

De l'écrit à l'oral : analyses et générations

Fabrice Maurel

IRIT – Université Paul Sabatier
118, route de Narbonne 31062 Toulouse - France
fmaurel@irit.fr

Résumé – Abstract

Longtemps considérée comme ornementale, la structure informationnelle des documents écrits prise en charge par la morpho-disposition devient un objet d'étude à part entière dans diverses disciplines telles que la linguistique, la psycholinguistique ou l'informatique. En particulier, nous nous intéressons à l'utilité de cette dimension et, le cas échéant, son utilisabilité, dans le cadre de la transposition automatique à l'oral des textes. Dans l'objectif de fournir des solutions qui permettent de réagir efficacement à cette « inscription morphologique », nous proposons la synoptique d'un système d'oralisation. Nous avons modélisé et partiellement réalisé le module spécifique aux stratégies d'oralisation, afin de rendre « articulables » certaines parties signifiantes des textes souvent « oubliées » par les systèmes de synthèse. Les premiers résultats de cette étude ont conduit à des spécifications en cours d'intégration par un partenaire industriel. Les perspectives de ce travail peuvent intéresser la communauté TAL en reconnaissance de la parole, en génération/résumé de texte ou en multimodalité.

Considered for a long time as ornamental, the informational structure of written documents carried by texts morpho-disposition becomes a full object of investigation in various disciplines such as linguistic, psycholinguistic or computer sciences. In Particular, we are interested in the utility of these aspects of documents and, if the need arises, their usability, within the framework of their oral transposition. In the objective to provide solutions which make it possible to react effectively to this “morphological inscription”, we propose the synoptic of an oralisation system. We modelled and partially realized the module specific to the oralisation strategies, in order to render some signifying parts of the text often “forgotten” by synthesis systems. The first results of this study led to specifications in the course of integration by an industrial partner. The prospects of this work can interest NLP community in voice recognition, text generation/summarization or multimodality.

Mots Clés – Keywords

Architecture textuelle, synthèse de la parole, stratégies d'oralisation
Textual architecture, speech synthesis, oralisation strategies

1 Introduction

La question que nous abordons dans cet article est celle de l'importance d'une prise en compte de la structure informationnelle qui est prise en charge par la morpho-disposition des textes. Ces structures architecturales, plus ou moins combinées à d'autres composantes structurelles, peuvent avoir une fonction et une valeur significative. Ainsi, il est important que cette composante soit étudiée et intégrée dans les applications en traitement automatique des langues (Scott, 2003). En particulier, nous nous intéressons à l'utilité des aspects morpho-dispositionnels des documents et, le cas échéant, leur utilisabilité, dans le cadre de leur transposition à l'oral par des moteurs de synthèse de parole à partir de textes (TTS). Dans ce contexte, la question vient naturellement : quels sont les modèles linguistiques et les outils technologiques à notre disposition pour rendre compte de ces aspects à l'oral en fonction des caractéristiques des usagers (handicaps, ...), de l'environnement (bruité, type de support de l'interaction, ...) et du type de tâche/activité (recherche d'information, ...) ?

Notre recherche pose l'hypothèse d'une unité de traitement à la fois plus efficiente (plus « cognitive » ?) que la phrase mais malgré tout réaliste : l'Objet Textuel (OT). Aussi, la section suivante s'attachera à justifier ce choix en mettant en évidence cette dimension propre aux textes écrits pour être lus silencieusement qui, bien que peu traité par des logiciels de synthèse de la parole à partir de texte, nous semble incontournable pour une oralisation correcte en terme d'intelligibilité. Notre approche revêtant un caractère fortement pluridisciplinaire, nous étudierons la prise en compte de cet aspect du texte à travers la littérature de la linguistique textuelle, de la psycholinguistique et de l'informatique. Nous décrirons ensuite la synoptique d'un système d'oralisation utilisant l'OT comme unité de traitement, puis nous focaliserons sur l'étape visant à mettre en œuvre les différentes stratégies envisageables pour leur oralisation. Ce dernier module est exemplifié, et son opérationnalisation en partie démontrée, à travers l'exemple des structures énumératives, en s'appuyant sur différents modèles et typologies dont nous décrierons l'intérêt pour notre travail.

