



HAL
open science

Modéliser la prosodie pour la synthèse à partir du texte : perspectives sémantico-pragmatiques

Anne Lacheret-Dujour, Michel Morel

► To cite this version:

Anne Lacheret-Dujour, Michel Morel. Modéliser la prosodie pour la synthèse à partir du texte : perspectives sémantico-pragmatiques. Franck Neveu; Peter Blumenthal; Nicole Lequerler. Au commencement était le verbe : syntaxe, sémantique et cognition : mélanges en l'honneur du Professeur Jacques François, 97, Peter Lang, pp.299-325, 2011, (Sciences pour la communication), 978-3-03-431063-5. halshs-01687846

HAL Id: halshs-01687846

<https://shs.hal.science/halshs-01687846v1>

Submitted on 18 Jan 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Modéliser la prosodie pour la synthèse à partir du texte: Perspectives sémantico-pragmatiques

Anne LACHERET-DUJOUR, MODYCO, université de Paris Ouest Nanterre
Michel MOREL, CRISCO, université de Caen

Introduction

Le propos de cet article est de présenter un modèle fonctionnel¹ de génération semi-automatique des proéminences prosodiques non démarcatives, de nature pragmatique, en tenant compte de l'interaction entre la structuration prosodique d'un message et les unités lexicales qui le constituent. Celles-ci sont appréhendées selon 2 angles combinés, l'un lexicologique, le second textuel: (1) en tant qu'unités types dont la signification lexicale hors contexte constitue un potentiel d'accentuabilité pragmatique, (2) en tant qu'unités occurrences, qui, inscrites dans un contexte donné, sont tributaires a) des constructions syntactico-sémantiques dans lesquelles elles apparaissent et b) de leur statut occurrence dans le texte (1^{ère} ou nième occurrence).

La génération automatique de la prosodie pour la synthèse de la parole a donné lieu à deux générations de modèles, les premiers basés exclusivement sur la syntaxe, les seconds intégrant les contraintes rythmiques. Aujourd'hui encore les systèmes effectivement implémentés le sont essentiellement sur ces deux critères et force est de constater que la qualité du discours produit en sortie est loin de donner entière satisfaction. Les experts en sont conscients et for de ce nouvel enjeu, *i.e.* l'invention des modèles de troisième génération, l'effort commun porte essentiellement sur la simulation de

1 Cet article est donc centré sur le repérage des fonctions sémantico-pragmatiques de l'accent, la dimension formelle (corrélats acoustiques associés) n'étant qu'évoquée ici (note 23).

l'expressivité vocale. La capacité d'exprimer et d'identifier des émotions, des intentions ou des attitudes par la modulation de la voix est en effet fondamentale. Pour désigner cet ensemble, nous utilisons le terme générique *d'expressivité* tout en sachant que ces différents aspects, maîtrisés ou non, de la production d'un énoncé recouvrent plus d'une catégorie, du paraverbal au symbolique.

Si, dans le domaine du paraverbal, les travaux abondent pour modéliser et simuler l'expressivité vocale, ils sont envisagés la plupart du temps indépendamment de tout ancrage linguistique. Dans ce contexte, les recherches consacrées au balisage sémantico-pragmatique des textes en vue de la génération de structures prosodiques restent sous représentées malgré les hypothèses prometteuses de quelques expériences récentes. C'est dans cette seconde mouvance que se situe le travail ici conduit consacré à la modélisation de la saillance prosodique en discours.

Partant du principe de la naturalité des structures défendu par les grammaires fonctionnelles et cognitives, nous posons l'hypothèse que les processus de saillance, loin d'être linguistiquement contingents, dépendent à la fois de la projection dans le texte de certaines propriétés lexicales des unités et de contraintes pragmatiques qui imposent une sélection ciblée, contrôlée et fonctionnelle des unités à marquer. Nous nous intéressons donc au processus de saillance textuelle sous ses différents angles: sémantique, pragmatique et prosodique dont il s'agit de préciser aussi rigoureusement que possible l'interface.

Nous posons comme premier principe que produire un texte, c'est produire un réseau de connexions entre des unités de différents ordres (mot simple, collocation, cadre syntaxique complet). Étant donné les contraintes de sélection lexicale et fonctionnelle posées plus haut, la constitution du réseau n'est pas uniforme: tandis que certaines unités sont prosodiquement **saillantes**, *i.e.* accentuées, d'autres ne le sont pas et c'est bien la présence des secondes qui permet l'émergence des premières. Sous **l'angle interprétatif**, la perception des contrastes entre formes prosodiquement fortes et faibles permet de prendre la mesure des différences qualitatives d'un texte. En conséquence, sous **l'angle de la génération**, nous posons comme second principe la nécessité de disposer d'un modèle qui intègre les contraintes symboliques sous-jacentes à la production des structures et à la variation qualitative des formes.

De cette posture méthodologique dérive un balisage sémantico-pragmatique du texte à synthétiser qui permet de compléter notre modèle de génération actuel, en tenant compte, au-delà de la syntaxe et du rythme, des contraintes sémantiques et communicationnelles pour prédire l'ensemble des proéminences accentuelles d'un texte (accent primaire mais également secondaire²).

La constitution d'un modèle de génération des proéminences prosodiques contraintes par la structure sémantico-pragmatique du texte se heurte à une difficulté majeure: la variation potentielle importante des schémas accentuels. D'une part, le même énoncé, la même phrase, la même unité lexicale ne comportera pas les mêmes proéminences selon le contexte (situation discursive, cotexte et enjeux énonciatifs), ce qui pose notamment le problème de l'étude hors contexte. D'autre part, à contexte identique, nous sommes inéluctablement confrontés à une variation inter et intra-locuteur significative. Un tel modèle ne peut donc que représenter les proéminences probables et il faut tenir compte de cette incertitude non seulement lors de la conception du modèle mais aussi lors de son évaluation. Par conséquent, si l'approche met au cœur de son dispositif d'analyse les textes comme observatoire, les lignes qui suivent rendent compte également du problème classique de la représentativité des corpus, quelle que soit leur taille (Péry-Woodley 1995, Cappeau et Gadet 2007)³.

Notre démarche est partie de 4 questions centrales articulant contraintes syntaxiques, lexicales et discursives.

La première s'inscrit dans le champ des analyses qui réclament une appréhension globale de la forme et du contenu et, dans ce contexte, tentent de mieux cerner l'interface syntaxe-sémantique-prosodie. Elle repose sur le concept de *cadre syntaxique* et l'hypothèse, aujourd'hui classique, de composition holistique des structures (Legallois & François 2006 dir.). En pratique, l'hypothèse défendue est la suivante: un cadre syntaxique est doté d'une signification qui convoque en elle-même et par elle-même un certain type de schéma accentuel. Concrètement, lorsque nous analysons des données langagières sur corpus, on constate que certaines constructions syntaxiques sont en elles-mêmes porteuses de schémas expressifs et de ce point

² Tels qu'ils sont définis dans Lacheret & Beaugendre 1999.

³ Voir Habert « Gros c'est beau » vs « l'insécurité des grands ensembles » <http://atala.biomath.jussieu.fr/je/010428/Habert/Perpignan00/node3.html>

de vue contraignent les constructions prosodiques instanciées dans le message parlé (par ex: ***l'heure***, *c'est **l'heure!** c'est à mourir de rire! elle est à gifler! s'il faut en passer par là*. Il est donc possible de construire un lexique de telles constructions qui s'enrichit progressivement au fur et à mesure des observations conduites sur corpus. Reste à savoir comment représenter ces cadres syntaxiques dans le lexique ou, dit autrement, quelle entrée squelettale sélectionner? Ainsi, pour nos exemples ci-dessus, les entrées suivantes: <N c'est N>, <c'est à Vinf de N>, <{N,PRO} V à Vinf>; <en Vinf {par/de/à}> constituent-elles des entrées pertinentes, suffisamment cernées pour ne pas créer des effets de bord? Pour l'heure, faute de recherches massives sur corpus et de résultats statistiquement représentatifs, la question reste ouverte et nous ne l'aborderons pas plus en détail dans la suite de cet article.

