

Rythme et Accentuation en Français :

Invariance et Variabilité Stylistique

Rythme et Accentuation en Français :

Invariance et Variabilité Stylistique

~

Corine ASTESANO

L'Harmattan
Edition –Diffusion
5-7 rue de l'Ecole-Polytechnique
75005 Paris

- *Introduction* -

Alors que le rythme linguistique a fait l'objet d'une littérature abondante, et continue d'occuper une place conséquente dans les recherches en phonétique, il apparaît qu'un certain nombre de questions reste en suspens. Cela vaut particulièrement pour le français qui ne bénéficie pas de la place privilégiée de l'anglais, langue abondamment décrite et à laquelle on oppose souvent le français sur le plan prosodique. Nous verrons d'ailleurs que « l'hégémonie » de la langue anglaise dans le domaine de la prosodie porte tort aux recherches sur le français et ne favorise pas l'élaboration de théories indépendantes des présupposés issus des travaux sur l'anglais.

Notre recherche a donc pour objet la description du *rythme* du français, et plus particulièrement des *phénomènes accentuels*, qui sont au cœur même du rythme linguistique. Dans cette perspective, nous proposons de mettre en évidence les *invariants* rythmiques et accentuels de ce qui pourrait constituer le système noyau du français (cf. § « Un modèle de l'accentuation du français », p. 90), ainsi que la part de *variabilité* inhérente à ce système lors de la mise en énonciation. A cette fin, nous proposons une analyse comparée de différents types de *discours*, dont on suppose qu'ils partagent un système linguistique commun, mais qui manifestent une variabilité propre aux différentes situations de communication et aux différents processus d'encodage.

Quel rythme ?

Nous définissons le rythme comme « l'organisation temporelle des proéminences qui participent de la structuration prosodique des énoncés ». Par là même, nous intégrons les notions essentielles de 'structuration', 'd'organisation temporelle' (Fraisse, 1956 ; 1974 ; Konopczynski, 1986), mais également la notion de proéminence accentuelle, dont nous pensons qu'elle est centrale dans une définition du rythme.

Le rythme est une notion qui englobe une multitude de réalités. Il se retrouve dans tous les comportements humains : linguistiques, bien sûr, mais également biologiques et artistiques. Le rythme est un principe organisateur, ordonnateur des événements. En ce sens, il n'est pas simplement le *résultat* de l'agencement des événements les uns par rapport aux autres, mais plutôt la cause, le support actif organisateur.

« Rhythmic organization is an inherent part of all human activity. Rhythm energizes, structures, creates, and expresses temporal quality. It is as important in hammering a nail or playing basketball as it is in speech, music, or dancing. Time is not merely the passive medium within which events occur. Rather, time acts to shape and determine all phenomena. Moreover, rhythmic organization exists at all levels of activity. » (Handel, 1989: 383).

Comme le propose Handel, le rythme ne se manifeste pas uniquement au niveau vocal et verbal de l'activité linguistique, mais existe à de nombreux niveaux. Ainsi, les paramètres gestuels et posturaux sont également des indicateurs d'une structuration rythmique de l'énonciation et de l'interaction entre les participants (Guaïtella, 1991 ; 1995 ; ainsi que les nombreuses publications en ce sens dans les actes du colloque « Oralité et Gestualité », Santi et al., 1998). Cette approche unifiée du rythme avait déjà été énoncée en 1926 par Bally.

« Et que serait-ce, si nous élargissions encore la notion de rythme linguistique, en constatant qu'il n'est pas autre chose qu'une mimique

vocale, qu'il n'y a pas de limite tranchée entre cette mimique et celle qui est faite de contractions faciales, de gestes des bras ou des mouvements du corps, et qu'en définitive l'organisme tout entier est intéressé dans l'acte de parole ! Comme d'ailleurs toute mimique est rythmée d'une manière ou d'une autre, le rythme linguistique se trouverait ainsi placé dans son milieu naturel. » (Bally, 1926 : 254).

Cet aspect unifié du rythme, certes fondamental pour parvenir à une compréhension globale du phénomène rythmique, ne sera pas pris en compte dans cette étude. Nous nous concentrerons ici sur l'analyse du 'rythme dans le langage ou la parole' et laisserons de côté le rythme 'hors de la parole' (Konopczynski, 1986).

Il s'agit en effet pour nous de circonscrire le cadre de notre analyse, que nous limiterons à une analyse métrique de l'organisation rythmique et accentuelle du français.

Quel cadre théorique ?

Pour ce faire, nous avons choisi le cadre théorique développé par Di Cristo & Hirst (1997) et Di Cristo (sous presse). Ces auteurs proposent un cadre accentuel pour le français dans la tradition de description métrique des faits de langue. L'approche métrique postule que le rythme, en tant qu'ossature de l'organisation prosodique, est prééminent à l'intonation (Bolinger, 1985 ; Bruce, 1985 ; Rossi, 1985b ; Di Cristo, sous presse). A travers leur modèle de l'accentuation du français, Di Cristo & Hirst mettent l'accent sur l'analyse métrique du rythme, en tant que système noyau permettant de distinguer les langues.

Les théories métriques ont à l'origine tenté d'expliquer les relations de prééminence entre les syllabes dans le cadre du mot, puis dans le cadre du groupe. Ce faisant, le rythme et l'organisation accentuelle sont plus ou moins dépendants de l'analyse syntaxique. Le modèle de Di Cristo & Hirst n'échappe pas à cette tradition descriptive, dans la lignée théorique post-généraliste (Lieberman & Prince, 1977). Les corpus qui ont permis l'élaboration d'un modèle de l'accentuation pour le français sont essentiellement des corpus de phrases écrites lues. Cependant, comme le font remarquer les auteurs, il s'agit d'une étape préalable à l'analyse de parole en situation.

« Il est à noter cependant que ces travaux ont contribué au développement d'outils d'analyse performants (Hirst & Espesser, 1993) ainsi qu'à l'élaboration d'un cadre théorique et méthodologique applicable à l'étude de la prosodie indépendamment de la variation situationnelle (Hirst & Di Cristo, sous presse) » (Astésano et al., 1995a : 3)

Les développements du cadre accentuel proposés par Di Cristo (sous presse) sont d'ailleurs le reflet de cette volonté d'ouverture sur de la parole 'réelle', en situation, dans de véritables situations de communication.

L'approche métrique se situe dans la tradition de la grammaire générative, dont le but fondamental est la recherche des universaux linguistiques. Dans cette optique, elle est fondée sur la quête des éléments abstraits constitutifs des systèmes linguistiques (structures invariantes).

Invariance et variabilité

Les notions d'invariance et de variabilité sont intimement imbriquées. La recherche ou l'observation d'une certaine invariance suppose l'existence de la variabilité des formes canoniques.

« Qui dit variation, dit invariant ; qui dit diversité, dit unité. La variation ne peut être pensée ni décrite qu'à partir d'une notion de ce qui est invariant. D'autre part, l'invariant ne saurait être établi qu'à partir d'une étude consciencieuse des faits de la variation. » (Seiler, 1997 : 116).

C'est le processus de construction du sens qui, au fur et à mesure de l'énonciation, produit une modification, une altération des unités linguistiques virtuelles.

« Cette dynamique de construction du sens dans l'énoncé implique de ne pas prendre la catégorisation linguistique comme une catégorie mentale fixe mais seulement comme un élément intervenant à un certain niveau de construction des représentations linguistiques. Entre la pensée et le langage, comme entre les unités et l'énoncé, il y a toute une dynamique de construction qui passe par des ajustements, des approximations et d'éventuels ratages. » (Robert, 1997 : 37).

Robert (1997) met ici l'accent sur les niveaux de représentation des unités linguistiques, niveaux phonologiques (unités abstraites) et phonétiques (réalisations concrètes de ces unités). C'est précisément à ce dernier niveau d'analyse que nous nous situons dans le cadre de cette étude, tout en proposant un aller retour entre les niveaux phonologique et phonétique. Il s'agit en effet de caractériser sur le plan acoustico-phonétique les catégories accentuelles fonctionnelles définies dans le cadre du modèle accentuel de Di Cristo & Hirst.

Dans cette optique, nous postulons l'existence d'un système accentuel noyau (invariant) qui, dans la 'dynamique' du langage, va se trouver modelé suivant l'intention de communication et les contraintes liées à la situation de communication.

Quel type d'analyse ?

Le modèle de l'accentuation du français, dans sa version actuelle (Di Cristo, sous presse), propose de faire le lien avec les propositions émanant des théories du discours, et plus précisément avec le courant de la linguistique pragmatique. Le modèle tente de rendre compte de la diversité de la langue dans des situations d'énonciation différentes. C'est pourquoi ce type de modèle nous paraît adapté aux objectifs de recherche que nous nous sommes fixés.