2 Valeur et prise en compte de la morpho-disposition des textes

2.1 Du point de vue des linguistes

Longtemps la mise en forme des textes n'a pas été considérée comme un objet d'étude à part entière. Aujourd'hui encore, peu de manifestations scientifiques proposent de traiter de ce sujet. Malgré tout, nous remarquons depuis quelques années un intérêt croissant pour l'étude de la morpho-disposition des textes afin de prendre en compte cette dimension pour la génération de document, le résumé automatique ou la synthèse de parole. Cependant peu de modèles permettent d'appréhender globalement cet objet d'étude, d'un point de vue théorique et formel, et donc d'envisager des traitements informatiques. A notre connaissance, une des seules tentatives pour parvenir à un tel résultat est décrite par le Modèle d'Architecture Textuelle (MAT) (Virbel, 1989), (Pascual, 1996). Dans le cadre de notre étude, ce modèle nous a semblé un point de départ intéressant pour soutenir la mise en œuvre de stratégies d'oralisation de l'architecture du texte.

Ce modèle définit l'**architecture de texte** comme une composante abstraite du texte qui prend sa source dans la notion de **Mise en Forme Matérielle** (MFM) et dans l'hypothèse de l'existence d'une équivalence fonctionnelle entre des phénomènes typographiques, dispositionnels et lexico-syntaxiques. La MFM est un sous-ensemble de propriétés morpho-dispositionnelles du texte possédant des équivalents langagiers. Les constituants des textes, **les Objets Textuels** (OT), sont perceptibles par le jeu de contraste de la MFM et l'architecture de texte est l'ensemble des objets textuels et les propriétés qu'ils entretiennent entre eux. L'approche générale de l'architecture textuelle est sous-tendue par les travaux de (Harris, 1971) qui pose que la langue peut être décrite par elle-même, c'est à dire que l'on peut exhiber des phrases à portée méta-textuelle explicitant le fonctionnement de la langue et obéissant aux mêmes lois de constructions grammaticales que les phrases de la langue. L'ensemble de ces phrases forme le **métalangage**. Par exemple, la phrase « Max est sujet de *Max mange une pomme* » est une phrase du métalangage. Harris pose qu'en règle générale ce métalangage est réduit mais que cette réduction laisse des traces dans la langue. Des travaux ont été menés par l'équipe de M. Gross (Gross, 1990) et ont pu valider cette approche pour le français. Ce principe a été appliqué aux propriétés morpho-dispositionnelles du texte : elles peuvent être décrites à l'aide de phrases méta-textuelles (par exemple : « je segmente mon texte en trois parties », « je commence par une introduction consacrée à la linguistique », etc.) et la réduction de ces phrases laisse des traces dans le texte qui sont des propriétés typographiques et dispositionnelles. L'ensemble des phrases décrivant les propriétés morpho-dispositionnelles forment le **métalangage architectural**. Cette approche a pour conséquence l'existence d'un continuum entre des formes de MFM entièrement discursives et des formes visuelles : on passe d'une extrémité à l'autre de ce continuum en effaçant (au profit de marques visuelles) ou en reconstruisant (par interprétation) le métalangage architectural.

2.2 Du point de vue des psycholinguistes

Du fait que les marques de MFM sont considérées par la majorité des modèles de la compréhension comme des traits de surface, leur rôle est réduit à une fonction quasi ornementale. Ainsi, peu (ou pas) de modèles de la compréhension du texte se sont intéressés à intégrer l'implication des marques structurales visuelles dans le traitement cognitif du texte (Schmidt, 2002) et par conséquent, on note une pauvreté de travaux ayant étudié l'existence d'une telle implication et les mécanismes qui y sont sous-jacents. Parmi les travaux qui se sont intéressés à l'impact des traits de surface sur le traitement cognitif, une première catégorie s'est penchée sur le rôle des signaux dans la gestion des traitements mis en œuvre durant la lecture (Schmidt, 2001). D'autres ont montré leur implication dans le macro-traitement du texte (Goldman, 1995), (Lorch, 1996) en mettant en évidence leur participation à l'identification des thèmes. (Nespoulous, 2002) va même jusqu'à avancer la notion d'avenance textuelle comme une transposition de la notion d'affordance au texte ; cette notion renvoie au fait que ce sont les propriétés de l'environnement qui conditionnent notre traitement de l'information. Le texte posséderait à des degrés divers et à travers l'utilisation des marques de MFM une affordance conditionnant son traitement. A ce titre, ces marques seraient bien plus qu'ornementales, elles constitueraient une enveloppe spatiale que le lecteur appréhenderait pour orienter son traitement du texte.