La seconde question a trait à l'inventaire de l'expression linguistique de la subjectivité: unités affectives ou évaluatives, expression du point de vue de l'énonciateur dans le discours, en général des noms, adjectifs et adverbes modaux consignables dans un lexique dorénavant dénommé *lexique énonciatif* (*infra*, § 2.1).

Quant à la troisième, il s'agit de modéliser les relations entre la fréquence d'occurrence d'une unité telle qu'elle peut être fixée à l'issue de l'exploration de grosses bases de données textuelles et son statut informationnel dans le discours. L'hypothèse est la suivante: si les items fréquents sont peu variables prosodiquement et en général atones, les items rares sont sujets à variation dans le discours et donc potentiellement porteurs d'un accent pragmatique. Il est donc possible de construire un lexique informationnel basé sur la fréquence dans le discours (*infra* §2.2).

Pour finir, la prise en compte des processus de mémorisation discursive nous a permis d'affiner notre modèle. L'approche est, là aussi, classique: il s'agit de tenir compte des phénomènes de progression textuelle et d'enchaînement thématique: l'essentiel des mots mis en relief dans leur première occurrence, parce qu'ils sont informationnellement denses (fonction de focus), ne le sont plus par la suite puisqu'ils assument la fonction de reprise thématique (*infra*, § 2.3).

1. Conception du modèle

Notre modèle a été construit progressivement autour de 6 étapes qu'il s'agit de résumer ici pour ensuite détailler les différents maillons de la chaîne.

La première étape a consisté à créer une liste d'unités à valeur de quantification et/ou contrastive, dans leur dimension essentiellement évaluative et affective pour constituer le lexique énonciatif.

Dans la seconde, fondée sur l'exploitation de la base de données Frantext, il s'est agi de développer un lexique des 4000 formes les plus courantes du français, flexions comprises. Outre les informations fréquentielles fournies par Frantext, comme dans d'autres travaux (Baumann & Riester 2010), elle repose sur l'hypothèse d'une corrélation fonctionnelle entre l'opposition lexicologique générique/spécifique, l'opposition pragmatique unité non informationnelle/informationnelle et l'opposition prosodique unité non proéminente/proéminente.

Dans la troisième étape, nous avons pris en compte le statut occurrence d'une unité et sa saillance mémorielle en nous fondant sur l'hypothèse classique que la saillance prosodique d'une unité est inversement proportionnelle à sa saillance cognitive (Lacheret 2003, Lacheret et François 2004) et qu'il y a donc un lien direct entre le statut d'un référent accessible textuellement⁴ parce que déjà évoqué et son profil accentuel.

La quatrième étape était consacrée à l'implémentation de ce premier jeu de règles combinant informations lexicologiques hors contextes et informations contextuelles (statut occurrence d'une unité dans un texte et phénomène de mémorisation) dans notre système de synthèse Kali et au test de leur rendement fonctionnel sur un corpus de 5400 mots. A l'issue de ce travail, l'importance relative de chaque type de règle a pu être évaluée.

A la lumière des résultats obtenus, environ 50 % des erreurs ont été corrigées (cf. exemples en *infra*). Elles ont été validées (cinquième étape) par la recherche de chaque item dans un corpus journalistique de 46000 mots et dans la base de données Frantext. Leur accentuation pragmatique dans différents contextes a été examinée: les corrections n'ont été validées que si elles s'avéraient pertinentes dans une majorité de contextes (cf. détails en *infra*). Cette analyse des erreurs a permis de nuancer la relation posée

4 Cf. l'échelle d'accessibilité des référents posée par Chafe (1987).

comme a priori bi-univoque entre la fréquence d'une unité et sa saillance prosodique, distinguant ainsi 4 classes d'unités: a) les unités effectivement fréquentes et non saillantes prosodiquement, b) les unités effectivement non fréquentes et saillantes prosodiquement, c) les unités saillantes prosodiquement bien que fréquentes et d) les unités non saillantes prosodiquement bien que non fréquentes.

Dans la sixième et dernière étape, une mise à l'épreuve du modèle dans son état actuel a été conduite sur un extrait de la base de données PFC (Phonologie du français contemporain), la section 2.6 lui est consacrée.

2. Description détaillée du modèle

2.1. Etape 1: lexique énonciatif

Le lexique énonciatif comporte à ce jour 180 mots constituant un ensemble apparemment disparate, qui possèdent en réalité une propriété commune de quantification, soit temporelle et aspectuelle⁵ (*aussitôt, bien avant, brusquement*), soit spatiale (*énorme, immense, minuscule*), soit relative au nombre et à la quantité (*au moins, aucun, beaucoup, chaque, ni, tous*), soit modal (*bonnêtement*). Cette propriété a fréquemment deux caractéristiques d'ailleurs liées: 1) elle dérive de la force illocutoire et de l'engagement énonciatif (*absolument*), 2) elle relève souvent de l'expression linguistique de la subjectivité et de l'affect (*c'est formidable! vs c'est affreux! je suis si heureux!*). De ce point de vue, les constructions dans lesquelles s'insèrent ces mots marquent l'expression d'un contraste par rapport à l'attente supposée du co-énonciateur et sont donc prosodiquement saillantes⁶. En outre, ces mots ne s'inscrivent pas dans le schéma de la progression textuelle classique (*infra*, § 2.3.) et ne subissent pas d'effet de thématization; au contraire, leur répétition crée un effet d'insistance parfaitement maîtrisé sur le plan rhétorique (*le même jour, à la même heure; ni les poètes, ni surtout les savants; autant il est radin, autant sa sœur est généreuse*). Ce lexique est le premier consulté par le programme, qui

5 Valeur inchoative.

6 Pour ce type d'hypothèse, voir aussi Morel & Danon-Boileau 1998.

génère ainsi environ 33 % des balises pragmatiques du modèle. La marque pragmatique porte sur l'ensemble du groupe prosodique – ou *mot phonologique* (Selkirk 1978, Nespor & Vogel 1983)⁷: *aucun homme, absolument pas*, etc.

2.2. Etape 2: lexique informationnel basé sur la fréquence

Pour mettre en place le balisage pragmatique, nous avons, dans un second temps, eu recours à l'hypothèse classique aujourd'hui, selon laquelle moins une observation est probable plus cette observation est porteuse d'information (voir dans la théorie de l'information, Shannon 1948). Reformulée pour la linguistique (Martinet 1960), cette théorie démontre que plus une unité lexicale est rare statistiquement, plus elle a de chances d'être informative. C'est ainsi que nous avons créé une première version d'un lexique informationnel basée sur la fréquence d'apparition des mots dans un large corpus construit à partir de la base de données Frantext limitée au XX^e-XXI^e siècle, tout genre discursif confondu: essais, romans, théâtre, poésie, correspondance, articles de journaux, mémoires, etc. Nous avons obtenu un classement des 75 millions de mots de ce corpus, répartis suivant la loi de Zipf⁸ par ordre décroissant du nombre d'occurrences. Par exemple, la première unité classée, *de*, a été rencontrée 3 millions de fois. Nous avons situé la frontière entre items fréquents et moins fréquents de telle façon que le taux de saillances prosodiques prédit par le modèle soit voisin de celui observé dans les corpus oraux. En pratique, les 4000 mots les plus fréquents de Frantext⁹ sont étiquetés « non saillants prosodiquement » par le modèle,

7 Dans la phonologie prosodique le mot phonologique correspond à une unité accentuable et une seule. Ainsi, la construction <adj+nom> correspond à un mot phonologique alors que <N+adj> génère deux mots phonologiques modulo les contraintes rythmiques et la règle de clash accentuel (Lacheret & Beaugendre 1999, *le chat siamois*: deux mots phonologiques, *le chat noir*: un mot phonologique).

