" text:name="Drawing" /> </text:sequence-
decls>
<text:p text:style-name="Standard">
<text:span text:style-name="T1">Temoignage</text:span>
</text:p>
<text:p text:style-name="Standard">
<text:span text:style-name="T2" />
</text:p>
<text:p text:style-name="Standard">
<text:span text:style-name="T2">Au TP, mes copains ont toujours fini de lire l'
</text:span>
<text:span text:style-name="T3">énoncé</text:span>
<text:span text:style-name="T2">avant moi. Et ils n'arrêtent pas de</text:span>
<text:span text:style-name="T4">trépigner</text:span>
<text:span text:style-name="T2">à côté en me demandant si j'ai fini. J'essaye d'
</text:span>
<text:span text:style-name="T4">accélérer</text:span>
<text:span text:style-name="T2">ma lecture, ça me stresse. Je lis en diagonale.
</text:span>
</text:p>
<text:p text:style-name="Standard">
<text:span text:style-name="T2" />
</text:p>
<text:p text:style-name="Standard">
<text:span text:style-name="Emphasis">
<text:span text:style-name="T2">Une fois fini je n'ai rien compris car j'ai lu des mots
mais je n'ai pas essayé de comprendre ce qu'il disait.</text:span>
</text:span>
<text:span text:style-name="Emphasis">
<text:span text:style-name="T5">De toute façon dans l'urgence je suis incapable de
lire.
</text:span>
</text:span>
<text:span text:style-name="T5" />
</text:p>
</office:text>
</office:body>
</office:document-content>

```


Ce n'est rien d'autre qu'un fichier XML avec tout le contenu textuel de notre fichier OpenOffice. De même que le fichier **styles.xml** qui contient que des informations de mise en page.

Le stockage des informations (celles qui nous intéressent : contenu et style) se fait de la manière suivante : le contenu du document et le balisage sont stockés dans le fichier content.xml alors que la définition des styles le sont dans styles.xml. Néanmoins, il est possible que certains styles soient définis dans le fichiers content.xml.

3.3. Stockage des informations de mise en page

Voici un exemple d'une définition de styles :

```
<style:style style:name="MonStyle" style:family="text">
  <style:properties fo:color="#0000ff"
    style:text-underline="double"
    style:text-underline-color="font-color"/>
</style:style>
<style:style style:name="AutrePolice" style:family="text"
  style:parent-style-name="MonStyle">
  <style:properties fo:color="#ff00ff"
    style:font-name="Courier" fo:font-style="italic"
    style:text-underline="none" fo:font-weight="normal"/>
</style:style>
```

L'exemple d'un contenu utilisant ces styles peut s'écrire de la manière suivante:

```
<text:span text:style-name="MonStyle">
  Ceci est un exemple de
</text:span>
<text:span text:style-name="AutrePolice">
  phrase
</text:span>
<text:span text:style-name="MonStyle">
  avec plusieurs attributs de formatage
</text:span>
```

En reprenant notre fichier content.xml, nous constatons aussi cette structure de stockage.

```
<office:automatic-styles>
  <style:style style:name="T1" style:family="text" style:parent-style-name="Emphasis">
    <style:text-properties fo:font-variant="normal" fo:text-transform="none" style:font-
      name="Arial" fo:font-size="12pt" fo:font-style="normal" style:text-underline-
      style="solid" style:text-underline-width="auto" style:text-underline-color="font-color"
      fo:font-weight="bold" style:font-weight-asian="bold" style:font-weight-complex="bold"
    />
  </style:style>
  ...
  ...
```

Il s'agit d'une définition d'un style nommé **T1** de type **Text** dont les propriétés sont entre autres : police : Arial, taille :12, style : gras, souligné... L'utilisation de ce style dans notre document OpenOffice donne :

```

<office:body>
  <office:text>
    <office:forms form:automatic-focus="false" form:apply-design-mode="false" />
    <text:sequence-decls>
      ...
    <text:p text:style-name="Standard">
      <text:span text:style-name="T1">Temoignage</text:span>
    </text:p>
    ...
  </office:text>
</office:body>

```

Agir sur la mise en page d'un document OpenOffice revient à manipuler un document XML. Nous verrons plus loin comment nous allons nous y prendre.

3.4. Les extensions des fichiers associés

OpenOffice.org 2.0 utilise les formats de fichier suivants (formats de fichier OpenDocument) :

Format de document	Extension
Texte OpenDocument	*.odt
Modèle de texte OpenDocument	*.ott
Document maître OpenDocument	*.odm
Document HTML	*.html
Modèle de document HTML	*.oth
Classeur OpenDocument	*.ods
Modèle de classeur OpenDocument	*.ots
Dessin OpenDocument	*.odg
Modèle de dessin OpenDocument	*.otg
Présentation OpenDocument	*.odp
Modèle de présentation OpenDocument	*.otp
Formule OpenDocument	*.odf
Base de données OpenDocument	*.odb

NB : Le format HTML n'est pas un format OpenDocument.