2.3 Du point de vue des informaticiens

Un texte ne peut être réduit à une suite linéaire de mots qui se constituent en phrases. En particulier, des procédés spécifiques peuvent se constituer en configurations particulières, intentionnelles et signifiantes (titres, exemples, insistances, ...), appelées Objets Textuels (Cf. 2.1), auxquelles les technologies actuelles de synthèses de parole à partir de textes ne savent que partiellement réagir.

Par exemple, le texte à gauche¹ exprime des différences logico-linguistiques dans les titres d'un même document : le premier titre est totalement indépendant (syntactiquement) de la suite du développement. Dans le cas du 2^{ème} titre, le texte commence par une reprise anaphorique du titre (il → le directeur). Enfin, dans le 3^{ème} cas, on a une continuation syntaxique entre le titre (sujet) et la suite du texte. Chacun de ces titres devrait avoir un impact différent sur l'oralisation.

<p>A) LE CONSEIL D'ADMINISTRATION</p> <p>La Résidence Borde Haute est un Etablissement géré par l'APMP à but non lucratif.</p> <p>B) LE DIRECTEUR</p> <p>Il est votre interlocuteur et a pour attribution la gestion de l'Etablissement.</p> <p>Il ordonne les dépenses et les recettes dans le cadre du budget visé par le Conseil d'Administration.</p> <p>Il est responsable du bon fonctionnement de la Maison de Retraite.</p> <p>C) LE PERSONNEL</p> <p>rattaché à l'établissement est présent 24 heures sur 24.</p> <p>Les professions médicales et paramédicales sont joignables à tout moment.</p>

Même certaines marques de ponctuation (lorsqu'elles sont traitées) sont en relation quasi-bijective avec les effets que leur traitement par un système de TTS produit : une marque entraîne toujours le même effet. Or, par exemple, un simple tiret ne devrait pas aboutir à une oralisation équivalente selon qu'il est utilisé :

- pour marquer le début d'un item d'énumération,
- comme une marque d'incise – parenthétique ou commentaire –,
- en lieu et place d'un séparateur de type « , » – « ; » – « . », ...

Nous avons exploré dans un premier temps la capacité des systèmes de TTS à réagir à certaines spécificités des textes écrits pour être lus silencieusement (et non pour être dits). Des résultats de ces observations a émergé la synoptique d'un système qui nous semble plus adapté à la perspective d'une transposition automatique à l'oral des structures visuelles des textes (Figure 1).

L'originalité de ce système réside dans l'unité d'analyse considérée pour le traitement des documents écrits. En effet, elle ne correspond plus à la traditionnelle phrase ponctuée et n'a pas non plus l'ambition d'intégrer le texte globalement avec son co-texte et ses contextes. En fait, le processus d'oralisation décrit par cette architecture permet de faire « traverser » à un OT plusieurs phases de transformation qui s'appuient sur un ensemble de connaissances associées à chacune d'elles. Ces trois étapes (normalisation, extraction, reformulation) permettent, à partir d'un objet textuel « écologique », de fournir une représentation, associée à une stratégie d'oralisation particulière, interprétable par un TTS. Notre intérêt s'est porté dans un premier temps sur la faisabilité de l'automatisation du processus de reformulation à travers l'acquisition des connaissances nécessaires et suffisantes pour l'élaboration des

¹ Le texte rend compte de l'hétérogénéité des textes réels étudiés. Sa qualité médiocre est due à la mauvaise facture du document original.

représentations permettant un traitement informatique de cette étape (parties grisées de la Figure 1).