Nous proposons en effet, dans le cadre de notre étude, de confronter le modèle de l'accentuation du français à la réalité de la parole continue 'réelle', en situation. Nous avons choisi de comparer pour cela les réalisations accentuelles et rythmiques de trois types de discours, dont nous postulons qu'ils sont situés sur une échelle de 'spontanéité' : la lecture d'un texte (style *Lecture*), la lecture 'interprétée' journalistique (style *Journal*) et l'interview radiophonique (style *Interview*).

Par cette analyse comparée, nous cherchons à mettre en évidence la part de variabilité participant de la caractérisation d'un type de discours. Il faut pour cela faire la part de la

variabilité *aléatoire*, et de la variabilité *structurelle* pertinente sur le plan linguistique (selon la terminologie proposée par Rossi, 1985a). Le choix d'un cadre théorique et d'un modèle descriptif réside dans cette recherche de caractérisation de la variabilité stylistique (alloformes stylistiques), autour d'un système que l'on présuppose noyau (invariants).

Au cours de notre exposé, nous utilisons les termes de 'types de discours' et de 'styles de parole' pour désigner les réalisations orales dans différentes situations de communication. Nous privilégions néanmoins le terme 'types de discours', qui englobe la notion de 'styles de parole' mais qui présente l'avantage de ne pas mélanger les niveaux d'analyse. La notion de 'styles' revêt en effet plusieurs dimensions, sources de confusions terminologiques. Nous explicitons ce choix dans le chapitre consacré à la présentation de notre corpus (cf. § « Notions de 'styles de parole' et de 'types de discours' », p. 130).

Enfin, nous tenons à préciser que le terme '*Discours*' est pris ici dans le sens donné par Brown & Yule (1983) et Hirschberg (1993). Pour Hirschberg (1993), le 'discours' se définit comme tout énoncé en contexte. C'est également la définition que donnent Brown & Yule (1983), en insistant sur les fonctions de la communication et les situations de communication.

« The analysis of discourse is, necessarily, the analysis of language in use. As such, it cannot be restricted to the description of linguistic forms independent of the purposes or functions which those are designed to serve in human affairs. » (Brown & Yule, 1983 : 1).

C'est cette acception du discours que nous adoptons : la notion de 'contexte' et de 'situation de communication' est en effet pour nous centrale.

Nous avons choisi d'analyser la parole dans différentes situations de communication car nous pensons que l'étude de la parole en situation pousse plus loin les limites de la recherche de l'invariance. En particulier, si l'analyse de parole spontanée permet dans une certaine mesure de *tester* les modèles et théories inférés des données contrôlées de laboratoire, elle est aussi le moyen pour le chercheur d'élargir la recherche des sources de variabilité, non observables dans la parole contrôlée.

« (...) the unconstrained nature of spontaneous speech put a premium on the search for invariances. » (van Santen, 1997 : 226).

Les nouvelles possibilités techniques qui s'offrent au chercheur pour étudier la parole continue (puissance des ordinateurs et des méthodes de calcul, notamment) facilitent cette démarche et encouragent la recherche d'invariants.

En proposant d'analyser les phénomènes accentuels dans différentes situations de communication, nous tentons par là même d'établir un lien entre la théorie métrique, inspirée du courant de pensée post-généraliste (théorie Auto-Segmentale Métrique ; pour une discussion, voir Ladd, 1996), et les descriptions prosodiques issues de la théorie comportementaliste (Bolinger, 1972 ; 1985 ; 1989 ; Guaitella, 1991 et suivantes, notamment). Il s'agit pour nous de proposer une alternative descriptive dans une approche unifiée des faits de langue. En proposant l'analyse des phénomènes accentuels dans diverses situations de communication, nous nous rapprochons des théoriciens comportementalistes. En ce sens, nous envisageons le rythme de manière autonome, 'musicale' et métrique, et non pas dans sa relation ou sa corrélation à la structure

syntactique, à la différence de certains travaux (Martin, 1980 ; Mertens, 1993 ; Delais, 1994, notamment). Nous partageons néanmoins avec les tenants de l'approche métrique la conception selon laquelle il existe un système noyau de la langue, et selon laquelle la variabilité observée lors de la mise en énonciation dans diverses situations de communication n'est que l'expression d'alloformes de ce système noyau.

Notre travail se divise en deux parties principales :

– dans une première partie, nous proposons de présenter le *cadre théorique et méthodologique* de notre recherche. Cette première partie comprend 4 chapitres.

♦ Le Chapitre 1 est consacré à la présentation des notions fondamentales liées au '*Rythme*'. Nous présentons notamment les différentes théories du rythme linguistique dont nous nous inspirons. De même, nous proposons une réflexion sur la notion d'isochronie accentuelle et syllabique, dont on verra qu'elle nuit aux descriptions du rythme du français. Une alternative à la dichotomisation des langues sur la base de la notion d'isochronie est proposée, qui permet de mettre en valeur la coexistence du rythme accentuel et du rythme syllabique en français.

♦ Le Chapitre 2 est consacré à la présentation des '*Paramètres du Rythme*', parmi lesquels nous retenons l'accentuation, les pauses, le débit de parole et le timbre. Nous insistons particulièrement sur la notion d'accentuation, qui constitue le paramètre central de nos analyses empiriques. Enfin, nous situons le modèle de l'accentuation du français de Di Cristo & Hirst dans l'approche métrique, et présentons l'évolution du modèle jusqu'à sa version actuelle, sur laquelle notre recherche est fondée (Di Cristo, sous presse).

♦ Le Chapitre 3 concerne les notions de '*Variabilité et Discours*'. Nous présentons dans un premier temps les notions d'invariance et de variabilité, qui constituent les éléments de réflexion préliminaires à une analyse de la parole en situation. Nous proposons enfin de caractériser les types de discours sur le plan phonétique et prosodique, en insistant particulièrement sur le style journalistique. Ce dernier présente en effet une hétérogénéité linguistique le situant dans une position intermédiaire entre les styles *Lecture* et *Interview* sur une échelle de spontanéité.

♦ Le Chapitre 4, qui conclut cette première partie, présente nos objectifs de recherche et notre '*cadre méthodologique*'. Nous exposons les paramètres que nous avons retenus dans notre analyse de l'accentuation du français. Une grande partie de ce chapitre est également consacrée à la description de notre corpus et aux critères linguistiques à la base de la sélection des locuteurs et des types de discours.

– la deuxième partie de ce travail concerne la présentation de nos *analyses expérimentales*, et est composée de 3 chapitres.

♦ Le Chapitre 5 est consacré à l'analyse des '*durées segmentales*' du français, dans le cadre du modèle de Campbell (1992 et suivantes). Ce modèle propose de rendre compte des phénomènes d'allongement différentiels des segments dans le

cadre de la syllabe. Nous proposons de comparer 3 niveaux d'allongement sous l'effet des contraintes accentuelles et stylistiques : le niveau de la syllabe, le niveau infrasyllabique Attaque + Rime et le niveau infrasyllabique Attaque + Noyau + Coda. L'analyse de plusieurs niveaux d'actualisation des phénomènes d'allongement nous permet de rendre compte de manière optimale des stratégies stylistiques mises en œuvre dans l'organisation accentuelle en français, ainsi que du système accentuel noyau du français.

◆ Le Chapitre 6 présente l'analyse des '*configurations tonales des prééminences*'. Cette étude est plus particulièrement orientée vers la recherche de l'invariance. Dans cette optique, notre corpus multistyle se justifie dans la recherche d'invariants tonals, susceptibles de rendre compte du système noyau du rythme du français, transversal à tout style de parole.

◆ Le Chapitre 7 enfin, est consacré à l'analyse de '*l'organisation temporelle de la parole*'. Il nous semblait en effet intéressant de compléter l'analyse paradigmatique des prééminences en français par une analyse des paramètres rythmiques globaux, dont on pense qu'ils sont plus révélateurs de la variabilité stylistique (Hirschberg, sous presse). Nous montrons, à travers l'analyse de plusieurs niveaux de rythmicité et des phénomènes de pauses, que les styles de parole ne constituent pas des systèmes différents. Par contre, la variabilité stylistique observée par rapport au système rythmique noyau contribue à situer les styles de parole sur une échelle de métricit , congruente   l' chelle de spontan it .

Nous concluons enfin par une discussion g n rale o  nous argumentons sur l'int r t d'une analyse multistyle dans la compr hension et la caract risation du syst me accentuel du fran ais. Cette discussion ouvre des perspectives sur une approche perceptive de l'accentuation, qui constituerait une  tape compl mentaire incontournable dans la validation de nos hypoth ses et de nos r sultats.