8 Selon cette loi, la fréquence d'une unité est inversement proportionnelle à son rang dans la liste obtenue, autrement dit, le produit de la fréquence de n'importe quel mot par son rang est constant. Cette égalité, qui n'est vraie qu'en approximation, est indépendante des locuteurs, des types de textes et des langues.

9 Le nombre d'occurrences dans Frantext des 4000 mots les plus fréquents du lexique est de 50 millions selon la loi de Zipf, soit 67 % du corpus, ce qui donne un taux de 67 % de mots non saillants et 33 % de mots saillants.

les autres étant étiquetés « saillants prosodiquement ». La démarche fait également écho aux travaux récents des grammaires cognitives selon lesquelles le lexique mental des locuteurs est façonné par l'usage (voir notamment Barlow & Kemmer (éds, 2000), Bybee & Hopper (éd. 2001), Bybee (2006)). Autrement dit, les mots stockés prioritairement dans ce lexique correspondent à ceux qui sont les plus souvent utilisés dans les productions discursives et rarement saillants.

Précisons que les mots du lexique énonciatif restent étiquetés « saillants » quelle que soit leur fréquence d'usage, leur rôle étant prépondérant dans la structure informationnelle.

Tableau 1. Pour illustrer la façon dont le lexique informationnel a été conçu, les 3 colonnes du tableau présentent respectivement 10 unités parmi les plus fréquentes (faible probabilité de saillance prosodique), 10 unités aux alentours de la 4000^e position dans le classement par fréquence décroissante (unités assez fréquentes, probabilité de saillance moyenne) et 10 unités peu fréquentes (probabilité de saillance forte). Les unités de la colonne centrale se situant à la frontière entre fréquents et peu fréquents, leur classement dans l'une ou l'autre catégorie reste arbitraire. Autrement dit, la prévision d'accent pragmatique sur ces unités n'est pas fiable.

Unités fréquentes	Unités assez fréquentes, aux alentours de la 4000 ^e position	Unités peu fréquentes
<i>Vous</i>	<i>retrouvait</i>	<i>rétinien</i>
<i>Lui</i>	<i>chevelure</i>	<i>sacrificielle</i>
<i>Avec</i>	<i>adultes</i>	<i>sphéricité</i>
<i>Y</i>	<i>confondre</i>	<i>tentaculaire</i>
<i>Etait</i>	<i>blesure</i>	<i>vengeresses</i>
<i>Sa</i>	<i>ruisseau</i>	<i>anaphylaxie</i>
<i>Avait</i>	<i>diversité</i>	<i>prohibitif</i>
<i>Ou</i>	<i>tendue</i>	<i>narcotiques</i>
<i>Me</i>	<i>exercices</i>	<i>moellenses</i>
<i>Si</i>	<i>sauvé</i>	<i>intersion</i>

2.3. Etape 3: mémorisation textuelle

Pour les unités peu fréquentes, nous posons l'hypothèse que le principe majeur qui agit relève du traitement de l'information dans la progression textuelle et est lié aux processus d'enchaînements thématiques tels qu'ils sont définis dans Chafe (1987) et Combettes (1988). Concrètement, en se fondant sur l'échelle d'accessibilité des référents (Chafe 1976, 1987, Prince 1981), nous pouvons postuler qu'une unité renvoyant à un référent actif ou accessible textuellement et répondant donc à la fonction de topic n'a plus besoin de recevoir une saillance accentuelle par la suite. Autrement dit, la saillance prosodique serait inversement proportionnelle à la saillance cognitive (Lacheret 2003, Lacheret & François 2004)¹⁰. Cette hypothèse explique pourquoi dans les corpus observés les unités prosodiquement mises en relief dans leur première occurrence ne le sont généralement plus dans la suite du texte. En d'autres termes, l'information apportée par la première occurrence est ensuite thématisée (*cette mâle assurance (...) mâle assurance éditée donc chez (...)*)¹¹ Notre modèle mémorise donc les unités prosodiquement saillantes dans un texte (hormis les unités du lexique énonciatif) afin de ne pas activer la saillance de leurs occurrences subséquentes.

La question qui se pose alors est celle de la distance textuelle à partir de laquelle il ne s'agit plus d'une reprise thématique mais d'une nouvelle information. Nous avons fixé dans un premier temps cette distance à 250 mots – soit environ une minute de parole. Au vu de l'évaluation réalisée (très peu d'erreurs sur les items concernés, cf. *infra*, étape 5), cette valeur pourrait être légèrement augmentée.

10 Cette hypothèse n'inclut pas le cas des détachements topicaux dont la variété des profilages prosodiques (Avanzi & al. 2010) illustre l'insuffisance du critère d'accessibilité comme indice de classement, les paramètres sous-jacents à l'instanciation du topic étant de différents ordres, discursif, énonciatif mais aussi rythmiques.

11 Ce phénomène de thématisation ne touche pas les unités qui relèvent de la propriété commune de quantification (*supra*, 3.1.).

2.4. Etape 4: évaluation sur un corpus de 5400 mots

Afin de n'être tributaire ni d'un genre discursif particulier, ni d'un phonostyle spécifique à un locuteur, notre corpus de test a été construit autour de 6 échantillons textuels actualisés par 8 sujets (5 hommes, 3 femmes). Trois types de production ont été retenus: du texte à oraliser, de la parole professionnelle et de la parole publique. En pratique, le premier texte correspond à une séquence narrative (Harry Potter lu par B. Giraudeau, Rowling 2002, 1576 mots), le second est constitué d'un matériel expérimental centré sur l'analyse du contenu affectif des énoncés (Beaucousin 2003, 2 locuteurs, 735 mots), le troisième est une lecture d'un article politique du journal *Libération* (2005, 1 locuteur, 1112 mots). Pour la parole professionnelle, il s'agit d'un corpus prescriptif (conseils en temps réels à un skippeur, 2005, 1 locuteur, 852 mots). Quant à la parole publique, nous avons sélectionné une chronique radiophonique de B. Guetta (2003, 504 mots) ainsi qu'une interview radiodiffusée de B. Groult par R. Fayard (2002, 2 locutrices, 609 mots), cette dernière étant essentiellement constituée de séquences argumentatives¹².

La pose de balises pragmatiques a été simulée par programme, en se basant sur les deux lexiques – énonciatif et informationnel – et le principe de mémorisation textuelle, sur ce corpus de 5400 mots. Le texte du corpus a ensuite été reporté sur tableur afin d'effectuer la comparaison entre le balisage et les items effectivement saillants. Pour chaque occurrence, nous avons ainsi obtenu le nombre de proéminences prédites et réalisées, le nombre de proéminences non prédites mais réalisées et le nombre de proéminences prédites mais non réalisées (tableau 2).