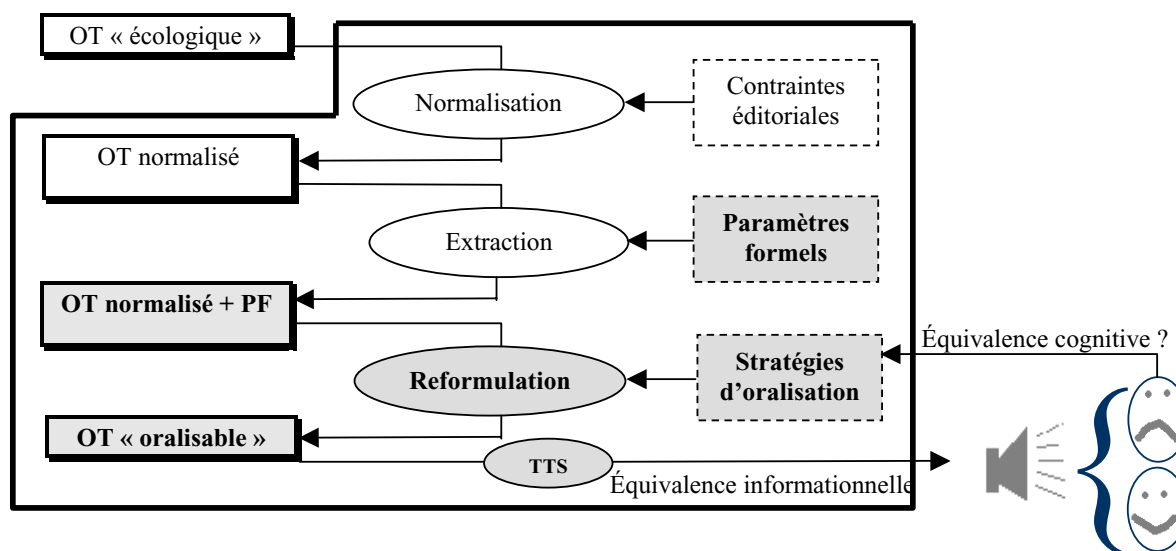


Figure 1 : Synoptique du système d'oralisation des structures visuelles des textes

3 Le module des stratégies d'oralisation

Notre premier objectif pour aborder la phase de reformulation a été de concevoir un modèle des stratégies d'oralisation prenant en compte la morpho-disposition des textes. L'axiome principal qui sous-tend le modèle prétend qu'il est nécessaire d'obtenir une forme discursive de l'OT commune aux deux modalités et que, de là, pourra être décrit un ensemble de stratégies utilisant des procédés spécifiques à la modalité orale. Sur ces bases, nous avons conçu le Modèle d'Oralisation des Textes Ecrits pour être Lus Silencieusement (**MOTELS** – Figure 2). La forme discursive d'un Objet Textuel (OT) particulier peut être reconstruite selon trois stratégies : par une description objective de ses marqueurs : forme discursive descriptive (**S1 - FDDC**), par un premier niveau d'interprétation, en attribuant un nom à cette configuration morpho-dispositionnelle : forme discursive dénominative (**S2 - FDDN**) ou par une interprétation plus profonde des intentions du rédacteur qui se traduisent par l'utilisation de cet OT particulier : forme discursive interprétative (**S3 - FDI**). Ces trois opérations pourront chacune conduire ensuite à 4 continuums de nouvelles formes d'oralisation : la FDDC pourra se réduire grâce à l'utilisation de conventions « typophoniques » par un procédé d'association bijective entre un ensemble de marques et un ensemble de sons (**C1**) ; la FDDN pourra se réduire par une convention « typophonique » qui associe de manière injective les marques morpho-dispositionnelles définissant un type d'objet particulier à un son (**C2a**) ou par un schéma prosodique particulier, repéré comme étant associé au type d'OT étudié indépendamment de tout contexte d'énonciation (**C2b**) ; la FDI pourra se réduire par l'utilisation de schémas prosodiques dépendant de la situation d'énonciation et donc de la structure informationnelle de la FDI (en termes de thème, topique, focus, rhème - **C3**).