- 1^{ERE} PARTIE -

Cadre Théorique et Méthodologique

<i>Chapitre 1 - le Rythme -</i>	<i>19</i>
1. Introduction	21
2. Le rythme : quelques notions fondamentales	27
3. Résumé	49
<i>Chapitre 2 - Les paramètres du rythme -</i>	<i>51</i>
1. L'accentuation	53
2. Le timbre	97
3. Le débit de parole	98
4. Les pauses	99
5. Résumé	104
<i>Chapitre 3 - Variabilité et Discours -</i>	<i>105</i>
1. Introduction	107
2. Invariance et variabilité	107
3. Caractérisation des types de discours	109
4. Résumé	122
<i>Chapitre 4 - Aspects Méthodologiques -</i>	<i>123</i>
1. Objectifs de l'étude et cadre méthodologique	125
2. Paramètres retenus pour l'analyse de l'accentuation	126
3. Protocole expérimental	129
4. Résumé	160

- *CHAPITRE 1* -

Le Rythme

I. INTRODUCTION

La notion de rythme recouvre une multitude de réalités. Il semble alors nécessaire de présenter les différentes acceptions de cette notion, et de nous situer par rapport à une définition du rythme.

1.1. Éléments de réflexion

La question fondamentale qu'il convient de se poser concerne la place du rythme dans la prosodie. Plus précisément, quelles relations le lient à l'accentuation, à l'intonation et aux autres paramètres prosodiques que sont les pauses, le timbre et le débit ? Nous verrons que le rythme, quelle que soit la réalité à laquelle il renvoie, est foncièrement ancré dans une conception duale (structuration temporelle vs. structuration accentuelle), et la place qu'il occupe dans la prosodie demeure floue. C'est pourquoi nous proposerons une définition du rythme qui nous paraît unificatrice, et qui facilite sa relation aux paramètres temporels et harmoniques (mélodie, timbre) de la prosodie.

1.1.1. *Le rythme : réalité perceptive*

Certains auteurs s'accordent à penser que le rythme appartient essentiellement au domaine de la perception. Selon Malmberg (1967) par exemple, le rythme se situe au niveau de la perception d'événements qui se répètent périodiquement selon un schéma ou un modèle donné.

« La dite réalité que nous voulons communiquer aux autres en parlant ou en écrivant [...] est un continuum indivisible qui ne se laisse soumettre à une catégorisation ou à une délimitation quelconque qu'à condition qu'un modèle, une structure, y soit appliquée. » (Malmberg, 1967 : 35)

Plus qu'une forme, une structure, le rythme serait une capacité du système cognitif à "former", à structurer l'information perçue ; plus le contexte linguistique est ambigu, plus on a tendance à projeter ce que nous percevons, en fonction de variables socioculturelles ou de variables individuelles telles que l'âge, le sexe, l'origine géographique...(Paseloup, 1990). Pour Guaïtella (1991) enfin, le rythme serait l'interprétation perceptive des indices émanant du potentiel rythmique.

1.1.2. *Différents niveaux d'actualisation du rythme*

La structuration rythmique est complexe car elle est constituée par l'interaction des niveaux lexical, syntaxique, accentuel, énonciatif et intonatif, niveaux par eux-mêmes déjà porteurs de sens et fonctionnant en relation les uns par rapport aux autres. Fraisse (1974) considère que l'organisation temporelle d'événements est hiérarchique et que, par conséquent, le rythme se situe à chacun des niveaux de la parole que sont les mores, les syllabes, les pieds, les syntagmes, les pauses... Bertinetto (1989) précise qu'au fur et à mesure des recherches, on trouve de plus en plus d'éléments composant le rythme, à des degrés différents. Il est donc clair qu'à la fois les éléments prosodiques (phonétiques et phonologiques) et les éléments segmentaux contribuent à l'organisation temporelle de la parole.

Le rythme est une organisation qui obéit à certains principes d'agencement : principes de récurrence, d'équilibrage, de symétrie-miroir, de contraste, de rupture ou de gradation (Pasdeloup, 1990). Pasdeloup ajoute cependant une dimension paralinguistique à cette définition : le rythme serait une structure qui résulte des interactions entre les niveaux linguistiques et paralinguistiques. Il existe trois composantes du rythme : une composante prosodique, une composante verbale (c'est à dire l'organisation syntaxique et lexicale) et un composante gestuelle, qui interagissent mais qui peuvent aussi intervenir indépendamment.

1.1.3. Conceptions duales du rythme

Le rôle structurel du rythme n'exclut pas une conception duale de ce phénomène. Ainsi, Allen (1975) distingue deux types de rythmes : le rythme *temporel* et le rythme *général*. Le rythme temporel est synonyme de mètre, alors que le rythme général est constitué par la répétition d'éléments atemporels comme les assonances, les répétitions etc. Il faut remarquer cependant que les phénomènes atemporels dont parle Allen sont d'autant plus prégnants sur le plan de la Gestalt (cf. § « Théorie de la Gestalt et rythme parolier », p. 31), qu'ils sont coextensifs au rythme 'temporel'. S'ils sont capables de créer effectivement une structure rythmique, l'intégration temporelle est souvent le gage d'une plus grande prégnance, selon les fondements de la psycholinguistique.

Le rythme est la structure d'une séquence, structure supposant les relations qu'entretiennent les éléments la composant (Allen, 1975). L'intérêt de cette définition est qu'elle suppose la prééminence de la *structure* sur le *temporel*. Bien que le temps soit nécessaire à la production et à la perception de la structure, c'est la structure elle-même qui génère le rythme, le temps n'étant qu'une composante de la structure. Cette définition nous satisfait pleinement en ce sens qu'elle suppose que les relations entre prééminences ainsi que leur hiérarchisation est centrale dans la notion de rythme. Elle correspond donc parfaitement à une analyse métrique du discours. Allen & Hawkins (1980) élargissent encore la notion de rythme, en postulant que tout élément (qu'il soit syntaxique, sémantique ou pragmatique) participe également à la structuration rythmique du message, si tant est qu'il participe d'une structuration.

« One can similarly argue that the organizing rhythm of a spoken utterance can be influenced by virtually any element of speech that is subject to sequential patterning. » (Allen & Hawkins, 1980 : 228).

Le rythme se déroule dans le temps et en tant que principe organisateur, il participe de l'agencement d'unités dans le temps (unités syllabiques, phonémiques, accentuelles, pausaes etc.). L'organisation temporelle des unités est également fonction du débit de parole, qui participe à la mise en valeur des unités en créant des ruptures temporelles locales.

« l'organisation temporelle des séquences (...) relève d'un principe d'équilibre sous des formes variées qui font intervenir simultanément ou alternativement, entre autres, le nombre de syllabes, les différents types de pauses, les allongements, les ruptures de tous ordres. » (Wioland, 1984 : 319-320)

Dès lors qu'on délimite des unités (forcément abstraites), la durée se voit conférer une valeur organisationnelle majeure : intrinsèquement, et à quelque niveau d'abstraction

auquel on se situe, la durée a donc un rôle ordonnateur essentiel.

L'organisation rythmique se caractériserait en effet par l'organisation *temporelle* des éléments distingués par leur accent, leur hauteur ou leur timbre (Fraïsse, 1967). La durée serait la base du rythme car elle assure l'ordre de succession. C'est ainsi que dans la musique, on observe cette prééminence de la structuration temporelle sur les autres composantes du rythme (hauteur, harmonie notamment). Un exemple frappant se trouve dans les 'Partiten' de Bach où le thème du morceau est sans cesse transposé, et où toutes les voix (parfois plus de 4 voix) reprennent dans des tonalités différentes, ou dans des mouvements mélodiques inversés, la structure du thème : la structuration temporelle joue dans ce cas un rôle prééminent dans l'identification du thème, plus que l'harmonie par exemple.

Nous avons vu à travers les définitions de Allen ou Fraïsse, que le rythme est considéré comme l'addition (plus ou moins simultanée) de deux structures principales : la structure temporelle et la structure harmonique et intensive, ou séquentielle. Cette dualité du rythme constitue une définition classique du rythme linguistique. En effet, certaines études présentent le rythme comme étant le résultat soit d'une structuration accentuelle, soit d'une structuration temporelle. Dans le premier cas, les schèmes accentuels seuls constituent la structure rythmique (Dell, 1984), alors que la deuxième acception définit le rythme comme une structuration de durées. C'est en partie ce genre de vision dichotomique du rythme qui induit la classification binaire des langues selon la prépondérance de l'une ou l'autre des caractéristiques rythmiques (cf. § « Rythmicité accentuelle vs. rythmicité syllabique », p. 32).