12 Pour la classification en séquences discursives, voir Adam 1999.

Tableau 2: extrait de 11 lignes du tableau de bilan sur le corpus de 5400 mots, totaux (en nombre de mots) et résultat de l'évaluation (en % du nombre de proéminences)

Unités	Proéminences prédites réalisées	Proéminences non prédites réalisées	Proéminences prédites non réalisées
(...)			
<i>tout</i>	1	0	0
<i>ceci</i>	0	0	0
<i>est</i>	0	0	0
<i>rassemblé</i>	0	0	1
<i>dans</i>	0	0	0
<i>un</i>	0	0	0
<i>livre</i>	0	1	0
<i>qui</i>	0	0	0
<i>a</i>	0	0	0
<i>pour</i>	0	0	0
<i>titre</i>	0	0	0
(...)			
Totaux			
Nombre d'entrées 5388	1074	153	146
Taux	78,22 %	11,14 %	10,63 %

2.5. Etape 5: diagnostic des erreurs

Nous avons adopté la méthodologie suivante pour diagnostiquer les erreurs de balisage: chaque mot mal étiqueté du corpus d'évaluation (pour mémoire: 5400 unités) a été recherché dans la base de données Frantext ainsi que dans un corpus journalistique de 46000 items¹³. Pour les 20 premières

13 Précisons que Frantext manque de textes journalistiques, ce qui a motivé notre ajout du corpus de 46000 mots (30 extraits d'articles de Libération et du Monde dans les rubriques les plus diverses, année 2000).

occurrences trouvées¹⁴, il a été demandé à 4 locuteurs de 30 à 50 ans (2 hommes, 2 femmes, français langue maternelle) de lire chaque occurrence avec son contexte gauche et droit¹⁵. Le taux de proéminences réalisées a été noté et les corrections n'ont été validées que si elles étaient conformes à l'observation statistique: l'intégration *a posteriori* d'un item au lexique informationnel pouvait être effectuée si et seulement si une proéminence pragmatique était réalisée dans plus de 50 % des cas; pour le reste, aucune modification n'était effectuée. Au total, environ 50 % des erreurs recensées ont été corrigées.

Suite à ce premier jeu de corrections, nous avons dorénavant des items fréquents pour lesquels un accent pragmatique est malgré tout généré et, à l'inverse, des items peu fréquents pour lesquels l'absence d'accent pragmatique est prédite (voir tableaux 3 et 4 pour une illustration).

14 Le nombre de 20 échantillons correspond statistiquement à un intervalle de confiance de ± 11 % sur le taux de proéminences prosodiques mesuré.

15 Empan du contexte variable de quelques mots à quelques dizaines de mots en fonction de leur caractère nécessaire et suffisant pour la compréhension.

Tableau 3: mots fréquents mais accentués après correction. Les taux de proéminences prosodiques sont exprimés par rapport aux 20 occurrences étudiées sur corpus (cf. supra).

Mots	taux de proéminences	Trait sémantique
Bonne	90 %	Evaluation
Chaque	80 %	Récurrence
Grand	80 %	Évaluation
Première	75 %	Quantification numérique
dès lors	100 %	Quantification temporelle

Tableau 4: mots peu fréquents, mais non accentués après correction

Mots	taux de proéminences	Trait sémantique
Constitue	0 %	Verbe transitif ¹⁶
Déposer	15 %	ibid.
Etats-Unis	30 %	Cadrage thématique ¹⁷
Sondage	30 %	Thème de discours dans le contexte politique de nombreux articles
Election	20 %	ibid.

L'observation des corrections met en lumière deux phénomènes. D'une part, les items fréquents mais prosodiquement saillants sont des mots du dictionnaire énonciatif oubliés dans la version initiale, et plus précisément soit des quantifieurs, soit des marqueurs de l'expression linguistique de la subjectivité et de l'engagement énonciatif. D'autre part, les mots peu fréquents mais prosodiquement non saillants sont majoritairement des verbes transitifs (voir note 16). Cette observation confirme la mesure effectuée par Goldman et al. (2010) sur les proéminences en français: moins de 50 % des verbes de leur corpus d'étude C-PROM¹⁸ sont porteurs de proéminence –

16 L'hypothèse ici est la suivante: il s'agit à chaque fois de constructions transitives dans lesquelles l'information n'est pas portée par le prédicat mais par le complément qu'il appelle. C'est donc celui-ci qui sera prosodiquement saillant.

17 Les toponymes reçoivent rarement un accent pragmatique. Ils ont la plupart du temps une valeur de cadrage thématique dès leur première occurrence (voir la notion d'introducteur de cadre chez Charolles 1997).

18 Corpus oral de 70 mn constitué de 24 échantillons dans 7 styles différents (<http://sites.google.com/site/corpusprom>).

alors que 70 % des noms et adjectifs sont proéminents. Quant aux proéminences potentielles qui ont été neutralisées par le principe de mémorisation, elles ne représentent que 5 % des proéminences totales et n'ont été à l'origine que de 1 % des erreurs¹⁹.

2.6. Etape 6: Mise à l'épreuve du modèle sur un enregistrement PFC de 386 mots

Le corpus utilisé pour un premier bilan sur le modèle est un texte lu, extrait de la base de données *PFC phonologie du français contemporain*, projet de recherche initié en 2003 pour travailler à grande échelle sur le français oral dans ses usages attestés et dans sa variation au sein de l'espace francophone (Durand, Laks & Lyche 2002, Durand & Lyche 2003, <http://www.projet-pfc.net/>). Actuellement, la base de données dispose de 32 points d'enquêtes (ou régions de la francophonie) en ligne. Chaque enquête a été réalisée à partir d'un protocole commun soumis à des groupes d'environ une douzaine de personnes, enregistrées de façon à mettre en évidence divers paramètres de variation sociolinguistique (lecture de mots, lecture de texte, conversation guidée et conversation spontanée). En conséquence, le réservoir PFC constitue un potentiel extrêmement riche pour effectuer des analyses contrastives non seulement inter-locuteurs pour un type de discours produit mais également pour différents modes de production (texte lu vs conversation).

2.6.1. La base de données PFC comme outil d'évaluation

Le corpus sélectionné pour notre étude est un extrait de l'enquête parisienne qui porte sur 12 locuteurs appartenant tous à la haute bourgeoisie ou à l'aristocratie de la capitale. Il s'agit d'un groupe de locuteurs très homogène: ils constituent un réseau social fermé, fréquentent les mêmes milieux,

19 Le lexique informationnel a cependant ses limites: lorsqu'un mot est accentué dans 60 % des cas, peut-on estimer qu'une prosodie saillante sera recevable y compris dans les contextes inappropriés, *i.e.* qui ne convoquent pas la saillance? Autrement dit, ce type d'erreur est-il négligeable pour la perception où au contraire totalement incongru? La question reste ouverte et les expériences à construire pour y répondre.

se retrouvent dans les mêmes lieux de villégiature, les mêmes clubs et manifestent tous un très haut degré de sécurité linguistique (Lyche & Østby 2009). Il s'agit du texte lu par une locutrice de 70 ans. La taille modeste de ce texte permet de le présenter dans son intégralité et de suivre le comportement du modèle en essayant, là encore, de comprendre l'origine des erreurs rencontrées

2.6.2. *Le texte de test: proéminences prédites/proéminences réalisées*

Pour annoter les proéminences du texte, nous avons eu recours à deux outils: (1) le logiciel PRAAT (Boersma & Weenink 2009), pour le traitement du signal de parole, sa visualisation et son annotation, (2) le script easy-align (Goldman 2008) pour la transcription et la segmentation semi-automatique (alignement en phones, syllabes, mots et groupes)²⁰.

Les figures 1 et 2 ci-dessous présentent deux extraits du corpus avec sur l'axe des abscisses le temps en millisecondes et sur l'axe des ordonnées de haut en bas: (1) le signal de parole, (2) l'évolution de la ligne mélodique (ou variation de la fréquence fondamentale en repérant par des valeurs nulles les zones non voisées et les pauses, (3) la chaîne d'annotation. Pour cette dernière, nous voyons 6 tires (ou lignes d'annotation) avec de haut en bas toujours: (a) la segmentation en phonèmes, b) la segmentation en syllabes, (c) viennent ensuite les annotations des proéminences (P), enfin la segmentation en (d) mots graphémiques, (e) chaîne phonologique et (f) chaîne orthographique.