Pour échanger un document avec des utilisateurs utilisant des versions plus anciennes de OpenOffice.org, enregistrez-le dans l'un des formats suivants :

Application	Extension
OpenOffice.org Writer antérieur à OpenOffice.org 2.0	*.sxw
Modèles OpenOffice.org Writer antérieurs à OpenOffice.org 2.0	*.stw
Documents maîtres antérieurs à OpenOffice.org 2.0	*.sxd
OpenOffice.org Calc antérieur à OpenOffice.org 2.0	*.sxc
Modèles OpenOffice.org Calc antérieurs à OpenOffice.org 2.0	*.stc
OpenOffice.org Impress antérieur à OpenOffice.org 2.0	*.sxi
Modèles OpenOffice.org Impress antérieurs à OpenOffice.org 2.0	*.sti
OpenOffice.org Draw antérieur à OpenOffice.org 2.0	*.sxd
Modèles OpenOffice.org Draw antérieurs à OpenOffice.org 2.0	*.std
OpenOffice.org Math antérieur à OpenOffice.org 2.0	*.sxm

Tous ces fichiers sont en fait des archives ZIP. En conclusion, OpenOffice.org propose de nombreuses fonctionnalités pour faciliter la génération de documents bureautiques. Parmi elles, l'omniprésence des styles dans tous les éléments de la suite permet de faciliter le formatage tout en assurant la cohérence interne du document. Les outils permettant de les utiliser sont accessibles et efficaces.

4. Comparaison avec WordML de Microsoft Office

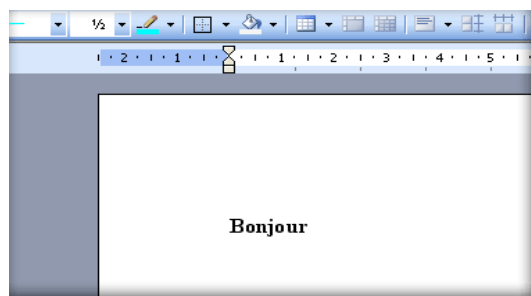
Microsoft Office Word 2003 permet de gérer des documents au format XML : ce sont des documents ne contenant que des données XML, sans aucune information de mise en forme, ou bien des documents Word classiques, avec toute l'information de mise en forme nécessaire, mais représentés avec le format XML.

Dans le dernier cas, il s'agit d'un dialecte XML appelé **WordML**, suffisant pour décrire un document Word dans toute sa richesse. Il s'ensuit que, d'un point de vue de la programmation, il est possible de générer un document Word complet en WordML, à condition de respecter le schéma XML (la grammaire) associé à WordML.

Voyons de plus près comment ressemble la structure XML de Word. En effet si nous mettons le contenu suivant dans un fichier XML :

```
<?xml version="1.0"?>
<w:wordDocument xmlns:w="http://schemas.microsoft.com/office/word/2003/wordml">
  <w:body>
    <w:p>
      <w:r>
        <w:rPr>
          <w:b w:val="on"/>
        </w:rPr>
        <w:t>Bonjour</w:t>
      </w:r>
    </w:p>
  </w:body>
</w:wordDocument>
```

En renommant ce fichier en « .doc », nous obtenons un fichier Word au format WordML, structurellement bien formé dont le contenu est :



Sans aller plus loin, nous remarquons qu'il est tout à fait « facile » de modifier un document Word en passant par sa structure XML. Il faut noter que le format WordML n'est connu que par la version 2003 d'Office et sera étendu dans les futures versions de la suite bureautique de Microsoft. Par ailleurs, il existe une multitude d'éléments XML pour décrire un document WordML qu'il est impossible de tous les connaître. C'est pour cela que nous vous conseillons d'utiliser le fichier d'aide fourni par Microsoft dans le fichier d'installation des schémas Office System dont l'adresse de téléchargement se trouve dans les annexes de ce rapport.

En conclusion, comme pour OpenOffice, il est tout à fait possible de manipuler et même de générer un document Word en passant par son format WordML à l'aide des outils du domaine XML (API, parseurs, transformation XSLT), au lieu par exemple de passer par les objets automation de Word, propres au monde Microsoft.

VI. La structure de mise en page de DysBloc

Nous avons vu dans le chapitre précédent qu'il était beaucoup plus aisé de choisir une structure XML pour manipuler les fichiers texte issus des applications bureautiques les plus utilisées. C'est en effet, la solution que nous avons choisi pour pouvoir stocker puis appliquer les effets de mise en évidence de DysBloc. Plus précisément, nous avons choisi la structure utilisée par le model OpenDocument à cause de sa souplesse d'utilisation, sa syntaxe facile à déchiffrer car proche du HTML et la disponibilité des outils de transformation qu'offre OpenOffice.

1. Les solutions de mise en évidence choisies

Nous rappelons que le but de DysBloc est de parvenir à modifier un document texte par des règles de mise en page et ajout d'information pour permettre une meilleure lisibilité du contenu et sa compréhension. Le tableau suivant présente les styles de mise en évidence que l'utilisateur peut appliquer dans une règle de mise en évidence

Attributs de Formatage retenus	
style:font-name	nom de la police ex: "Arial"
fo:font-size	taille de la police ex: "20pt"
fo:color	Couleur de la police ex : "#000000"
fo:font-weight	style gras ex: "bold" / "normal"
style:text-underline-style	souligné ex: "solid" / "none"
style:text-underline-width	taille du soulignement ex: "auto"
style:text-underline-color	couleur de soulignement ex: "font-color" / "#00FF33"
fo:background-color	Couleur d'arrière plan ex: "transparent" / "#44AB24"

Le choix de ces styles est influencé par les recommandations sur l'accessibilité. Nous remarquons par exemple que le style italique n'a pas été retenu car il est difficilement lisible pour les personnes dyslexiques.

Il existe plusieurs outils écrits avec des langages différents pour gérer les documents XML. Cependant, nous avons choisi de développer notre outil avec le langage Java. Les raisons de ce choix sont les suivantes :

- Portabilité : Java est un langage reconnu sur toutes les plateformes.