Une majorité d'auteurs s'accordent cependant à penser que le rythme est constitué à la fois par la structuration des schèmes accentuels et la structuration temporelle. Selon Fraïsse (1956), le rythme aurait à la base une structure temporelle sur laquelle se développeraient les variations intensives et tonales. Il ajoute, en 1974 : « *L'analyse psychologique montrera que les structures accentuelles ne peuvent être séparées des structures temporelles. Le rythme est dans le temps et il joue des durées sur le plan quantitatif et qualitatif.* »

1.1.4. Conception unificatrice du rythme

Une telle conception (Fraïsse, 1974) conserve tout de même intrinsèquement l'idée d'une dualité rythmique. Il nous paraît plus fédérateur de considérer le rythme comme « *l'organisation temporelle des proéminences* » (Di Cristo & Hirst, 1997), qui permet de dépasser la dichotomie entre 'ordre temporel' et 'ordre accentuel'. Les proéminences réfèrent à tout élément mis en valeur dans son contexte par les moyens tonals, intensifs et temporels, ou une combinaison de ces paramètres. Ainsi, structurations temporelle et accentuelle sont unifiées et peuvent être analysées comme relevant d'une même réalité sous-jacente.

Avec cette définition du rythme, on répond par là même à la question de la prééminence du rythme sur l'accentuation : le rythme est l'élément fondateur, organisateur des proéminences. En ce sens, ce ne sont pas les proéminences, les accents, qui constituent le rythme, mais c'est plutôt le rythme qui est révélateur de ces proéminences. L'accent est en effet perçu lorsqu'il tombe sur un battement (mètre) (Handel, 1989).

Dans notre acception, les phénomènes accentuels seraient simultanés aux procédés

d'organisation temporelle. Pour Bolinger par contre, ce sont deux procédés prosodiques différents. Pour Bolinger (1981, 1985) en effet, accentuation et rythme sont des phénomènes indépendants. Si l'accent est manifesté différemment suivant le locuteur, par la f0, la durée ou l'intensité, le rythme est un autre indice d'accentuation. Il postule deux niveaux de rythmicité : une rythmicité syllabique (sensible à la qualité vocalique, caractéristique harmonique d'une langue) et une rythmicité accentuelle (durée, f0, intensité). L'accentuation n'est donc pas pour lui un procédé rythmique, ce serait même l'inverse : le rythme serait un procédé accentuel.

L'approche métrique adoptée par Di Cristo & Hirst (1993 ; 1997) permet d'intégrer les notions de rythme et d'organisation accentuelle. C'est en situant le rythme dans ses fonctions essentiellement hiérarchique, relationnelle et génératrice de structures périodiques, que les auteurs dépassent la dichotomie traditionnelle. Le rythme est selon eux à la fois 'relations de prééminences entre elles', mais aussi principe générateur de structures ou de groupements plus ou moins récurrents.

1.1.5. Place du rythme dans la prosodie

Une approche métrique du rythme parolier permet également de distinguer clairement entre les composantes de la prosodie que sont le rythme et l'intonation. Ainsi, Dabrowska (1967) définit la parole comme étant un mouvement sonore ordonné, contenant deux éléments : l'intonation (le mouvement) et le rythme (l'ordonnance). Boudreault (1968) imbrique plus explicitement rythme et intonation: « *Le rythme de la phrase est constitué par le retour à intervalles réguliers des temps marqués et la mélodie, par la succession des hauteurs musicales ordonnées par le rythme.* ».

Di Cristo (sous presse) se propose d'éclaircir le lien parfois flou et peu explicite existant entre le rythme, l'intonation et la prosodie. Il définit le rythme comme « *l'organisation temporelle des syllabes métriquement fortes et des syllabes métriquement faibles* ». En intégrant le rythme à une définition métrique, l'auteur insiste sur l'indépendance entre le rythme et l'intonation, qui se définit en termes de séquences de segments tonals. Cette indépendance est relative en ce sens que le rythme est prééminent à l'intonation, les syllabes métriquement fortes constituant les points d'ancrage des segments tonals. En ce sens, toute syllabe accentuée au moyen d'un changement de hauteur participe à la fois à la manifestation du rythme et à celle de l'intonation.

Nous adoptons par conséquent la définition du rythme selon laquelle « *le rythme est l'organisation temporelle des proéminences* », car elle permet d'une part de dépasser les conceptions duales du rythme, et d'autre part de situer clairement le rythme dans la prosodie, en tant que structure noyau prééminente à l'intonation.

Nous proposons maintenant de situer l'analyse du rythme parolier par rapport aux divers courants linguistiques. Nous verrons que, suivant le point de vue que l'on adopte, les objets d'études et les analyses effectuées varient en fonction du but que l'on s'est fixé et les moyens que l'on utilise pour y parvenir.

1.2. Différentes approches de l'analyse du rythme parolier

Selon le point de vue théorique que l'on choisit de favoriser, l'analyse du rythme parolier va se trouver teintée d'a priori conceptuels. Tout est fonction de ce choix théorique

préliminaire : application d'une méthodologie (analyse déductive ou inductive), choix du type de corpus (parole de laboratoire, parole lue, parole spontanée dans différentes situations de communication) etc.

Guaïtella (1991) propose un lien très étroit entre le choix conceptuel (généralistes vs. comportementalistes) et le cadre théorique (analyse métrique vs. analyse 'rythmique'), qui détermineraient, mais seraient également déterminés par, le choix d'un corpus. Ce constat a amené cet auteur à avoir une position très tranchée, et à proposer une vision duale de l'analyse du rythme parolier.

«Les auteurs qui se sont intéressés au rythme de la parole peuvent être regroupés selon 2 tendances : la tendance issue de la grammaire générative, où les auteurs considèrent que le texte est premier dans la parole, et se reconnaissent comme partisan de l'analyse métrique, et la tendance comportementaliste où les auteurs considèrent que la réaction de l'individu à la situation de communication est le facteur premier de la parole et que le texte n'est qu'une des conséquences de cet état, ceux-ci se reconnaissent dans l'analyse « rythmique ». Les généralistes pensent qu'il existe des structures a priori dans le texte et que ces structures doivent se retrouver dans toute situation de communication. Les comportementalistes pensent que ces structures sont des mythes, véhiculés par des usages conventionnels d'analyse de l'écrit, et que la parole est régie, non pas par des structures, mais par des règles globales du comportement humain. » (Guaïtella, 1991 : 25).

Il est vrai que les théoriciens de l'analyse métrique fondent leurs recherches sur la base d'un présupposé puissant de prééminence du texte sur la parole. Le rythme de la parole serait avant tout issu de la structuration textuelle, alors que pour les comportementalistes (parmi lesquels Guaïtella inclut Goldman-Eisler, 1968 et Bolinger, 1972, 1985 etc.), rythme et texte sont co-occurents. Ainsi, dans un but de description du système prosodique d'une langue donnée, les théoriciens de l'approche métrique utilisent des phrases ad hoc, créées par eux mêmes pour les besoins de l'élaboration de leur formalisme. Il est évident qu'à ce titre, des pans entiers de la réalité linguistique sont escamotés, notamment tout ce qui touche aux aspects pragmatiques et énonciatifs. Cependant, ces approches, certes limitatives, ont permis de mettre en évidence les systèmes noyaux des langues, notion rejetée par Guaïtella, mais que nous défendons pour notre part.

Nous ne souscrivons pas à la dichotomie stricte proposée par Guaïtella (1991), non pas que nous refusions d'admettre l'existence de ces deux courants à l'évidence antagonistes, mais plutôt parce que nous croyons qu'il serait réducteur de se situer exclusivement par rapport à l'un ou à l'autre. Nous pensons qu'il est possible, dans une certaine mesure, de réconcilier et d'unifier méthodologiquement les deux courants de pensée. Certains cadres d'analyse se proposent en effet (certes encore à un stade novateur et préliminaire) d'allier des concepts issus des théories métriques et des considérations issues des courants comportementalistes et de la linguistique interactionnelle. Le cadre accentuel proposé notamment par Di Cristo (sous presse) nous semble répondre à cette volonté unificatrice.

Guaïtella (1991 ; 1996) refuse la notion de structure dans la parole, qui selon elle, est trop fortement liée à la notion de phrase, de texte écrit. La parole spontanée, de par son

caractère improvisé et l'impossibilité de procéder à des retours en arrière, ne permettrait pas l'élaboration de structures. L'approche générativiste, et de fait, l'analyse métrique, présuppose l'existence d'une structure sous-jacente, qui ne serait pas adaptée à la réalité de la parole spontanée. Elle présuppose également la détermination d'unités d'observation, de principes préétablis qui limitent et contraignent l'observation des données, au risque d'ignorer certains phénomènes, ou de les écarter purement et simplement de l'analyse, alors qu'ils constituent des phénomènes rythmiques à part entière. Un événement non attendu par exemple, se verra exclu de l'analyse s'il ne correspond pas aux formes canoniques présupposées. Également, et c'est en ce sens que les réflexions de Guaïtella (1991) nous paraissent particulièrement justifiées, des phénomènes non textuels (les composantes rythmiques para-verbales notamment) vont être systématiquement ignorés, alors qu'ils forment un tout, indissociables des composantes linguistiques dans la fonction de communication (Bailly, 1926 ; Handel, 1989, notamment).