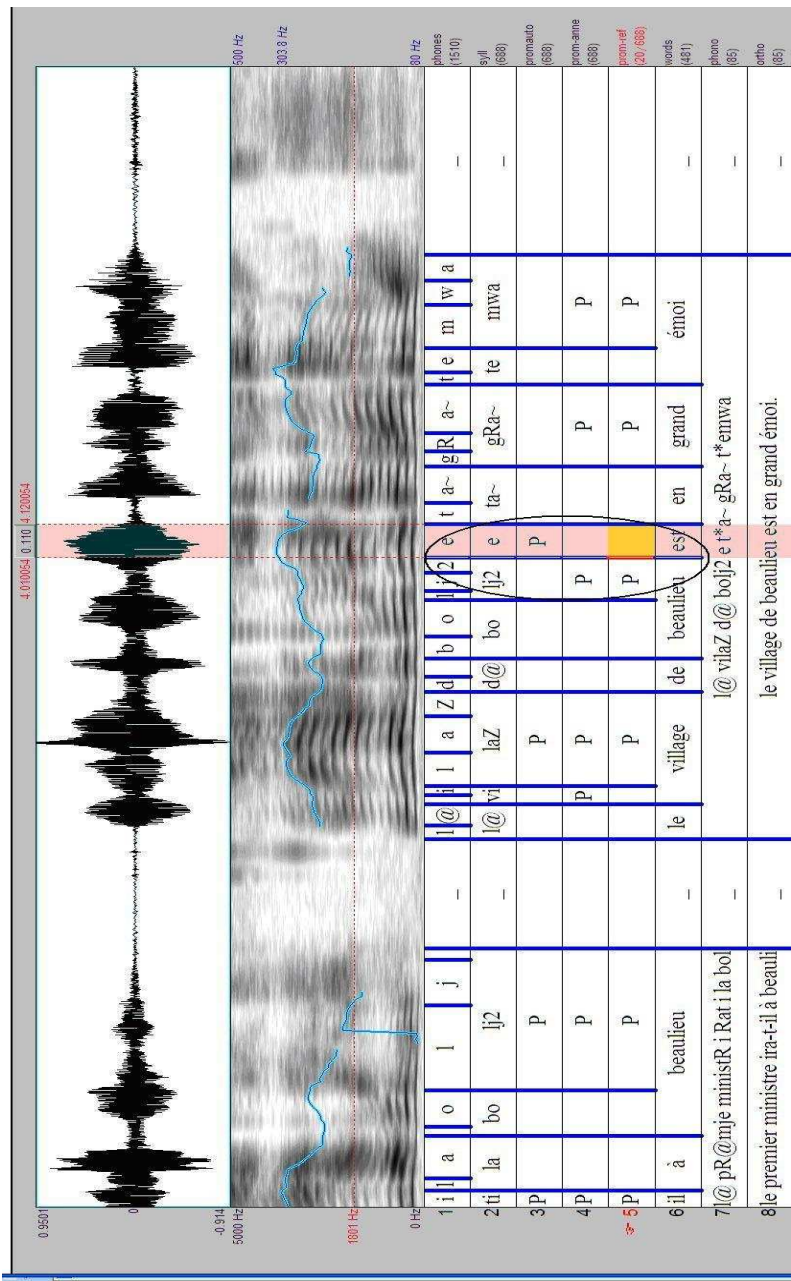
L'analyse des proéminences accentuelles du texte supposait d'abord d'effectuer une annotation de référence²¹. Pour ce faire, nous avons comparé l'étiquetage délivré par l'outil Prosoprom²² sur les seules bases d'indices acoustiques (Obin & al. 2008) et l'annotation manuelle réalisée par un an-

20 Semi-automatique car l'alignement et la transcription ont été vérifiés et nettoyés manuellement.

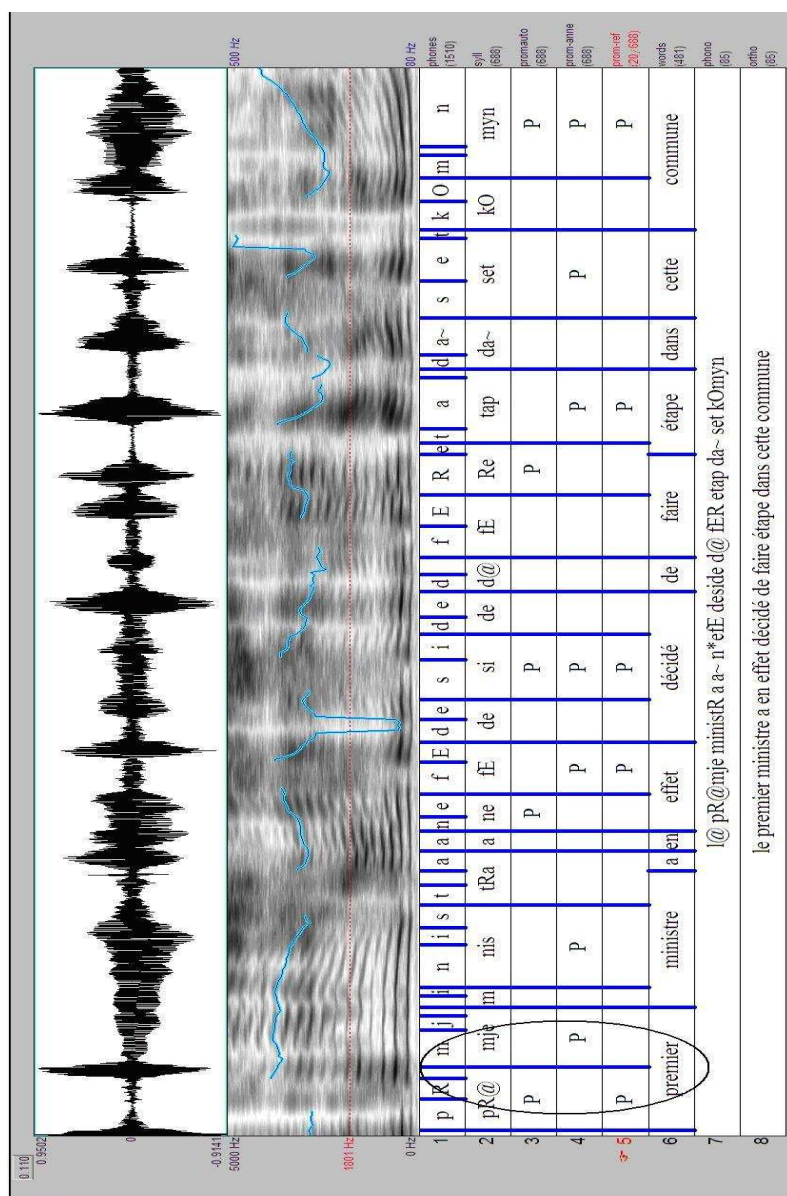
21 Nous entendons par « annotation de référence », une annotation issue de la confrontation de plusieurs annotations et d'un calcul du consensus inter-annotateur, que les annotateurs soient des humains et ou des automates (pour la nécessité de fixer une annotation de référence, voir Morel et al. 2006 et pour des travaux récents sur la mise au point d'une annotation de référence, voir Obin & al. 2008, Avanzi & al. 2010, Lacheret & al. 2010).

22 Script développé sous PRAAT par J.Ph. Goldman à l'université de Genève.

notateur humain. Ces deux annotations sont fournies respectivement dans les tires 3 et 4 de Praat, la tire 5 résulte de l'alignement manuel des deux annotations. Les proéminences entourées d'une ellipse ont été sélectionnées pour illustrer deux prises de décision différentes. Dans la première, l'annotation humaine est retenue (dans la chaîne: *le village de Beaulieu est en grand émoi*, c'est bien la syllabe terminale de *Beaulieu* qui est proéminente et non la syllabe subséquente *est*). Dans la seconde, en revanche, la proéminence affecte bien la syllabe initiale et non terminale de *premier*, c'est donc le calcul automatique qui est conservé dans l'annotation de référence.



le village de beaultieu est en grand émoi.