- Java et XML : Il existe plusieurs API Java pour manipuler les documents XML dont nous connaissons déjà le fonctionnement : DOM, SAX, JDom...
- Java et XSL : Il est aujourd'hui facile d'appliquer des transformation XSL à l'aide des outils (Processeurs, Analyseur...) Java comme Xalan, Xerces...
- Etc.

2. Gestion des informations de mise en évidence

Les informations de mise en évidence sont stockées dans des objets Java que nous avons nommés **Style**.

Style
name fontName fontSize fontColor weightValue underlineValue underlineColor ...
setFontName setFontColor ... getFontName getFontColor ...

Les Styles sont stockés dans un fichier XML (styles.xml) dont un extrait de sa structure est illustré sur la figure suivante :

```

- <style:style style:name="FONTE-JAUNE" style:family="text">
  <style:text-properties fo:color="#FFFF00" style:font-name="Arial" fo:font-size="16pt" style:text-underline-
    style="none" style:text-underline-width="auto" style:text-underline-color="font-color" fo:font-
    weight="16pt" fo:background-color="transparent" style:font-size-asian="16pt" style:font-weight-
    asian="none" style:font-size-complex="16pt" style:font-weight-complex="none" />
</style:style>
- <style:style style:name="FONTE-ORANGE" style:family="text">
  <style:text-properties fo:color="#FF9900" style:font-name="Arial" fo:font-size="16pt" style:text-underline-
    style="none" style:text-underline-width="auto" style:text-underline-color="font-color" fo:font-
    weight="16pt" fo:background-color="transparent" style:font-size-asian="16pt" style:font-weight-
    asian="none" style:font-size-complex="16pt" style:font-weight-complex="none" />
</style:style>
</DYSTYLE>

```

Dans une balise **<style>**, nous retrouvons l'ensemble des informations nécessaires pour appliquer une mise en forme d'un texte de type OpenOffice. Bien que cette mise en forme soit limitée en possibilités, nous avons évoqué les raisons de cette restriction qui sont dues aux recommandations sur l'accessibilité.

3. Analyse et traitement de texte

Les étapes de traitement d'un contenu texte

Pour transformer du texte (à partir de l'interface DysBloc ou d'un fichier texte) en contenu accessible selon le profil de destination, il faut appliquer au maximum les quatre étapes suivantes :

1^{ère} étape (facultative) :

Transformer le document d'entrée en fichier OOo (on utilisera les extensions .ODT ou .SXW).

2^{ème} étape :

Extraire et modifier les fichiers :

- **content.xml** : pour l'ajout de contenu ou l'application des mises en page et/ou
- **styles.xml** : pour l'application des mises en page

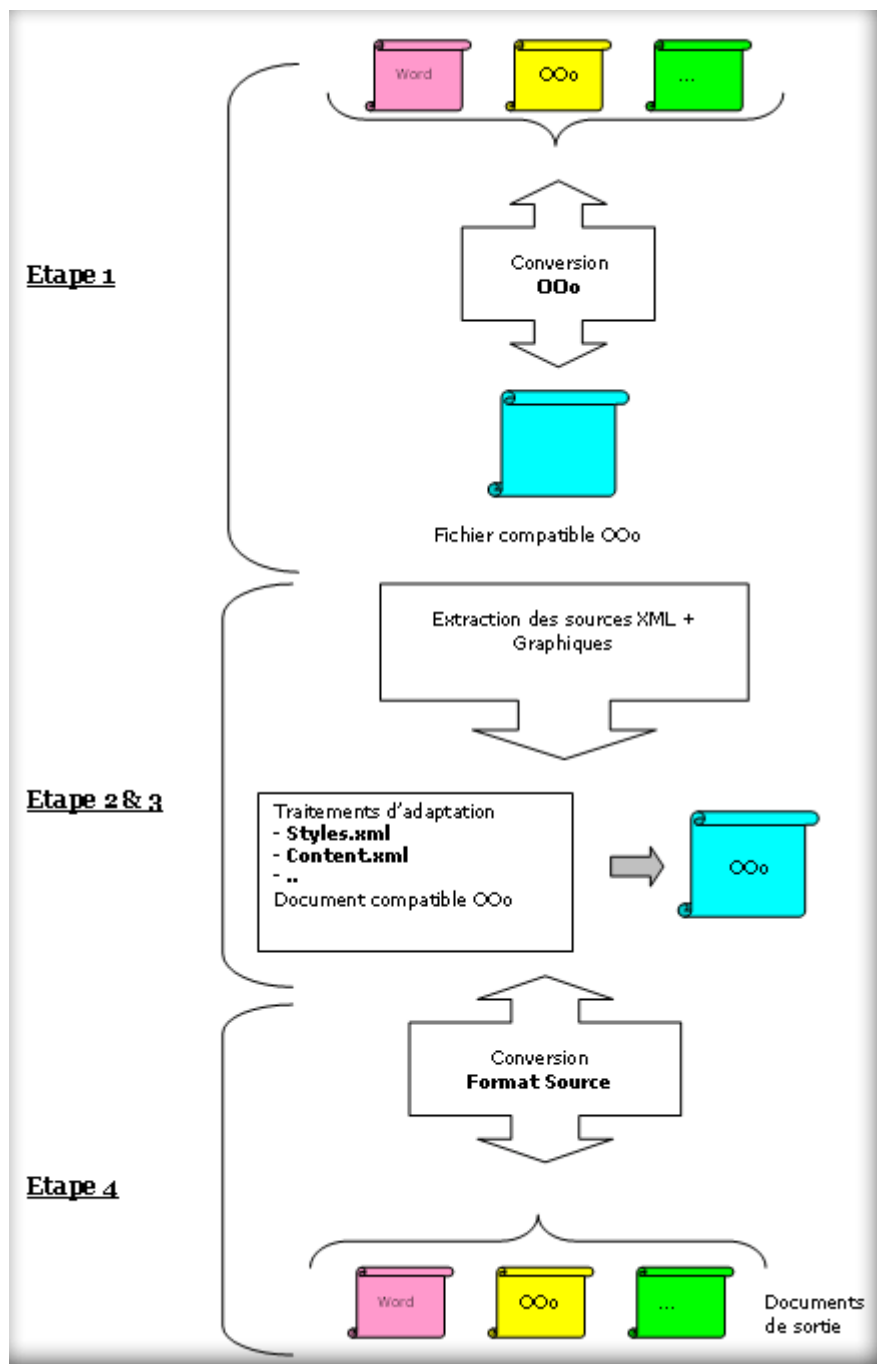
3^{ème} étape :

Régénérer le fichier OOo (.ODT ou .SXW).