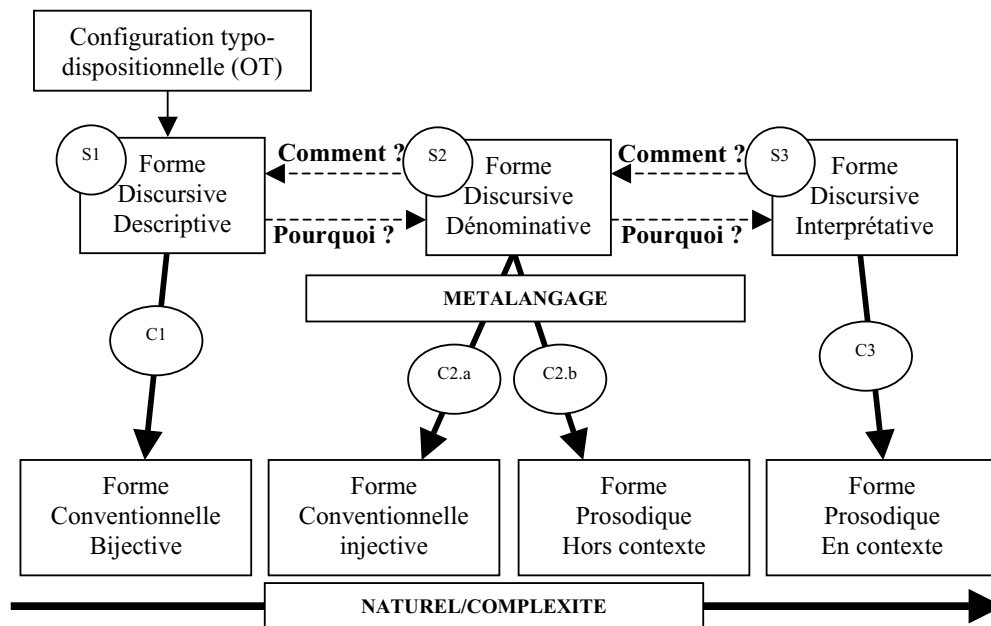


Figure 2 : Modèle d'Oralisation des Textes Écrits pour être Lus Silencieusement (MOTELS)

Une illustration des 3 stratégies dans leur forme discursive peut être donnée à travers l'exemple suivant. Soit le titre d'un document :

L'ORALISATION DES TEXTES

S1 : « **le texte écrit en gros, centré, en gras et en capitale est l'oralisation des textes** »

S2 : « **Le titre de niveau 1 est l'oralisation des textes** »

S3 : « **Je vais vous parler maintenant de l'oralisation des textes** »

Les parties en gras seront les marques lexico-syntaxiques restituées à partir de la morpho-disposition du texte. Ce sont aussi celles qui pourront potentiellement être effacées au profit de marques sonores telles que décrites plus haut. Nous remarquerons, comme l'indique le schéma de gauche à droite, que les trois formes discursives sont en relation à travers une sémantique de « pourquoi » (respectivement de droite à gauche une sémantique de « comment », probablement plus proche du travail cognitif du rédacteur d'un texte). Dans l'exemple du titre : (S3) comment parler maintenant de X ? → (S2) en faisant de X un titre ; comment titrer ? → (S1) en centrant, soulignant, ...

4 Application aux structures énumératives

Le modèle d'oralisation présenté ci-avant démontre la multitude de formes équivalentes d'un OT qu'il est envisageable d'obtenir du point de vue du contenu informationnel qu'il véhicule. Un certain nombre de critères scientifiques et applicatifs ont déterminé les OT et les stratégies sur lesquelles focaliser notre recherche dans un premier temps. Nous avons étudié les **structures énumératives**, les **titres** et les **structures d'insistance** afin de développer un traitement informatique capable de les restituer selon les trois formes *a priori* les plus intéressantes : **discursive dénominate**, **discursive interprétative** et **prosodique en contexte**. Nous décrirons ici la démarche appliquée et les résultats obtenus pour la transposition à l'oral actuellement la plus aboutie : celle des structures énumératives.

4.1 Paramètres formels et représentation des structures énumératives

En s'appuyant sur des observations de corpus et des travaux existants (Luc, 2001), (Maurel, 2002) nous avons pu établir les différents paramètres formels des structures énumératives à prendre en compte. Cet objet est défini par : une amorce (phrase introductive précédant et annonçant l'énumération qui peut être syntaxiquement complète ou incomplète ; dans le cas d'incomplétude, les constituants manquants sont fournis par un ou plusieurs items), un item (entité co-énumérée), une énumération (un ensemble d'items), une structure énumérative (l'ensemble composé d'une amorce et d'une énumération).