Figures 1 & 2: comment obtenir une annotation de référence pour l'étiquetage des proéminences accentuelles, illustration

A partir de ce premier travail, nous avons repéré en gras sur le texte les proéminences pragmatiques réalisés – à l'exclusion des proéminences purement démarcatives indiquant la finale, voire initiale d'un groupe syntagmatique (Lacheret, Lyche, Tchobanov à par.)²³ – et par une astérisque* les proéminences prédites. Le résultat est le suivant:

Le **Premier Ministre* ira-t-il à **Beaulieu*?

Le village de Beaulieu est en **grand* **émoi*. Le *Premier Ministre* a **en effet* décidé de faire étape dans **cette commune* au **cours* de sa **tournée* de la région en *fin* d'année. **Jusqu'ici* les **seuls* **titres* de gloire de Beaulieu étaient *son vin blanc sec*, ses **chemises* en soie, un **champion local* de *course* à *pied* (**Louis* **Garret*), **quatrième* aux jeux olympiques de **Berlin* en **1936*, et **plus* **récemment*, son usine de **pâtes* **italiennes*. Qu'est-ce qui a donc valu à Beaulieu ce **grand* honneur? Le **hasard*, **tout* **bêtement*, car le *Premier Ministre*, **lassé* des **circuits* **habituels* qui tournaient **toujours* autour des mêmes villes, *veut découvrir* ce qu'il appelle « la campagne **profonde* ».

Le **maire* de Beaulieu - **Marc Blanc* - est **en revanche* **très* **inquiet*. La **cote* du *Premier Ministre* ne **cesse* de **baisser* depuis les élections. **Comment* en *plus*, éviter les **manifestations* qui ont eu *tendance* à se **multiplier* lors des visites officielles? La *côte* **escarpée* du *Mont* **Saint-Pierre* qui mène au village connaît des **barrages* **chaque fois* que les **opposants* de **tous* les *bords* manifestent leur **colère*. D'un **autre* côté, à **chaque* voyage du *Premier Ministre*, le *gouvernement* prend contact avec la **préfecture* la **plus* *proche* et s'assure que **tout* est *fait* pour le **protéger*. **Or*, un **gros* **détachement* de police, comme on en a vu à **Jonquières*, et des **vérifications* **d'identité* risquent de provoquer une **explosion*. Un *jeune membre* de l'opposition **aurait* déclaré: « *Dans* le *coin*, on est **jaloux* de notre *liberté*. S'il faut *montrer* **patte blanche* pour **circuler*, nous ne **répondons pas* de la **réaction* des gens du pays. Nous avons le **soutien* du village **entier*. » **De plus*, quelques articles parus dans La **Dépêche* du Centre, **L'Express*, **Ouest Liberté* et Le **Nouvel* **Observateur* **indiqueraient* que des **activistes* des *com-*

23 Une proéminence est considérée comme pragmatique à partir de la conjugaison de critères distributionnels et acoustiques le cas échéant. En premier lieu, toute proéminence localisée sur une syllabe autre que les syllabes terminales ou initiales de groupes syntagmatiques est pourvue d'une fonction pragmatique. Ensuite, nous affectons également une valeur pragmatique aux syllabes initiales et terminales de syntagmes dont la hauteur, l'intensité ou la durée dépasse un certains seuil (plus 50 % des proéminences voisines). Pour les différents attributs phonétiques, phonologiques et fonctionnels des proéminences syntagmatiques et pragmatiques, voir Lacheret & Beaugendre 1999.

*munes voisines préparent une journée *chaude au Premier Ministre. Quelques *fanatiques auraient *même entamé un *jeûne *prolongé dans l'église de Saint *Martinville.*

*Le *sympathique maire de Beaulieu ne sait *plus à quel saint se vouer. Il a le sentiment de se trouver dans une *impasse *stupide. Il s'est, en *désespoir de *cause, décidé à écrire au Premier Ministre pour *vérifier si son village était *vraiment une étape *nécessaire dans la tournée prévue. Beaulieu préfère être *inconnue et tranquille plutôt que de se trouver au centre d'une *bataille politique dont, par la télévision, seraient *témoins des *millions d'électeurs.*

2.6.3. Les résultats et leur interprétation

Pour l'analyse des données, la méthode utilisée est identique à celle de l'étape 4. De la même façon, lorsqu'un item est prédit proéminent, c'est le groupe accentuel, ou *mot phonologique*, dans son ensemble qui est affecté par l'accent pragmatique (à *chaque voyage*).

Tableau 5: bilan du modèle sur le corpus PFC de 386 mots. Les taux sont rapportés au nombre total de proéminences.

Type de traitement	Taux de réussite de chaque traitement
Lexique énonciatif	28 %
Lexique informationnel (fréquence dans le discours)	31 %
Mémorisation	7 %
Erreurs	34 %

Le résultat global (tableau 5) peut sembler décevant: le taux d'erreurs est de 34 % alors qu'il était de 22 % dans l'évaluation étape 4. Mais il s'agit ici d'un texte lu générateur d'effets phonostylistiques incontournables²⁴. En outre, même si on observe dans l'enquête parisienne de PFC dont est extrait ce corpus une certaine évolution par rapport à la norme décrite par exemple par Fouché (1959), il s'agit là, nous l'avons vu, d'une variété conservatrice de français, situation qui pourrait expliquer en partie du moins ce décalage entre prédiction et réalisation accentuelle. Pour l'heure, la question reste

24 Nous entendons par « effets phonostylistiques » les stratégies utilisées par les locuteurs « mettre le ton ». Sans instructions précises (didascalies) sur le mode de lecture, la variation phonostylistique est incontournable.

posée mais le réservoir PFC tel qu'il a été conçu (*supra*, 2.6) doit nous permettre de valider cette hypothèse.

Venons-en au bilan détaillé des occurrences et à leur analyse pour le début du texte. Cette dernière étape nous permettra de voir dans quelle mesure les erreurs rencontrées mettent au jour des contraintes sur les constructions accentuelles de nature pragmatiques dont nous n'avions pas tenu compte jusqu'alors²⁵. L'ensemble des exemples commentés ci-dessus contient les 4 combinaisons entre prédiction (ou non) et réalisation (ou non): proéminences a) prédites réalisées, b) non prédites non réalisées, c) non prédite réalisées, d) prédites non réalisées.

- (*Le*) *Premier Ministre*: proéminence prédite sur *premier* (lexique informationnel²⁶), propagée sur le groupe accentuel (un seul mot phonologique) et réalisée sur celui-ci.
- (*ira-t-il à*) *Beaulieu*: proéminence prédite (lexique informationnel) mais non réalisée. Nous sommes là face à un cas de variation phonostylistique tel que nous l'avons exposé dans la note 24. On peut néanmoins émettre l'hypothèse que la locutrice a pris connaissance du texte préalablement à sa lecture et a considéré connue la première occurrence de *Beaulieu*.
- (*Le*) *village*: proéminence non prédite (lexique informationnel) et non réalisée.
- (*de*) *Beaulieu*: proéminence non prédite (mémorisation) et non réalisée.
- (*est en*) *grand émoi*: proéminence prédite sur les deux mots (lexique informationnel) donc sur le groupe accentuel et réalisée.
- (*Le*) *Premier Ministre*: proéminence non prédite (mémorisation de *premier*) mais réalisée. Cette deuxième occurrence de *Premier Ministre* est mise en relief alors que les suivantes ne le sont pas. C'est la seule erreur de la règle de mémorisation sur ce texte (pour 9 autres cas où la prévision a été bonne). Nous pourrions donc la considérer comme quantité négligeable mais en même temps elle nous amène à réfléchir sur le principe de mémorisation tel que nous l'avons conçu. Il paraît

25 Les unités prosodiquement vides (non proéminentes) sont indiquées entre parenthèses.

26 *Premier* pourrait être placé dans le lexique énonciatif au titre de la quantification. Mais il nous semble préférable que *Premier Ministre* soit proéminent dans sa première occurrence et thématiqué dans les occurrences suivantes.

nécessaire, dans les étapes suivantes de notre modélisation, de tenir compte des acquis en analyse du discours en considérant ce dernier non pas comme une entité homogène ou un bloc générique mais plutôt comme un dispositif qui se déploie linéairement et progressivement, et correspond de ce point de vue à une suite de séquences discursives au statut fonctionnel variable (séquence introductive, descriptive, explicative, etc.). Pour notre corpus, en l'occurrence, il s'agit de distinguer deux séquences majeures: introductives vs illustratives, la séquence introductive se limitant en fait à la première phrase du texte. Une telle approche nous permettrait alors de filtrer la règle de mémorisation qui ne pourrait s'appliquer qu'à l'intérieur d'une séquence²⁷.