Cependant, les prises de positions théoriques de Guaïtella concernant le rythme parolier peuvent dans une certaine mesure desservir la cause même que l'auteur défend. La dichotomie opérée par l'auteur entre l'approche générativiste et comportementaliste tend à séparer radicalement et à situer aux deux extrémités d'une échelle la parole lue et la parole spontanée. Si nous pouvons concevoir que la parole n'inclut pas uniquement la notion de texte au mépris des autres composantes rythmiques de la parole, nous ne souscrivons pas au postulat selon lequel la parole spontanée n'est pas structurée. Deux types d'encodage différents ne signifient pas pour autant deux types différents de systèmes. Si l'on allait jusqu'au bout de l'argumentation de Guaïtella, l'approche comportementaliste ne pourrait pas rendre compte des procédés rythmiques mis en œuvre dans la parole lue ! Cette incapacité à décrire la parole lue d'un point de vue comportementaliste impliquerait que l'on a affaire à un type de parole particulier, et rejoindrait paradoxalement, par un autre chemin, la notion selon laquelle la parole lue est en quelque sorte un type de parole à part, que les générativistes ont qualifié à tort de parole 'neutre'.

Nous penchons plutôt pour une approche unifiée, selon laquelle tout type de discours est structuré, bien qu'étant probablement soumis à des contraintes cognitives différentes. Tout est 'textuel' au sens large du terme, et non pas au sens de structuration de l'écrit. On ne peut pas séparer clairement les deux types de parole (lue vs. spontanée), mais on doit considérer qu'il existe un continuum sur lequel les types de discours se situent en fonction des contraintes pragmatiques, linguistiques, de processus d'encodage etc. Selon nous, une limite de la proposition de Guaïtella est que la dichotomisation des approches théoriques et méthodologiques *impose* un choix *exclusif* au linguiste : le désir d'ouverture originel peut aboutir paradoxalement à une fermeture à un certain type d'analyse et un certain type de discours ; le but inverse risque alors d'être atteint. Nous reconnaissons cependant que l'approche préconisée par Guaïtella (1991 et suivantes) a l'immense mérite d'avoir su brillamment proposer une alternative théorique et méthodologique aux approches traditionnelles du rythme parolier. Le type d'étude que nous entreprenons est d'ailleurs inspiré d'une tentative de rapprochement entre des conceptions du rythme souvent opposées (Di Cristo, sous presse), que la démarche de Guaïtella a participé à initier.

En effet, si l'on admet la possibilité d'isoler des unités dans le flot de parole, un repérage auditif des unités rythmiques permet de croiser une analyse métrique et rythmique. On s'affranchit ainsi de la prééminence du texte sur le rythme, et de la prévisibilité de

l'accent en fonction de la structure syntaxique et textuelle. Cependant, cette démarche ne nie pas l'existence du 'texte', base même de toute parole (au sens composante verbale de la parole), qu'elle soit lue ou improvisée. Cette approche permet d'une part de se situer dans le courant issu notamment de Bolinger (1972) qui propose que l'accent n'est pas prévisible (car il n'est pas issu de contraintes syntactico-morpho-phonologiques), et d'autre part de ne pas nier l'existence d'une structure rythmique sous-jacente, présente dans tout style de parole, mais dont la réalisation de surface sera justement contrainte par le processus d'encodage.

Dans cette optique, la variabilité observée, les différences structurelles, qualitatives et quantitatives observées entre la parole lue et la parole spontanée deviennent alors quantifiables. Plutôt que de se focaliser sur les spécificités de l'un et l'autre processus d'encodage (lu et spontané), et par là même de risquer de marginaliser l'étude de l'un par rapport à l'autre, nous proposons pour notre part une méthodologie unificatrice des divers types de discours. Ainsi, une analyse métrique n'exclut pas d'emblée une analyse rythmique (au sens perceptif, selon Guaitella, 1991). De même, une analyse des mécanismes de perception n'exclut pas du tout une quantification des phénomènes rythmiques. Il semble plus prometteur de tenter de trouver des éléments d'analyse communs aux deux activités afin de mettre en évidence les différences et similitudes. Bien sûr, d'un point de vue conceptuel, ce parti pris méthodologique implique l'acceptation de la notion de 'système noyau' d'une langue donnée. Nous pensons que toute langue possède des représentations rythmiques sous-jacentes, et que les différents processus d'encodage et les contraintes diverses qui les caractérisent vont se distinguer sur le plan prosodique au moment de l'actualisation à un niveau de surface des réalisations rythmiques (cf. § « Opposition entre 'mètre' et 'rythme' », p. 30).

2. LE RYTHME : QUELQUES NOTIONS FONDAMENTALES

Après avoir défini un certain nombre de termes autour de la notion de rythme, nous nous proposons de présenter les notions fondamentales liées au rythme. De même, l'évocation de quelques croyances tenaces qui ont donné lieu à une myriade de recherches nous conduira à présenter l'évolution du concept de rythme dans la description des systèmes linguistiques.

2.1. Psychologie du rythme

Le rythme est essentiel à la perception des événements temporels. Les spécialistes de psychologie expérimentale (Fraisse, 1956 et 1974 plus particulièrement) ont montré que tout phénomène se déroulant dans le temps subit une structuration rythmique imposée par l'auditeur. Cette structuration rythmique prend la forme de groupements, fondements du rythme.

2.1.1. Le rythme biologique

L'essence même du rythme a été abondamment étudiée par l'un des plus influents psycholinguistes français. Selon Fraisse (1967), tous les rythmes (biologiques etc.) sont acquis progressivement dans la petite enfance. Une fois acquis, les rythmes font partie intégrante de notre être, au point qu'on a pu constater le maintien des rythmes

nycthémeraux dans des situations extrêmes où l'homme est coupé de tout repère cosmique ou social. L'homme assimilerait donc les synchroniseurs externes, dans ce qu'on appelle communément « l'horloge biologique ». Les rythmes, le Rythme, est donc naturellement inscrit en nous, ce qui nous conduirait à pouvoir développer la capacité de créer également le rythme à travers les arts comme musique, danse, poésie etc.

Le fondement biologique du rythme a été étudié très tôt. Ainsi, Lévy (1926) montre que la perception des notions de 'rapide' et 'lent' est directement liée au rythme cardiaque moyen de l'homme (entre 60 et 80 battements par minute) : les événements d'une fréquence inférieure à 60 sont jugés lents, et ceux dont la fréquence dépasse 80 sont jugés rapides. Également, le rythme iambique serait favorisé dans la nature car il correspond au rythme respiratoire.

Le rythme aurait également des origines neurologiques, ce qui expliquerait que l'on trouve dans le proto-langage (langage émergent) d'enfants de toutes origines linguistiques, les mêmes tendances rythmiques notamment l'isochronie et l'allongement final. Le langage se construirait ensuite en fonction des contraintes linguistiques propres à la langue maternelle, où ces tendances premières prendraient alors plus ou moins d'importance en fonction du système linguistique.

« Il semblerait que le passage par une isochronie initiale, suivi par l'apparition d'un net allongement final serait commun, dans le langage émergent, à des langues de structuration rythmique totalement différente. On pourrait alors, avec Allen 1973, poser l'hypothèse d'une contrainte rythmique très généralisée due à l'existence d'une horloge neurale, au rythme régulier, contrôlant la production de la parole à sa base, mais dont le fonctionnement serait contrarié par toute une série de facteurs résultant des contraintes phonétiques, mais aussi phonologiques, lexicales, syntaxiques, prosodiques, caractéristiques de chaque langue et donc acquises. Il existerait donc une structure rythmique sous-jacente régulière dans le tout premier Proto-Langage, et ce, dans des langues à rythmicité totalement opposée telles que l'anglais et le français. » (Konopczynski, 1991 : 336).

En d'autres termes, Konopczynski se situe dans le débat traditionnel entre structures linguistiques acquises et innées, postulant qu'une certaine forme de rythmicité est innée (neuro-biologique), mais que l'acquisition du langage dans un système linguistique particulier viendrait par la suite s'imposer et modifier cette base innée. Elle se situe ainsi entre les deux principaux adversaires des fondements du langage, Chomsky et Piaget, dont les arguments respectifs sont présentés dans leur fameux débat « Théories du langage, Théories de l'apprentissage » (Piattelli-Palmarini, 1979). Elle réconcilie en quelque sorte les deux tendances (innéisme et constructivisme ou comportementalisme), puisqu'elle propose que le langage fait partie intégrante de l'humain, mais que son expérience du langage dans un contexte linguistique donné va modeler cet aspect 'neuro-biologique' du langage. La position de l'auteur par rapport aux théories de l'apprentissage est par ailleurs amplement discutée dans son ouvrage sur l'acquisition de la structuration prosodique (Konopczynski, 1986).