4^{ème} étape (facultative) :

Restituer le fichier au format souhaité par transformation XSL.

La figure suivante illustre cette procédure :



4. Les règles d'adaptation

La tâche principale du logiciel est de faire « une transformation » d'un document texte en document texte accessible. Pour ce faire, nous avons besoin de savoir quel type de transformation on doit faire subir notre document d'entrée. Cela dépend comme on l'a dit précédemment des difficultés du lecteur final.

Pour représenter ces difficultés ou plutôt les solutions utilisées pour les contourner, nous avons créé le concept de « règle d'adaptation ».

4.1. Définitions

a. Une règle d'adaptation

C'est une fonction de mise en évidence de traits particuliers d'un texte. Elle comprend la règle elle-même et les paramètres de mise en évidence visant à accentuer la présence d'une difficulté donnée.

Exemple : Une règle : la distinction des syllabes

Paramètres : première syllabe en vert, la prochaine en bleu

Ainsi on peut définir :

b. Une transformation

Elle consiste à générer à partir d'un document texte, un autre document qui sera mise en page de manière à faciliter son accessibilité. Le document de sortie, qui peut être de format différent, doit garder le même contenu, la même sémantique. Une transformation est une succession de règles d'adaptation (une ou plusieurs).

4.2. Structure d'une règle d'adaptation

Il existe deux types de règle d'adaptation :

a. Les règles statiques

Ce sont les règles qui ne tiennent pas compte du contexte d'apparition des groupes de lettres. Elles s'appliquent à toutes les occurrences dont elles font référence.

Exemple : la règle : séparation des mots

Tout ensemble de lettres séparé par des espaces

(Au début et à la fin) sera considéré comme étant un mot.

b. Les règles génériques

Il s'agit plutôt des règles qui tiennent compte du contexte orthographique, grammatical... du groupe de lettres.

Exemple : la règle : distinction des syllabes

Une syllabe n'est pas forcément un ensemble de deux lettres (consonne/voyelle). L'entité syllabe change de forme selon des contextes (par exemple, le mot dans lequel elle se trouve).

4.3. Gestion des règles d'adaptation

Une règle d'adaptation est représentée par un objet Java que nous avons appelé **Rule**.

Rule
name type (static/generic) List (objet encapsulant les informations sur le contexte) ...
getName() getType() getRows() setName(...) isGeneric()

Les règles d'adaptation sont stockées dans un fichier selon un schéma XML (rules.xml) dont la structure est la suivante :

```
<?xml version="1.0" encoding="UTF-8" ?>
- <RULE>
- <frule name="Mise en évidence des graphemes 'an', 'en', 'in', 'on', 'un'">
  <regex val="en" style="FONTE-BLEU" />
  <regex val="an" style="FONTE-JAUNE" />
  <regex val="in" style="FONTE-VIOLET" />
  <regex val="on" style="FONTE-VERT" />
  <regex val="un" style="FONTE-ORANGE" />
</frule>
- <grule name=... >
  ...
  ...
</grule>
```

Le fichier rules.xml est en quelque sorte la base de données des règles d'adaptation. Cette structure nous offre des avantages considérables surtout celui de réutilisation des règles. En effet, il est possible d'exporter ou d'importer des règles d'adaptation pour ne pas avoir à recréer celles qui existent déjà. De même la modification des règles reste facilement faisable. Elle permettra de créer rapidement des règles à partir des models.

5. Interface de communication JAVA/XML

La communication entre les bases XML et le programme est assurée par une interface mise au point à l'aide de JDom.

JDOM est une API du langage Java développée par des développeurs indépendants. Elle permet de manipuler des données XML plus simplement qu'avec les API classiques. Son utilisation est pratique pour tout développeur Java et repose sur les API XML de Sun.

- JDOM utilise des collections SAX (Simple API for XML) pour parser les fichiers XML. Ce type de parseur utilise des événements pour piloter le traitement d'un fichier XML.
- JDOM utilise DOM (Document Object Model) pour manipuler les éléments d'un Document spécifique (créé grâce à un constructeur basé sur SAX). DOM est une spécification du W3C pour proposer une API qui permet de modéliser, de parcourir et de manipuler un document XML.