- (a) *en effet décidé*: proéminence prédite sur la locution *en effet* (lexique énonciatif) et réalisée.
- (de) *faire étape*: proéminence non prédite (lexique informationnel) et non réalisée.
- (dans cette) *commune*: proéminence prédite (lexique énonciatif *cette* = déictique) et réalisée.
- (au) *cours (de)*: proéminence prédite (lexique énonciatif *au cours de*) et réalisée.
- (sa) *tournée*: proéminence prédite (lexique informationnel) et non réalisée. C'est le cas typique d'un mot dont le statut informationnel tourne autour de 50 %. La présence ou l'absence de proéminence sur ce mot ne nous semble pas contrainte.
- (de la) *région*: proéminence non prédite (lexique informationnel) et non réalisée.
- (en) *fin d'année*: proéminence non prédite mais réalisée. La valeur inchoative de la construction entraîne un accent pragmatique sur *fin* (cf. *supra*, 2.1, note 2).
- *Jusqu'ici*: (locution considérée comme un mot de 3 syllabes) proéminence prédite et réalisée.
- (les) *seuls titres*: proéminence prédite sur *seul* (lexique énonciatif) et *titres* (lexique informationnel) donc sur le groupe accentuel et réalisée.

27 Pour des réflexions sur ce concept de séquence et son utilisation en prosodie, voir aussi Lacheret 2007, Lacheret & al. à par.

Conclusion

Dans cet article, nous avons proposé un modèle de génération automatique des proéminences accentuelles de nature sémantico-pragmatique pour la synthèse de la parole à partir du texte. L'objectif était en premier lieu d'explorer leur dimension fonctionnelle, *i.e.* les contraintes sous-jacentes de nature lexicale et communicationnelles et leur portée, avant de se saisir de la dimension formelle²⁸. La modélisation formelle suppose en effet d'avoir résolu la question des contraintes fonctionnelles avant d'être entreprise. En pratique, elle repose sur l'identification des corrélats acoustiques de ces contraintes et consiste à interroger les indices dont on dispose sur le signal pour distinguer clairement ce type de proéminences, quand elles se trouvent en début ou fin de syntagme, des proéminences démarcatives dont la fonction est toute autre: segmenter le continuum sonore en une série d'unités de nature variable (syntaxique, rythmique, sémantique) et non mettre en relief une unité plus qu'une autre.

Pour effectuer cette première tâche fonctionnelle, nous avons utilisé 3 outils: deux lexiques et une règle de mémorisation. Le premier lexique rend compte des marqueurs de l'engagement énonciatif au sens large, il relève de la dimension illocutoire et expressive et également des phénomènes de repérages spatio-temporels et aspectuels. Le second tente de mieux cerner la relation implicite entre l'usage, *i.e.* la fréquence des unités dans le discours et leur statut informationnel potentiel. La règle de mémorisation enfin est utilisée pour modéliser l'activation de la saillance prosodique d'une unité en fonction de son statut textuel (plus ou moins accessible ou actif).

Ce travail et les premières évaluations auxquelles il a conduit nous amènent à préciser les étapes nécessaires pour enrichir notre modèle. En premier lieu, un troisième lexique semble s'imposer rapidement, exclusivement centré sur les constructions syntaxiques associées au vocabulaire de l'expressivité et de la subjectivité (Bally 1913, Hubler 1998, Foolen 2004). Cette nécessité soulève une question théorique de fond: comment présenter les entrées traitées, ou *squelettes lexicaux* afin de poser une règle de prédiction accentuelle nécessaire et suffisante, *i.e.* de manière à éviter une généralisa-

²⁸ Même si nous en avons dit quelques mots dans la note 23.

tion excessive (*supra*, introduction)²⁹. En second lieu, ce lexique doit pouvoir prédire les différents schémas accentuels des détachements topicaux tenant compte à la fois de l'accessibilité référentielle de ces derniers mais également de l'engagement énonciatif et des greffes expressives le cas échéant.

Concernant les corrélats formels et distributionnels de nos proéminences, nous avons posé comme première hypothèse très générale que la proéminence affecte l'ensemble d'un mot phonologique sans creuser l'hypothèse plus avant. Il semble pourtant évident qu'elle doit être modulée en fonction des règles de projection du focus sur la structure accentuelle; soit par exemple *un chat noir*, en situation de focus contrastif, l'accent sur le nom ou sur l'adjectif dépend du statut informationnel de l'un et de l'autre (Magne & al. 2005, Wagner & al. 2010).

Au sujet de l'exploitation statistique des données textuelles, le développement du lexique informationnel basé sur la fréquence souligne l'importance des textes comme observatoire et la nécessité de mettre les corpus au cœur du dispositif d'analyse, mais elle met en même temps en exergue les limites de la démarche et la nécessité d'explorer des volumes toujours plus gros de données sans pour autant être jamais sûrs d'avoir fait le tour de la question³⁰. Dans ce contexte, il est nécessaire de garder à l'esprit l'enjeu global du projet: la performance du modèle ne résulte pas de sa capacité (illusoire) à rendre compte de toutes les variations possibles des proéminences pragmatiques pour tout type de genre, de texte et de locuteur mais de son aptitude à rendre compte de l'usage moyen. Et c'est là où la tâche d'évaluation doit continuer et peut continuer grâce au réservoir PFC. Nous avons vu, en effet, que la locutrice testée représentait une variété de français relativement conservatrice et, à ce titre, sans doute pas si représentative de cet « usage moyen ». Pour s'en assurer, il nous faudra réitérer l'expérience pour le même texte lu sur d'autres points d'enquêtes de la base situés de préférence dans le centre de la France³¹.

29 Autrement dit, comment être sûr que le bon cadre syntaxique a été retenu comme entrée lexicale (ni trop spécifique, ni trop générique)

30 Voir Gadet (à par.) qui pose la question: Quelles sources pour une diversification des phénomènes linguistiques?

31 Afin d'éviter l'introduction de variables hexogènes liées à l'accent régional qui viendrait complexifier l'analyse.

Pour le dernier point, le mécanisme de gestion de la mémorisation textuelle est sans doute encore trop lâche pour donner toute la mesure de son efficacité. Une nouvelle piste de travail s'ouvre ici: l'analyse de l'organisation séquentielle des textes et leurs corrélats prosodiques, en termes de segmentation prosodique d'une part³², de remise à zéro de la règle de mémorisation d'autre part, celle-ci ne s'appliquant plus à l'initiale d'une nouvelle séquence, du moins c'est l'hypothèse que nous souhaitons sonder.