La perception du rythme est donc fortement influencée par notre nature biologique. Notre capacité à percevoir des éléments comme appartenant à un groupement ou bien en

tant qu'éléments contingents, dépendrait entre autres de notre tempo 'naturel et 'spontané' qui se situe aux alentours de 600 ms (1,7 battements par seconde). Au delà de ce tempo, le rythme n'est plus perçu car la succession régulière des événements (base du rythme) n'existe plus : la cohésion est détruite (Handel, 1989). Les limites inférieures et supérieures de perception des événements rythmiques ont été particulièrement décrites par Fraisse (1956, 1967 et 1974 ; Miller, 1956, notamment). De ces fondements biologiques du rythme découlent les notions psychologiques de 'périodicité' et 'structure'.

2.1.2. Périodicité et structure

Les notions de périodicité et de structure sont étroitement liées. La structure se trouve toujours coulée dans une périodicité et la périodicité est toujours organisation de structures. La capacité de l'homme à synchroniser son mouvement à un stimulus périodique est une caractéristique du comportement humain. Notre capacité de perception de deux stimulations successives, i.e. deux stimuli perçus comme distincts et se succédant, est comprise dans une marge étroite : de 0,2 secondes à 2 secondes. En deçà, les événements sont perçus de manière holistique (i.e. comme constituant un seul élément indissociable, une 'forme' dans la terminologie Gestaltiste, autrement dit comme un continuum sonore) ; au dessus de ce seuil, les événements sont perçus comme ne relevant pas d'une structure commune (Fraisse, 1967).

L'une des lois fondamentales de structuration est la notion de 'rapport' de durée : soit les durées sont égales, soit elles se trouvent dans un rapport notable (par exemple du simple au double), afin qu'il y ait harmonie entre rythme perçu et rythme moteur. Si les liens entre des événements successifs (périodicité) sont dans un rapport quelconque, alors apparaît l'arythmie.

Enfin, les structures que l'on peut facilement percevoir et reproduire ne dépassent jamais 4 ou 5 éléments. Au delà de 6 éléments, il se crée des sous-structures. Dès lors, la durée entre deux éléments semblables (créateurs de périodicité et de structure) peut être plus longue.

Le rythme semble donc avoir une nature duale, dont les deux aspects sont la *périodicité* et la *structure*. La périodicité est mise en évidence par la répétition de certains éléments (sonores), alors que la structure a trait aux éléments constitutifs de ce qui est répété périodiquement.

Comment ces principes psychologiques fondamentaux sont-ils traités dans l'activité linguistique et formalisés dans les théories du rythme ? Nous proposons, dans les paragraphes suivants, de présenter les notions d'assimilation et dissimilation directement liées aux aspects de périodicité et de structure, ainsi que les notions de 'mètre' et 'rythme' qui constituent les deux faces d'une même réalité linguistique. Enfin, nous présentons les notions développées par la théorie de la Gestalt, rendant compte des phénomènes rythmiques psychologiques, afin d'établir un lien avec le rythme parolier.

2.1.3. Loi d'assimilation et de dissimilation

Guaïtella (1991 et suivantes) fonde une grande partie de son argumentation concernant la comparaison du rythme de la parole lue et spontanée, sur les principes de perception des formes issus de la théorie de la Gestalt et développés par Fraisse (1974). Il s'agit des lois

d'assimilation et de dissimilation, toutes deux lois d'économie en perception, qui postulent que les formes vont être perçues en fonction de comparaisons, où les notions de groupements (assimilation) et événements (dissimilation) vont permettre l'identification des objets les uns par rapport aux autres. Ainsi, les éléments compris entre deux événements dissimilés (i.e. dont la substance va les distinguer des objets environnants) vont-ils être perçus comme des groupes de durée plus ou moins comparable, car ils subissent une influence assimilante induite par leur substance non marquée.

Guaïtella (1991) en déduit donc que

« le rythme se décrit alors comme une succession d'éléments distinctifs [...], et l'impression de périodicité n'en est qu'une conséquence secondaire. »
(Guaïtella, 1991 : 38).

2.1.4. Opposition entre 'mètre' et 'rythme'

Guaïtella (1991) propose que l'opposition entre régularité objective et régularité subjective se retrouve dans l'opposition proposée par Lévy (1926) entre 'mètre' et 'rythme'. Le mètre pourrait alors être défini comme une occurrence rythmique à régularité quantifiable, et serait englobé dans la notion plus large de 'rythme'.

« J'ai déjà remarqué que le rythme est essentiellement non métrique. Sa régularisation est donc une suite de l'influence du mètre. La pénétration mutuelle peut être parfaite ; c'est ainsi que les temps battus revêtent toujours, en même temps que leur caractère métrique, celui du rythme. »
(Lévy, 1926 : 80).

Guaïtella (1996) estime que de l'opposition entre 'mètre' et 'rythme', découle la distinction conceptuelle entre deux théories du rythme : le concept générativiste (métrique) et le concept comportementaliste (rythmique).

« Il est possible d'estimer que l'unique raison pour dissimiler est de mettre en relief des éléments contrastés, afin de les renforcer en tant que bornes d'unités, unités à travers lesquelles la tendance à assimiler peut être évaluée ; c'est la perspective d'une approche métrique. Inversement on peut proposer que le processus d'assimilation permet d'obtenir un fond sur lequel les éléments contrastés sont mis en relief ; c'est une perspective rythmique. »
(Guaïtella, 1996 : 14).

D'autres auteurs opposent cependant le 'mètre' et le 'rythme', non pas sur la base d'une opposition d'actualisation des structures suivant les processus d'encodage (Guaïtella, 1991), mais plutôt en fonction des représentations sous-jacentes et superficielles des structures rythmiques. Ainsi, le mètre serait la structure normative et abstraite (structure sous-jacente), et le rythme serait l'actualisation de cette représentation métrique mentale lors du passage à la performance du message (Couper-Kuhlen, 1993 ; Di Cristo, sous presse).

Cependant, Couper-Kuhlen reconnaît également, à l'instar de Guaïtella (1991), que le mètre peut être l'actualisation de la notion d'isochronie (battements réguliers), plus typique selon ces auteurs du rythme de la poésie.

Handel (1989) enfin, distingue le mètre temporel (timing) et le mètre mélodique, fondements du rythme linguistique et musical. Le rythme émergent serait le résultat de l'interaction entre ces facteurs (timing et mélodie), ainsi qu'avec les facteurs perceptifs généraux comme les groupements de tailles similaires etc.

2.1.5. *Théorie de la Gestalt et rythme parolier*

Pour rendre compte de cette théorie fondamentale dans la compréhension des phénomènes linguistiques, nous nous rapportons aux travaux de Couper-Kuhlen (1993) et Handel (1989).

La théorie de la Gestalt a été développée à l'origine pour rendre compte du traitement cognitif visuel. En l'appliquant à l'audition, la notion de déroulement dans le temps a du être rajoutée pour prendre en compte les effets de mémoire des événements précédents. Deux **propriétés principales** constituent le fondement de la Gestalt (théorie de la forme):

- la '*supersomation*', qui pose qu'une forme n'est pas simplement équivalente à la somme de ses parties ;
- la '*transposition*', qui pose qu'une forme peut être reconnue comme identique alors même qu'elle est composée de données différentes ou dans une modalité sensorielle différente ;

La Gestalt est fondée également sur une **loi fondamentale**, selon laquelle les formes visuelles ou auditives ont tendance à être bien formées (régulières, symétriques, simples ...). Il s'agit de la **loi de *Prägnanz***. La perception que nous avons de ces formes, parfois imparfaites, tend à les régulariser, à les rendre plus homogènes qu'elles ne le sont en réalité. Cette capacité de régularisation est renforcée par la nature dynamique des formes auditives (idéalisation de la reproduction dans la mémoire).

La *Prägnanz* a une influence centrifuge ou centripète sur les événements. Ainsi, les éléments appartenant à un événement sont maximalelement similaires et prévisibles, alors que des éléments appartenant à des événements différents sont maximalelement dissimilaires.

Enfin, la théorie de la Gestalt pose des **principes organisationnels** qui déterminent la notion fondamentale de *groupement* :

- la similarité : les éléments similaires par leurs attributs physiques ont tendance à être regroupés ;
- la proximité : les éléments proches dans l'espace ou dans le temps ont tendance à être regroupés. Il existe une préférence naturelle pour certains types de groupements (voir les comptines ou les rythmes musicaux de base.) ;
- la continuité : les éléments se déroulant dans la même direction ont tendance à être regroupés ;
- le 'destin' commun : les éléments qui se déplacent ensemble tendent à être regroupés ;
- la symétrie et la fermeture : les éléments qui forment des objets symétriques et enfermés tendent à être regroupés.