JDOM permet donc de construire des documents XML, de naviguer dans leur structure et de modifier ou supprimer leur contenu. Il est en vérité très laborieux de développer des applications complexes autour de XML avec DOM, qui rappelons le, n'a pas été développé spécifiquement pour Java. Nous avons choisi d'utiliser JDOM pour sa simplicité.

Les objets **RuleManager** et **StyleManager** assurent l'interface de communication entre les données XML et le programme. Ils font appel à JDom pour parser les fichiers XML et les transformer en objets Java.

```
public static void parse(String file){
    cFile = file;
    SAXBuilder sxb = new SAXBuilder();
    try
    {
        document = sxb.build(new File(cFile));
        root = document.getRootElement();
    }
    catch(Exception e){...}
}
```

La suite des opérations se fait grâce aux méthodes de lecture et d'écriture directe aux données XML. Exemple la lecture de toutes les règles présentes dans la base :

```
public static Vector getRules() {
    Vector v = new Vector();
    List lf = getElementFixedRules();
    List lg = getElementGenericRules();

    for (int i=0; i<lf.size(); i++){
        v.addElement(elementToRule((Element)lf.get(i)));
    }

    for (int i=0; i<lg.size(); i++){
        v.addElement(elementToRule((Element)lg.get(i)));
    }
    return v;
}
```

6. Analyse du contenu

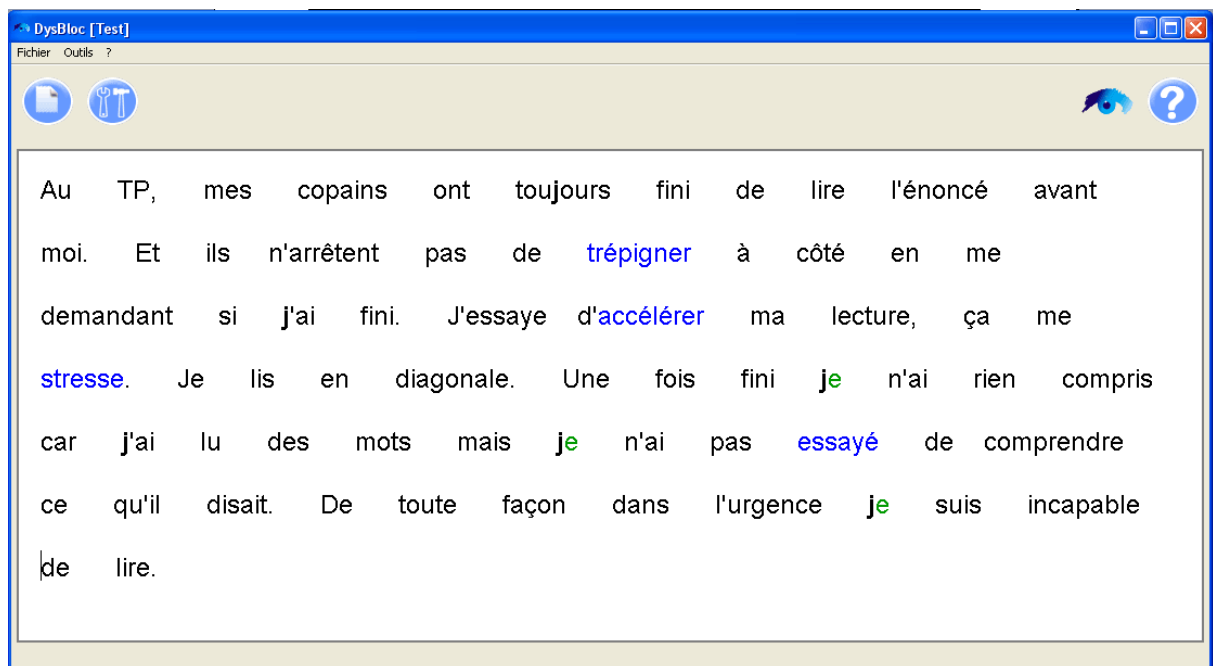
Pour analyser le contenu textuel des fichiers entrants, la classe TextTool.java s'appuie sur JDom pour récupérer le contenu à partir du fichier content.xml (voir la structure d'un document OpenOffice).

VII. L'interface de DysBloc

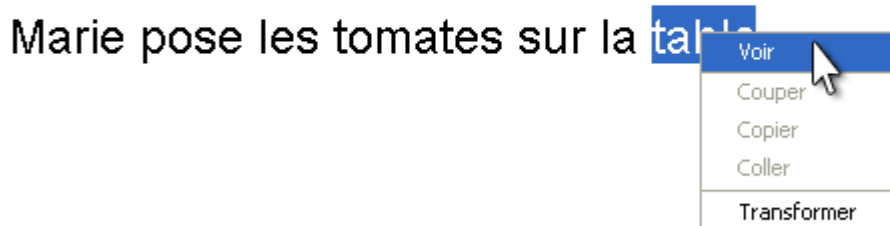
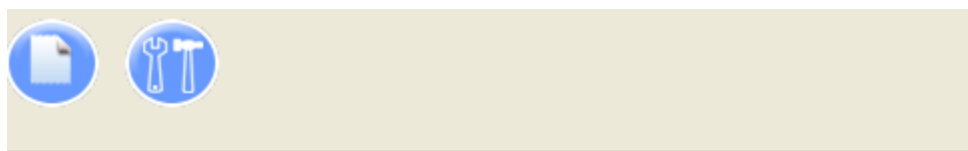
L'interface de DysBloc est faite en Java Swing. Elle se repose sur un modèle classique des interfaces des applications bureautiques. L'écran principal intègre une barre de menues et une barre d'outils pour les accès rapides.