Les structures énumératives sont signalées par une architecture impliquant des marques lexico-syntaxiques, typographiques et dispositionnelles variées (*ibid*). Ces marques constituent l'ensemble des paramètres formels des structures énumératives, en ce sens qu'ils participent à la forme visuelle de l'OT. Nous rajouterons que les critères de **parallélisme**, d'**exhaustivité** et d'**ordonnement** peuvent également influencer la structure visuelle de cet OT. Automatiser la phase de reformulation nécessite l'élaboration d'une représentation des structures énumératives qui tienne compte de ces paramètres formels. Nous avons démontré les lacunes du langage XHTML (Maurel, 2003a) et proposons une DTD XML pour les combler (Figure 3 pour un exemple).

Ce qui est (parfois!) fait en HTML	Ce que nous faisons
<p>Les activités du centre social sont:</p> <p></p> <p>- <i>la gestion de six appartements d'urgence que propose la commune</i> ; </p> <p>- <i>l'organisation de réunions de formation collectives</i> ; </p> <p>- <i>le montage des projets de formation.</i></p> <p></p>	<pre><SE type="parallèle" exhaus="oui" ordon="non"> <AMORCE classifieur="activités du centre social " vs="énumérées" gnspec="ci-après"/> <ENUMERATION> <ITEM format="absgn" repclass="activité"> <NOYAU> <i>la gestion</i> </NOYAU> <i>de six appartements d'urgence que propose la commune</i></ITEM> <ITEM format="absgn" repclass="activité"> <NOYAU> <i>l'organisation</i> </NOYAU> <i>de réunions de formation collectives</i> </ITEM> <ITEM format="absgn" repclass="activité"> <NOYAU> <i>le montage</i> </NOYAU> <i>des projets de formation</i> </ITEM> </ENUMERATION></SE></pre>

Figure 3 : Exemple d'une représentation HTML versus XML

De là peuvent être construites automatiquement les formes dénominative et interprétative à condition qu'elles s'inscrivent dans un *pattern* de métaphrases prédéfinies (Figure 4).

Version discursive dénominative	Version discursive interprétative
<p>L'auteur annonce une énumération de trois items concernant les activités du centre social. Le premier item de l'énumération concerne la gestion de six appartements d'urgence que propose la commune. Le deuxième item de l'énumération concerne l'organisation de réunions de formation collectives. Le troisième et dernier item de l'énumération concerne le montage des projets de formation.</p>	<p>Les trois activités du centre social sont énumérées ci-après. Une première activité possible est la gestion de six appartements d'urgence que propose la commune. Une deuxième activité possible est l'organisation de réunions de formation collectives. Une troisième et dernière activité possible est le montage des projets de formation.</p>

Figure 4 : Exemples de formes discursives dénominative et interprétative

4.2 De la FDI à la forme prosodique en contexte (FPEC)

Il s'agissait de déterminer les marques prosodiques qui ont la même fonction que les marques lexico-syntaxiques « architecturantes » de la version discursive. Des tentatives infructueuses d'étude par corpus pour faire émerger des régularités prosodiques nous ont conduits à une démarche pluri-fondée : **écoute** qualitative (attentive et critique) d'un recueil de discours oraux , **introspection** soumise à discussion avec des spécialistes et **projection des règles dictées par un modèle de l'intonation pragmatique du français** (Rossi, 1999) sur la relation entre des marques lexico-syntaxiques d'un OT et sa version prosodique. Cette méthode a amené à conférer à l'effacement des métaphrases au profit de marques prosodiques un **rôle de conservation de la structuration informationnelle** (en terme de thématisation, rhématisation, topicalisation ou focalisation). Nous avons appliqué cette démarche générale aux structures énumératives (Maurel, 2003b). Ce travail a permis d'évaluer le pouvoir descriptif, au niveau de notre problématique, du modèle intonatif de M. Rossi. Lorsque ce cadre théorique n'apportait de solutions satisfaisantes, nous avons eu recours à d'autres résultats issus de la littérature (Morel, 1998), (Vannier, 1999) et à un protocole expérimental (en particulier pour traiter la stratégie pausale et l'intonation de fin d'item). Nous avons alors procédé à l'examen systématique des spécifications du langage SSML² dont un rôle est de normaliser la représentation de phénomènes prosodiques. Trois cas se sont présentés : (1) utilisation de marqueurs SSML sans modifications des spécifications (e.g : <BREAK >); (2) utilisation de marqueurs SSML avec modifications des spécifications (e.g : <PROSODY pitch="calcul" param="item:i:n"); (3) mise en oeuvre de nouveaux marqueurs non décrits dans les spécifications SSML (e.g. : FOCUS type="étendu" | "sémantique"). Finalement, à partir de notre représentation des structures énumératives et une feuille de transformation XSLT il nous est possible d'obtenir notre forme prosodique. L'exemple de la figure 3 est repris ci-dessous.

```
<SE> Les <FOCUS type="étendu"> trois activités du centre social </FOCUS> sont <BREAK size="large"> <AUDIO src="uri"/> <PROSODY pitch="calcul" param="item:1:3"> <FOCUS type="sémantique"> la gestion </FOCUS> de six appartements d'urgence que propose la commune <BREAK type="medium"> </PROSODY> <PROSODY pitch="calcul" param="item:2:3"> <FOCUS type="sémantique"> l'organisation </FOCUS> de réunions d'information collectives <BREAK type="medium"> </PROSODY> <PROSODY range=Xst> <FOCUS type="sémantique"> le montage </FOCUS> des projets de formation </PROSODY> </SE>
```