Appliqués à la parole, ces principes de la Gestalt supposent la définition d'éléments discrets susceptibles de constituer des formes, des groupements. Les éléments primitifs perceptivement discrets dans la parole sont les **syllabes**. Les syllabes inaccentuées constituent un 'fond' sur lequel se détachent les syllabes accentuées ('figure'). A un premier niveau d'organisation perceptive, les syllabes saillantes forment des **groupes** comprenant les syllabes alentours moins saillantes. Ces groupes de bas niveaux peuvent à leur tour être regroupés entre eux pour former des **structures** plus larges à un niveau supérieur d'organisation (principe d'organisation hiérarchique). Selon Couper-Kuhlen (1993), les structures rythmiques perceptivement isochrones ont une tendance naturelle à former des gestalts auditives, la régularité temporelle étant un principe organisationnel de base.

C'est d'ailleurs vers cette notion fondamentale de 'Régularité' que nous nous tournons maintenant, notion qui a donné lieu à un très grand nombre de travaux sur la nature formelle et les unités constitutives de la rythmicité des langues.

2.2. Rythmicité accentuelle vs. rythmicité syllabique

2.2.1. Introduction : le mythe de l'isochronie et de l'isosyllabité

Traditionnellement, l'approche métrique s'est attachée à décrire le rythme de la poésie. De là découlent un certain nombre de croyances, selon lesquelles les langues s'inscrivent dans un cadre de rythmicité binaire (régularité accentuelle vs. régularité syllabique), rythmicités dont on pourrait trouver des traces objectives à l'aide d'analyses acoustiques. Or, de nombreuses recherches acoustiques ont été entreprises dans ce sens, qui ont échoué dans leur quête d'une régularité physique. Il apparaîtrait, d'après les travaux des psycholinguistes, que la régularité rythmique n'aurait de réalité qu'au niveau perceptif. Fraisse (1956, 1974) notamment, démontre que la perception du rythme nécessite la formation subjective de groupements, qui n'ont pas de réalité quantitative. C'est également la conclusion à laquelle sont arrivés des phonéticiens comme Lehiste (1973) et Benguerel & D'Arcey (1986) parmi de nombreux autres auteurs.

En réalité, les théories métriques décrivant les systèmes accentuels des différentes langues sont également fortement influencées par la place prépondérante qu'occupe l'anglais dans les descriptions linguistiques. Ainsi, la séparation traditionnelle entre les langues dites à isochronie accentuelle et les langues dites à isochronie syllabique (initiée par Pike, 1945, et reprise par Abercrombie 1964 et de nombreux auteurs), est une conséquence directe de cette prédominance. Selon Pike et Abercrombie, pour les langues à rythmicité accentuelle comme l'anglais, les intervalles entre les pieds (groupements de syllabes contenant une syllabe accentuée et un nombre n de syllabes inaccentuées) sont de longueur à peu près égale. Cette tendance à l'isochronie accentuelle (durée uniforme entre les syllabes accentuées, indépendamment du nombre de syllabes constituant l'unité rythmique) ne peut être réalisée que si la durée de chaque syllabe n'est pas égale. Pour les langues à rythmicité syllabique comme le français, toutes les syllabes sont théoriquement de même longueur, et ont tendance à se produire à intervalles réguliers.

« French syllables tend to come at more or less evenly recurrent intervals - so that, as a result, phrases with extra syllables take proportionately more time » (Pike, 1945 : 35).

S'il est difficile de rassembler des preuves objectives de la tendance à l'isochronie, cette notion n'est pas fondamentalement remise en cause, même si l'isochronie et l'isosyllabité parfaites n'appartiennent pas à la réalité linguistique. Des études ont en effet montré que les intervalles interaccentuels en anglais varient en fonction du nombre de syllabes (Faure et al., 1980 ; van Santen, 1997) et que les durées syllabiques ne sont pas constantes en français (Wenk & Wioland, 1982 ; Nishinuma & Duez, 1987). De même, le principe strict d'isochronie accentuelle ne tient pas compte de l'allongement final de syntagme (Lehiste, 1973 ; 1983), et le principe d'isochronie syllabique ignore l'allongement de la syllabe finale du groupe rythmique en français (Wenk & Wioland, 1982 ; Fletcher, 1991).

En revanche, certains auteurs rejettent depuis une quinzaine d'années la dichotomie traditionnelle de Pike et s'accordent à penser que les rythmicités accentuelle (régularité des pieds accentuels) et syllabique (régularité des syllabes) coexistent dans les langues. Ainsi, il semblerait que la régularité temporelle doive être considérée à trois niveaux simultanés et superposés, entrant en compétition perceptivement : le niveau syllabique, le niveau du pied accentuel, et le niveau de frontière syntagmatique majeure (Fant et al., 1991a, 1991b). Ces caractéristiques prosodiques (niveaux de rythmicité) ainsi que les caractéristiques phonologiques, phonétiques, segmentales et physiologiques (Bolinger, 1981 ; Dauer, 1987 ; Fant et al, 1991a et b ; Vaissière, 1990, 1991 ; Wenk, 1983) participent à distinguer les langues, et à les situer sur une échelle représentant leur 'déviaton' par rapport à un « *prototype idéal d'isochronie pure* » (Bertinetto, 1989).

L'organisation temporelle des accents ou la mise en valeur des proéminences par le rythme, spécifique à chaque langue, serait cependant le facteur de distinction le plus important, et permettrait de fournir une catégorisation scalaire des langues sur une échelle allant de l'isochronie 'pure' à une isochronie moindre.

2.2.2. Caractéristiques prosodiques : niveaux de rythmicité syllabique et accentuelle

Le mythe de l'isochronie syllabique et accentuelle a pris appui essentiellement sur la comparaison entre l'anglais et le français. Ces langues ont longtemps été décrites comme étant prototypiques de l'une et l'autre tendances rythmiques. Or, depuis de nombreuses années, certaines études sur l'anglais et le français font peser des doutes sur une telle dichotomie. A la lumière des résultats, on a d'ailleurs bien du mal à imaginer comment de telles croyances ont pu prendre racine.

2.2.2.1. Démystification de l'isosyllabité du français

C'est à l'intérieur du pied accentuel que l'agencement des syllabes caractérise les différentes langues : les contrastes entre syllabes accentuées et inaccentuées s'expriment en termes de complexité syllabique, d'étendue de l'allongement sous l'accent, de tension articulatoire etc. En français, la résistance à la réduction articulatoire en syllabes inaccentuées réduit le contraste perceptif accentué/inaccentué (Fant et al., 1991a ; 1991b). Pourtant, les résultats de Delattre (1966b) montrent paradoxalement que le contraste de durée entre syllabes inaccentuées et accentuées est plus fort en français qu'en anglais ou en espagnol (rapport de 1 à 2). Les différences entre ces langues se caractériseraient donc plus par la qualité vocalique (timbre), que par les contrastes de durée entre syllabes accentuées et inaccentuées.

Les variations de durée ne distinguent pas seulement les syllabes inaccentuées des accentuées, mais également les syllabes inaccentuées entre elles. Il existe un phénomène d'alternance syllabique brève/longue dans les groupes rythmiques (Nishinuma & Duez, 1987), ainsi qu'une tendance à la compression des syllabes inaccentuées dans les mots lexicaux longs (Vaissière, 1977).

Si la syllabe ne semble pas être une unité d'isochronie privilégiée en français, qu'en est-il du pied accentuel ? Mais avant tout, peut-on parler de pied accentuel en français ?

Dans leur comparaison entre le français, l'anglais et le suédois, Fant et al. (1991a ; 1991b) ont montré que la rythmicité accentuelle peut également caractériser le français. Leurs résultats montrent que, dans les trois langues, les intervalles interaccentuels ont une durée comparable (env. 550 ms). De plus, les locuteurs lecteurs tendent à planifier les pauses et l'allongement prépausal associé de sorte qu'ils coïncident avec un multiple entier de la valeur moyenne des pieds accentuels, et ce dans les trois langues. Le rythme accentuel est ainsi mis en valeur par les pauses. Certains aspects de rythmicité accentuelle existent donc également en français.

En termes de pieds accentuels, Beckman (1992) note que les différences rythmiques se situent plutôt dans le nombre de syllabes inaccentuées par rapport aux syllabes accentuées : ainsi, les langues à rythmicité syllabique auraient des pieds accentuels beaucoup plus longs (de 9 à 10 syllabes) alors que le nombre maximal de syllabes dans un pied pour les langues à rythmicité accentuelle est de 5 syllabes. Outre le caractère abusif de ces allégations (la plupart des auteurs travaillant sur le français estiment que la taille du pied accentuel est en moyenne de 3 ou 4 syllabes), cette comparaison présente le désavantage de cautionner la dichotomie traditionnelle en niant toute possibilité d'eurythmie en français.