Bibliographie

- ADAM J.M. (1999): *Linguistique textuelle: des genres de discours au texte*, Paris, Nathan.
- AVANZI M., GENDROT C. & LACHERET A. (2010): « Is there a prosodic difference between left-dislocated and heavy subjects? Evidence from spontaneous French », *Speech prosody 2010*, Chicago, Illinois.
- AVANZI M., LACHERET A. & SIMON A.C. (2010): « Proceedings of Prosodic Prominence », *Speech Prosody 2010 Satellite Workshop*, Chicago, May 10th.
- BALLY C. (1913): *Le langage et la vie*, Genève, Atar.
- BARLOW M. & KEMMER S. (éds, 2000): *Usage-based Models of Language* Stanford: Center for the Study of Language and Information.
- BAUMANN S. & RIESTER A. (2010): « Annotating Information Status in Spontaneous speech », *Speech Prosody 2010*, Chicago, Illinois.
- BEAUCOUSIN V. (2003): *Les bases neurales de la compréhension des prosodies affectives*. Mémoire de DEA de biologie cellulaire, Université de Caen.
- BOERSMA P. & WEENINK D., Praat (2009): *Doing phonetics by computer* (Version 5.1). www.praat.org.
- BYBEE J. & HOPPER P. (éd. 2001): *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins.
- BYBEE J. (2006): From usage to grammar: The mind's response to repetition. *Language* 82(4), 711-733.
- CAPPEAU P. & GADET F. (2007): « L'exploitation sociolinguistique des grands corpus. Maître-mot et pierre philosophale » *Revue Française de Linguistique Appliquée*, XII-1, 99-110.
- CHAFE W. L. (1976): « Givenness, contrastiveness, definiteness, subjects, topics and point of view » in C. N. Li (ed.) *Subject and Topic*, New York, Academic Press, 25-55.
- CHAFE W. L. (1987): « Cognitive Constraints on Information Flow », in *Coherence and Grounding in Discourse*, R. Tomlin (éd.), Amsterdam, Benjamins, 21-51.
- CHAROLLES M. (1997): « L'encadrement du discours: univers, champs, domaines et espaces », Document de travail, Landisco, p. 6.

32 Voir la notion de « période » chez Lacheret et Victorri 2002, Lacheret & al. (à par.)

- COMBETTES B. (1988): *Pour une grammaire textuelle: la progression thématique*, Paris, Bruxelles, De Boeck, Duculot.
- DURAND J., LAKS B. & LYCHE C. (2002): « La phonologie du français contemporain: usages, variétés et structure », in C. Pusch & W. Raible (eds), *Romanistische Korpuslinguistik- Korpora und gesprochene Sprache/Romance Corpus Linguistics - Corpora and Spoken Language*, Tübingen: Gunter Narr Verlag, 93-106.
- DURAND J. & LYCHE C. (2003): « Le projet 'Phonologie du Français Contemporain' (PFC) et sa méthodologie. Corpus et variation en phonologie du français », in E. Delais-Roussarie and J. Durand (éds), Toulouse: PUM, 213-276.
- FOOLEN A. (2004): « Expressive binominal NPs in Germanic and Romance Languages » in *Studies in Linguistic Motivation*, éds. Radden G. and Panther K. U., Mouton de Gruyter.
- FOUCHE P. (1959): *Traité de prononciation française*, Paris, Klincksieck.
- GADET F. (à par.): *Les corpus oraux et la diversité des productions langagières*, Verbum
- GOLDMAN J.-P. (2008): « EasyAlign: a semi-automatic phonetic alignment tool under Praat », <http://latlui.unige.ch/phonetique>.
- GOLDMAN J.-P., AUCLIN A., ROEKHAUR S., SIMON A.-C., AVANZI M. (2010): « Proeminence perception and accent detection in French. A corpus based account », *Speech Prosody 2010*, Chicago, Illinois.
- HUBLER A. (1998): *The Expressivity of Grammar. Grammatical Devices Expressing Emotion Across Time (Topics in English Linguistics)*, Benjamins.
- LACHERET A. & BEAUGENDRE F. (1999): *La prosodie du français*, Paris, Editions du CNRS.
- LACHERET-DUJOUR A. & VICTORRI B. (2002): « la période intonative comme unité d'analyse pour l'étude du français parlé: modélisation prosodique et enjeux linguistiques », *Verbum*, M. Charolles (éd), Nancy, 55-72.
- LACHERET A. (2003): *La prosodie des circonstants*, Louvain, Peeters.
- LACHERET-DUJOUR A. & FRANÇOIS J. (2004): « De la notion de détachement topical à celle de constituant thématique extrapositionnel », *Cahiers de praxématique*, Montpellier, 167-198.
- LACHERET-DUJOUR A. (2007): Prosodie-discours: une interface à multiples facettes, Colloque international Interface Discours-Prosodie 2007, Genève.
- LACHERET A., OBIN N. & AVANZI M. (2010): « Design and Evaluation of Shared Prosodic Annotation for Spontaneous French Speech: From Expert Knowledge to Non-Expert Annotation », *The Fourth Linguistic Annotation Workshop (LAW IV)* Held in conjunction with ACL-2010 Uppsala, Sweden, 15-16 July 2010.
- LACHERET-DUJOUR A., AVANZI M. & VICTORRI B. (à par.): « Schématisation discursive et schématisation intonative: Question de « genre? »
- LACHERET A., LYCHE C. & TCHOBANOV A. (à par.): « Schwa et position initiale revisitée: l'éclairage de la prosodie dans les corpus PFC, Langue française », *Phonologie du français contemporain: usages, variations, structures* (dir. Laks B.)
- LEGALLOIS D. & FRANÇOIS J. (dir., 2006): « Autour des grammaires de construction et des patterns », Cahiers du CRISCO N°21, D. LeGallois & J. François dir.
- LYCHE C. & ØSTBY K. A. (2009): *Le français de la haute bourgeoisie parisienne: une variété conservatrice?*, Durand J., Laks B. & Lyche C. (éds), 203-230.

- MAGNE C., ASTESANO C., LACHERET-DUJOUR A, MOREL M., ALTER K. & BESSON M. (2005): « on-Line processing of « pop-out » words in Spoken French dialogues », *Journal of Cognitive Neurosciences*, 17:5; 740-756.
- MARTINET A. (1960): *Éléments de linguistique générale*, Paris, Armand Colin 1960, 181-187.
- MOREL M.-A. & DANON-BOILEAU L. (1998): *Grammaire de l'intonation. L'exemple du français*. Ophrys, Paris/ Gap.
- MOREL M., LACHERET A. LYCHE CH., MOREL M. & POIRE F. (2006): « vous avez dit proéminence? », *Actes des 26èmes journées d'étude sur la parole*, juin 2006.
- NESPOR & VOGEL (1983): « Prosodic Structure above the word », in A. Cuttler & D.L. Ladd, *Models and Measurements in the Study o Prosody*, Berlin, Springer, Verlag 123-140.
- OBIN N., GOLDMAN J.-P., AVANZI M. & LACHERET-DUJOUR A. (2008): « Comparaison de trois outils de détection semi-automatique des proéminences dans les corpus de français parlé », *Actes des 22èmes journées d'étude sur la parole*, Avignon.
- PERY-WOODLEY M.-P. (1995): « Quels corpus pour quels traitements automatiques? », *TAL*, 36(1-2), 213-232. *Traitements probabilistes et corpus*, Benoît Habert (resp.)
- PRINCE E. F. (1981): « Towards a taxonomy of given-new information », in P. Cole (ed), *Radical Pragmatics*, New York, Academic Press, 223-255.
- RIEGEL M., PELLAT J.CH., RIOUL RENE (1994): *Grammaire méthodique du français*, Paris, PUF.
- ROWLING J. K. (2002): *Harry Potter* lu par Bernard Giraudeau, livres audio Gallimard Jeunesse, 8 CD, 8 heures de parole.
- SELKIRK (1978): « On Prosodic Structure and its Relation to Syntactic Structure », *Nordic Prosody*, II, T. Fretheim (ed.), Trondheim, TAPIR, 111-140.
- SHANNON C. (1948): « A mathematical theory of communication », *Bell System Technical Journal*, vol. 27, 379-423 and 623-656, october 1948.
- WAGNER M., BREEN M., FLEMMING E., SHATTUCK-HUFNAGEL S. & GIBSON E. (2010): « Prosodic Effects of Discourse Salience and Association with Focus », *Speech Prosody 2010*, Chicago, Illinois.

Annexe: organigramme des différentes étapes