Par exemple, (Rossi, 1999:56) décrit le « morphème intonatif de topique contextuel référentiel » comme « montrant du doigt » ce dont on va dire quelque chose ; l'allocutaire prenant ainsi connaissance du référent juste avant de recevoir l'information. Nous avons utilisé ce morphème intonatif dans la version prosodique des structures énumératives, à la fin du syntagme nominal de l'amorce contenant le classifieur des items (ci-dessus : « social ») : nous substituons par cette marque prosodique la marque lexico-syntaxique « ci-après » de la forme discursive dont le rôle est précisément de « montrer du doigt » les entités énumérées qui suivent. Les spécifications formulées pour l'interprétation de l'ensemble des balises utilisées ont été intégrées dans un système de synthèse de parole à partir de texte d'un partenaire industriel³.

² Speech Synthesis Markup Language - <http://fr.w3.org/TR/2003/CR-speech-synthesis-20031218/>

³ Elan Speech – Tempo Technology – www.elan.fr

5 Perspectives et conclusions

Nous avons montré qu'une équivalence en terme de contenu informationnel était envisageable entre un texte écrit pour être lu silencieusement et sa version orale. Idéalement, l'obtention de cette équivalence passe par une analyse fine de chaque objet textuel susceptible de se trouver dans un document. A partir de ces observations nous construisons manuellement des *patterns* aptes à « recevoir » les versions discursives de l'OT. Nous nous appuyons sur des observations et des cadres théoriques de la prosodie du français pour déduire, de ces versions discursives, des versions prosodiques. Nous étudions les possibilités de représentation de cette dernière dans le langage SSML (interprétable par un TTS). Nous proposons, le cas échéant, des spécifications pour de nouvelles balises qu'un partenaire industriel peut intégrer dans son système. Nous déduisons finalement une représentation de type XML de l'OT, adéquate pour obtenir par transformation (XSLT) nos deux versions. Aujourd'hui, ce travail a été partiellement réalisé pour trois objets textuels : les structures énumératives, les titres et les structures d'insistances.

Si il a été ainsi possible de construire des représentations qui opérationnalisent différentes stratégies d'oralisation d'OT, nous n'avons que peu de données sur les moyens à mettre en œuvre pour obtenir automatiquement ces représentations à partir d'un document original. Ainsi une perspective incontournable pour concevoir un système réellement automatisable se dégage : **l'extraction automatique, dans les documents tout-venant, des paramètres formels des OT** (e.g. : pour l'énumération - extraction du classifieur des items, du nombre d'items, ...).

Une seconde perspective concerne **l'évaluation de l'équivalence de nos formes non plus en terme de contenu informationnel, mais en terme de traitement cognitif** : un protocole expérimental est en cours pour évaluer l'impact des stratégies discursives interprétatives et prosodiques sur la mémorisation/compréhension de l'information par des jeunes versus âgés (les dégradations dans la perception de la prosodie et dans les capacités mnésiques pouvant potentiellement influencer sur l'efficacité des stratégies d'oralisation).