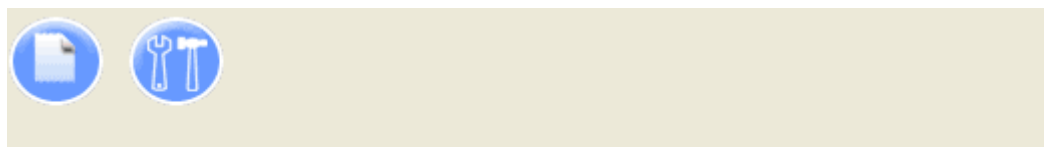
DysBloc est conçu en deux modules graphiques distincts mais parfaitement communicants. Le premier assure l'affichage du contenu principal et le second est composé des fenêtres de création de règle, de transformation et de configuration. Cette séparation permet une facilité de maintenance dans le futur et surtout une évolution de la structure des objets métiers (exemple : ajout de nouveaux paramètres dans les règles) sans toucher l'interface principale.

L'écran suivant présente l'interface principale de DysBloc.



Un module d'assistance à la compréhension (aussi au stade d'élaboration) permet de donner une illustration d'un mot voire d'une phrase.



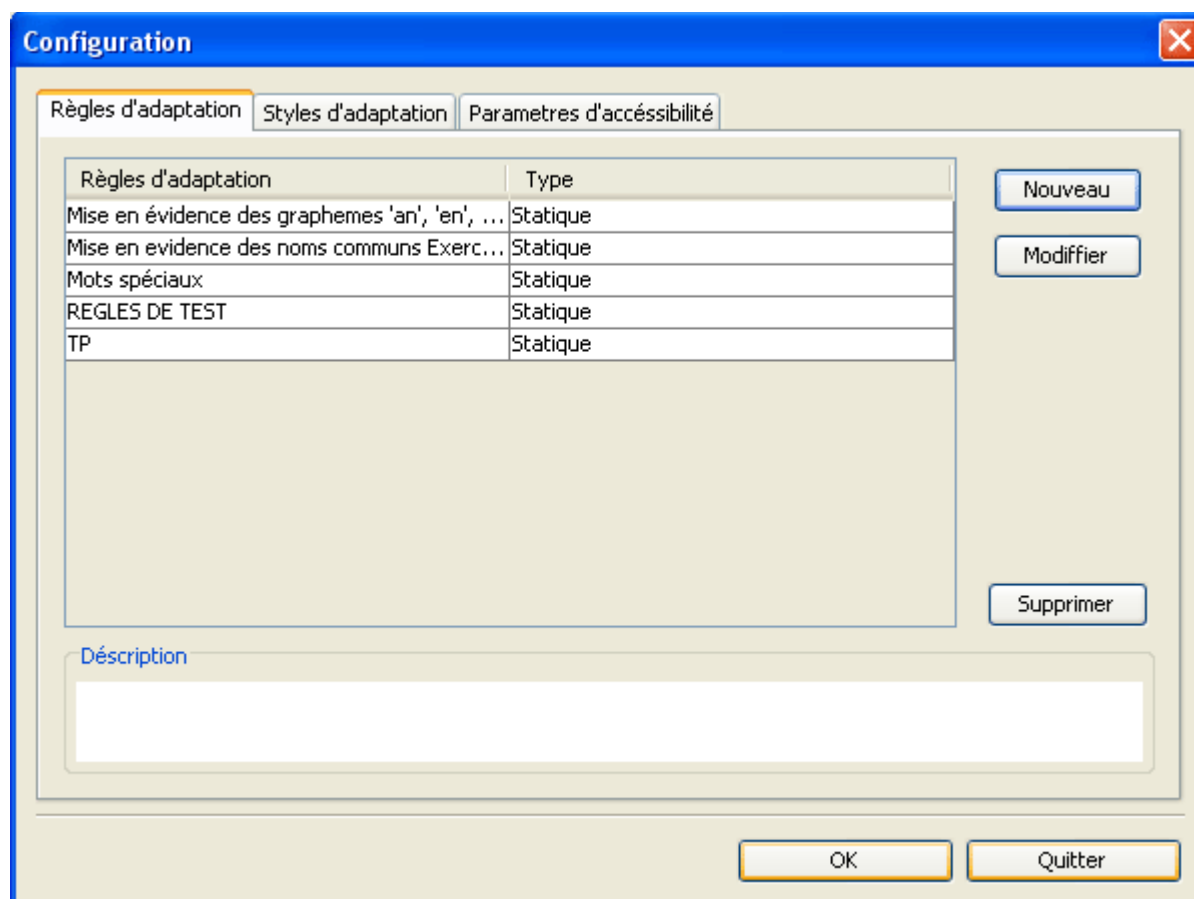


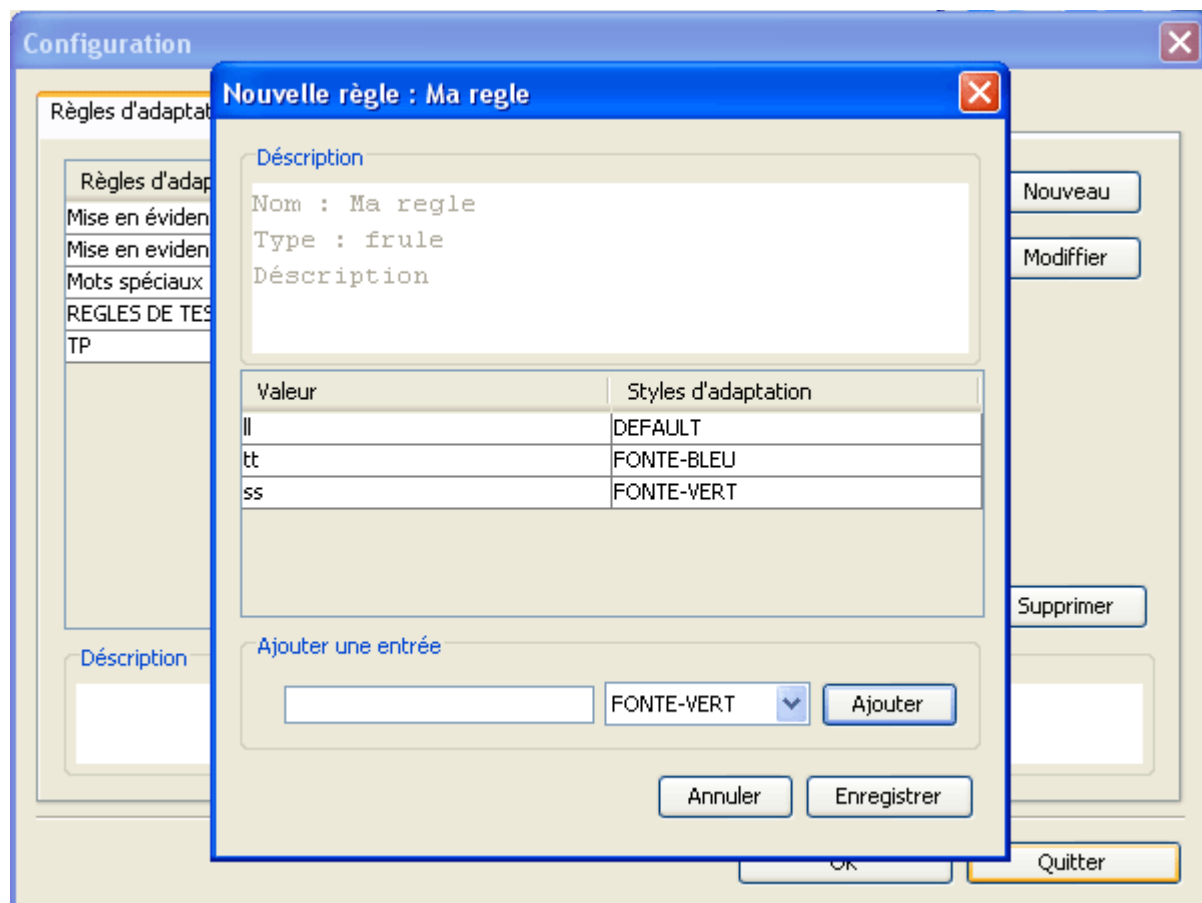
Marie pose les tomates sur la table



table

Autres capture d'écran :





Conclusion

Il existe un certain nombre de troubles qui empêchent ou retardent l'apprentissage du langage chez l'enfant. Les signes langagiers remarquables repérables à l'école ou dans le cadre familial peuvent constituer l'expression d'un trouble du langage.

Ces troubles spécifiques du langage regroupent ceux qui sont spécifiques au développement du langage oral (retards de parole et de langage, dysphasie) et ceux qui affectent le développement du langage écrit (dyslexie). L'origine exacte de ces troubles reste en très grande partie inconnue.

Ainsi les activités qui consistent à réveiller les réflexes langagiers, reprogrammer les compétences cognitives élémentaires ou renforcer les capacités préservées, doivent être privilégiées car elles ont montré leur efficacité dans l'apprentissage.

Le projet DysBloc s'inscrit dans ce cadre des applications d'apprentissage évolutif et d'accompagnement face aux difficultés de lecture.

Ce logiciel qui reste, à l'heure de la rédaction de ce rapport, à l'état de prématuré, présente un cahier de charge dont les objectifs visent à sensibiliser les éditeurs sur les difficultés d'accessibilité de l'information écrite. Il propose des modèles de présentation de données selon les difficultés du lecteur.

Des nouvelles technologies de présentation de l'information telles que le XML, le XSL... sont au centre du projet. Celles-ci doivent venir au secours du lecteur et jouer un rôle d'assistance pour adapter l'information au lecteur sans pour autant altérer le sens.

Il reste beaucoup de fonctionnalités à développer et d'autres à perfectionner. Aussi l'outil n'a pas encore été testé pour des raisons de période scolaire. Pour être complètement opérationnel, DysBloc doit être confié à des enseignants spécialisés, des personnes dyslexiques et à de nombreux acteurs pour construire un véritable gage de soutien à la lecture et la compréhension de texte. Il me paraît clair qu'un tel outil peut faire ses preuves très vite ou à moyen terme.

DysBloc est, en son état actuel, qu'un début de projet que j'espère sera poursuivi pour qu'il soit validé et adopté.

Références

CECIAA ceciaa.com

CLEVER AGE cleverage.com

CNEFEI "Les troubles spécifiques du langage" [CNEFEI]

CORIDYS <http://www.coridys.asso.fr>

DOCTISSIMO <http://www.doctissimo.fr>

INDESKO indesko.com

Marc Delahaie "L'évolution du langage chez l'enfant : De la difficulté au trouble"
[INPES 2004]

Ministère de la santé <http://www.sante.gouv.fr>

MICROSOFT microsoft.com

Mot par Mot <http://www.motparmot.com>

OASIS oasis.org

OPENOFFICE.ORG OpenOffice.org

SUN MICROSYSTEMS Sun.fr

The University of Manchester www.intranet.man.ac.uk

Universités de Grenoble grenoble-universites.fr

W3C / WAI W3c.org

Annexes

Il existe de nombreux logiciels et aides techniques qui assistent les enfants dyslexiques en difficulté et les aident à réintégrer un cycle scolaire normal.