Les perspectives de ce travail peuvent intéresser la communauté TAL en reconnaissance de la parole (repérage des marques architecturantes du discours oral pour une mise en forme automatique du texte), en génération de texte (par l'utilisation du MAT) ou en multimodalité (la transmodalité comme un moyen d'étude de la multimodalité (Maurel, 2003c)).

Références

GOLDMAN S.R., SAUL E.U., Coté N. (1995), Paragraphing, Reader, and Task, effects in discourse comprehension, *Discourse Processes*, Vol. 20, pp. 273-305.

GROSS M. (1990), Sur la notion harrissienne de transformation et son application au français, *Langage*, n°99, pp. 39-56.

HARRIS Z.S. (1971), *Structures Mathématiques du Langage*, Paris, Dunod.

HARRIS Z.S. (1982), *A Grammar of English on Mathematical Principles*, *Wiley-interscience*, New York.

- LORCH R.F., LORCH E.P. (1996), Effects of organizational signals on free recall of expository texts, *Journal of Educational Psychology*, Vol. 88, N°1, pp. 38-48, 1996.
- LUC C. (2001), Une typologie des structures énumératives basée sur les structures rhétoriques et architecturales du texte. Actes de *TALN'2001*, pp. 263-272.
- MAUREL F., LUC C., MOJAHID M., VIGOUROUX N., VIRBEL J., NESPOULOUS J.-L. (2002), De l'influence de la variabilité des textes sur leur transposition automatique à l'oral, Actes de *Conférence Internationale sur le Document Électronique 5*, pp. 211-225.
- MAUREL F., RAYNAL M., ORIOLA B., VIGOUROUX N. (2003a), Extension of W3C Standards for Real Web Accessibility, Actes de *Human - Computer Interaction 2003*, Vol. 4, pp. 995-999.
- MAUREL F., LEMARIÉ J., VIGOUROUX N. (2003b), Oralisation de structures visuelles : de la lexico-syntaxe à la prosodie, Actes de *Interface Prosodique 2003*, pp. 137-142.
- MAUREL F., VIGOUROUX N., RAYNAL M., ORIOLA B. (2003c), Contribution of the Transmodality Concept to Improve Web Accessibility, Dans M. MOKHTARI (eds). *Assistive Technology Series Vol.12 (ICOST'03 – Independent living for persons with disabilities and elderly people)*. Chapter 4 – computer accessibility, IOS Press, Amsterdam, pp. 186-193.
- MOREL M.-A., DANON-BOILEAU L. (1998), *Grammaire de l'intonation*, Paris, Ophrys.
- NESPOULOUS J.-L., DE MATTOS PIMENTA M.A., VIRBEL J. (2002), Des objets textuels à leur traitement par le cerveau humain, Actes de *Inscription Spatiale du Langage 2002*, pp. 77-87.
- PASCUAL, E. (1996), Integrating Text Formatting and Text Generation, *Trends in Natural Language Generation : an Artificial Intelligence Perspective*. Réds : ADORNI M. et ZOCK M., Eds: Springer-Verlag , pp. 205-221.
- ROSSI M. (1999), *L'intonation, le système du français : description et modélisation*, Paris, Ophrys.
- SCOTT D. (2003), *Layout in NLP: the case for Document Structure*, Invited presentation to ACL'03 Sapporo - <http://www.itri.brighton.ac.uk/~Donia.Scott/slides.html>, Japan, July 2003.
- SCHMID S. (2001), Représentation organisationnelle et compréhension : rôle de la mise en forme matérielle dans la lecture, Thèse de doctorat, Nice Sophia Antipolis.
- SCHMID S. (2002), Rôle de la représentation spatiale des textes dans la compréhension, Actes de *Inscription Spatiale du Langage*, pp. 289-296.
- VANNIER G. (1999), *Étude des contributions des structures textuelles et syntaxiques pour la prosodie : application à un système de synthèse vocale à partir de texte*, Thèse de doctorat, Caen.
- VIRBEL J. (1989), The Contribution of linguistic knowledge to the interpretation of text Structure, Dans J. ANDRÉ, V. QUINT et R. FURUTA (eds.), *Structured Documents*, Cambridge University Press, pp. 161-181.