De la synthèse vocale à la mise en page adaptée de texte, l'assistance des personnes dyslexiques s'appuie de plus en plus aux nouvelles technologies et notamment à l'informatique.

Nous présentons quelques outils traitant les critères les plus sensibles, dont une bonne partie est issue du catalogue des produits recommandés par CECIAA.

Aides à l'écrit

Kurzweil 3000 : Logiciel de traitement de texte très complet pour lire, écrire, orthographier, étudier particulièrement adapté aux besoins spécifiques des élèves et étudiants dyslexiques.

Points forts : vocalisation de l'écrit, dictionnaires extensibles

Dragon Naturally Speaking Professional: Suite logiciel très efficace pour la transcription automatique des paroles sur l'écran pour créer, modifier des documents textes (lettres, rapports, emails...) sans l'aide du clavier ou de la souris.

Il intègre des fonctionnalités permettant un usage professionnel de la vocalisation.

Points forts : reconnaissance/synthèse vocale

Penfriend FR : Outil de prédiction de mots. Penfriend s'adapte à la plupart des logiciels de traitement de textes et offre un retour vocal du contenu textuel.

Penfriend est aussi efficace pour l'aide à la lecture. Il reconnaît presque tout type de document numérisé. La lecture se fait sur une fenêtre indépendante du contenu qui peut provenir d'une page Web, email, document PDF...

Le suivi de la lecture est soutenu par une fonctionnalité de mise en évidence des mots en cours de lecture. Possibilité d'adapter le rendu visuel (police, couleur, arrière plan...) aux préférences de l'utilisateur.

Points forts : prédiction des mots

Médialexie (EditDys) : Logiciels facilitant la lecture et l'écriture. Les outils Médialexie sont particulièrement adaptés pour les enfants et les adultes dyslexiques ou bien qui souffrent de handicaps moteurs leur interdisant l'utilisation du clavier d'un ordinateur. Ces outils sont spécialisés dans le traitement automatique du langage naturel, la synthèse vocale, la correction automatique de l'écrit et de l'affichage adapté.

Aides à la lecture

Screenruler : Outil de grossissement de caractères, screenruler permet une meilleure visibilité de tout le contenu d'un document. En effet, les règles d'accessibilité du WAI (Web Accessibility Initiative / w3.org) telles que la police sans serif, les couleurs, double interligne... ne sont pas forcément appliqués aux documents. Ainsi la lecture à l'écran peut s'avérer fastidieuse et fatigante pour cette

catégorie de population. Screenruler masque certaines parties pour permettre au lecteur de concentrer sa vision sur la partie du document en cours de lecture. Cela lui permet de poursuivre ainsi une lecture évolutive et non encombrante. Ces fonctionnalités apportent une réduction considérable de la fatigue oculaire et une meilleure concentration.

ZoomText Magnifier : Dans la catégorie des outils de grossissement de caractères, zoomtext apporte une solution comparative à celle d'une loupe. Cela permet d'exploiter toute la surface de l'écran et de pouvoir appliquer un agrandissement allant jusqu'à 16 fois. Il permet aussi de personnaliser l'apparence du document (couleur de fond et texte, police, curseur...). Ce qui lui donne une mention spéciale en matière de respect des normes d'accessibilité.

Wordread : Simple d'utilisation et moins encombrant, Wordread se présente sous forme d'une barre d'outils simplifiée intégrant une synthèse vocale qui assiste au lecteur dans l'exploration de document textuels.

Lecture à la demande pour les document Word ou lors de la saisie de texte. Lecture depuis le presse-papier pour les autres documents (Pdf, Rtf...).

Wordread vous permet de créer vos propres podcasts (fichiers son) à partir de vos documents textes et de les réécouter n'importe où.

Logiciels multifonctionnels

De part les troubles liés à la lecture et à l'écriture pour les personnes dyslexiques, dans la plupart des cas, ils ont des difficultés à apprendre et comprendre des nouvelles connaissances. Ces difficultés sont dues du faite qu'ils ont des problèmes d'organisation des idées et des pensées.

Il existe des logiciels permettant de pallier aux difficultés de structuration de la pensée et de modélisation d'un scénario pour une visualisation globale et simplifiée.

Inspiration Dyslexie : est un outil qui permet de structurer les idées, d'organiser la pensée et d'avoir une présentation schématique de toute forme de données et informations.

A partir d'une idée principale, Inspiration Dyslexique permet de réaliser une structure d'organigrammes composée de diagrammes, de concepts, de brainstormings (remue-méningses), de schémas illustratifs... pour aboutir à un plan de travail. Ainsi, l'utilisateur peut facilement transformer cette structure modélisée en un rapport ou une présentation claire et concise.

PCA : Plateforme de Communication Alternative est un logiciel d'aide à la communication verbale qui est caractérisé par l'homogénéité de son utilisation quel que soit le moyen d'expression choisi. Le codage se fait avec du texte et/ou pictogrammes, soutenu par un moteur de prédiction et une base lexico pictographique riche.