Or, Wenk & Wioland (1982) montrent que la durée moyenne des syllabes d'un mot rythmique a tendance à diminuer lorsque l'intervalle entre deux syllabes accentuées successives est trop grand, ceci pour obtenir une meilleure eurythmie. Ce serait donc là bien la preuve d'une tendance à l'isochronie des mots rythmiques pour le français.

Pourtant, des phonéticiens tels Kohler (1986) continuent de penser que les phénomènes de compression syllabique n'existent pas en français. Cela justifierait l'irrecevabilité du niveau du pied métrique en français. Il semblerait qu'il y ait ici confusion entre 'qualité vocalique' (due à la forte tension articulo-voicatoire en français) et 'compression temporelle'.

Malgré certaines réticences, tout porte donc à penser que la classification du français comme langue à rythmicité syllabique est parfaitement inadéquate et fort éloignée de la réalité phonétique et phonologique. Nous verrons d'ailleurs que ces réticences proviennent de la prédominance des descriptions rythmiques sur l'anglais (cf. § « Caractéristiques phonétiques », p. 38).

2.2.2.2. Démythification de l'isochronie accentuelle de l'anglais

Nous avons vu que, dans son acception forte, l'isochronie accentuelle suppose en anglais une forte propension à l'élasticité syllabique, afin de favoriser la récurrence régulière des accents.

Or, la durée syllabique n'est pas forcément réduite dans les pieds accentuels longs en anglais (van Santen, 1997). Selon cet auteur, s'il existe bien une tendance à ce que les

syllabes dans les mots longs soient prononcées plus rapidement que dans les mots courts, ce ne serait pas dû à un quelconque rôle compensatoire (tendance à la rythmicité objective), mais plutôt parce que ces syllabes contiennent moins d'information.

D'autre part, pour certains auteurs tels Kohler (1986) et Faure et al. (1980), l'allongement final observé de manière universelle quel que soit le type de langues serait la preuve même de l'absence d'isochronie pure dans les langues dites à rythmicité accentuelle. En effet, quel que soit le nombre de syllabes inaccentuées dans un pied, les auteurs observent que les syllabes accentuées conservent une durée relativement stable. Les phénomènes de compensation entre syllabes ne s'opéreraient donc pas entre les syllabes inaccentuées et accentuées, mais uniquement entre syllabes inaccentuées.

Dès lors se pose la question de l'unité rythmique pertinente dans la description du rythme des langues dites accentuelles. L'absence d'isochronie accentuelle physique en anglais conduit en effet certains auteurs à considérer le mot comme seule unité rythmique pertinente plutôt que le groupe accentuel (Bell-Berti, 1991 ; Nootboom, 1991 notamment). Pour d'autres auteurs, cette alternative n'est pas satisfaisante en ce sens que l'unité accentuelle impose des contraintes structurelles au mot, auquel elle est supérieure. Ainsi, les patrons rythmiques long/bref sont-ils plus cohérents à l'intérieur du groupe accentuel qu'à l'intérieur du mot (Fant, 1991). Kohler (1991) propose de réconcilier les deux courants de pensée puisqu'il estime que l'interaction de ces deux unités joue un rôle important dans la structuration rythmique de la parole.

Enfin, certains auteurs issus de la tradition comportementaliste remettent en cause purement et simplement la notion d'unité en parole spontanée. Seules les unités de discours liées à des contraintes pragmatiques (groupes de souffle, tours de parole etc.) seraient à même d'être clairement délimitées (Guaïtella, 1996).

Sans remettre en cause la pertinence de telles remarques, nous pensons que dès lors qu'une prééminence a été perçue, on peut délimiter des unités dans le flot de parole (syllabe, groupe accentuel, groupe intonatif etc.). La délimitation d'unités rythmiques nous semble même incontournable dans la perspective d'une comparaison entre les types de discours, ou entre les systèmes linguistiques. Nous proposons d'ailleurs, dans le paragraphe suivant, de présenter une argumentation en faveur de la coexistence d'un certain nombre d'unités rythmiques pertinentes dans la description des systèmes linguistiques, quels qu'ils soient.

2.2.2.3. Coexistence de tendances rythmiques duales et niveaux de rythmicité

Il apparaît que les tendances de régularisation des intervalles accentuels et syllabiques sont communes à tous les types de langues.

Les recherches sur l'acquisition du langage indiquent d'ailleurs que le phénomène d'isosyllabité observé massivement dans le proto-langage, n'est pas un phénomène réservé aux enfants de langue française, mais est également présent dans le proto-langage d'enfants anglophones, germanophones etc. (Konopczynski, 1986). Le rythme du langage enfantin va évoluer vers plus ou moins d'isosyllabité suivant la langue, mais on peut dès lors envisager que ce phénomène, à caractère universel, se retrouve dans certaines proportions dans tout type de système linguistique, y compris dans les langues dites à isochronie accentuelle.

Wenk & Wioland (1982) et Guaïtella (1986) soutiennent que les deux mécanismes antagonistes coexistent en français:

- l'isochronisme entre les syllabes accentuées, c'est à dire la tendance à réduire la durée des groupes rythmiques longs et à augmenter celle des groupes rythmiques courts, afin de préserver une certaine régularité accentuelle;
- et l'isosyllabité, c'est à dire la tendance à conserver une durée égale pour toutes les syllabes, quel que soit leur nombre dans le mot.

Pour notre part, nous avons mis en évidence la 'cohabitation' de ces différentes tendances rythmiques en français dans trois styles de parole différents (Astésano, 1998b). Nos résultats indiquent même une nette tendance de régularité au niveau du Mot Prosodique, et plus précisément du Mot Prosodique complexe comprenant un accent initial rythmique, alors que le niveau syllabique par exemple présente une dispersion plus importante (cf. § « Niveaux de rythmicité », p. 268).

Pour l'anglais, Bolinger (1981) postule l'indépendance des rythmes syllabique (organisation temporelle, structuration des durées, et sensibilité à la qualité vocalique) et accentuel (structuration des patrons de f0). Les deux types de rythme coexistent selon lui en anglais, puisqu'on trouve une trace de structuration rythmique en parole monocorde par exemple. En conséquence, les voyelles réduites en anglais constituent un système séparé, participant du rythme de l'anglais.

En 1985, Bolinger précise sa conception duale du rythme. Pour lui, accentuation et rythme sont des phénomènes indépendants. L'organisation rythmique est un autre moyen de mise en valeur des syllabes dans la phrase, en groupements. Par conséquent, le rythme et l'accentuation, indépendants l'un de l'autre, sont deux moyens prosodiques de mise en valeur des syllabes proéminentes, et d'organisation des proéminences. La qualité vocalique en anglais constitue un indice 'négatif' de l'accentuation, lorsque les voyelles sont réduites et qu'elles constituent par conséquent le 'fond' dans la théorie de la Gestalt. Bien que le rythme syllabique se soit développé de façon à soutenir ou porter les accents (en permettant les changements de f0 par l'agencement des syllabes), les deux schémas dans la parole sont indépendants. Le rythme syllabique est déterminé structurellement par l'arrangement entre les syllabes pleines et les syllabes réduites), alors que le rythme accentuel est en partie déterminé par la mise en valeur accentuelle imposée par le locuteur pour l'intérêt du sens, et doit s'y ajuster. L'isochronie peut donc intervenir à deux niveaux, ou plutôt, l'isochronie accentuelle peut être réalisée sur une structure syllabique irrégulière. Dans le domaine du rythme accentuel, Bolinger note que les accents nucléaires (ou rhématiques) sont préférentiellement portés vers la fin de l'énoncé (cf. § « L'accent nucléaire », p. 69). La nécessité d'éviter des clashes accentuels vers la fin des énoncés, où le contenu informationnel est le plus important, pousse les locuteurs à déplacer certains accents sur des syllabes secondaires, voire non accentuées. Le rythme accentuel est donc plus variable et moins prévisible que le rythme syllabique, car les contraintes d'équilibrage des pieds accentuels (dont la tendance est notamment mise en valeur par les travaux de Cutler (1980) sur les erreurs de langage) sont nombreuses. Cette tendance crée en anglais un rythme iambique, dont le domaine est l'énoncé, contraire au rythme phonologique de base de cette langue (trochaïque). C'est en ce sens que Bolinger rejette la définition stricte d'isochronie accentuelle en anglais.

La conception de Bolinger (1981, 1985) du rythme dual de l'anglais n'est pas sans rappeler non plus la définition du rythme de Fraise, où structure et périodicité représentent les deux faces du rythme.

L'idée centrale de Bolinger est que le locuteur impose un équilibre rythmique sur sa production langagière. Le locuteur est donc à même de pouvoir contrôler et choisir l'occurrence des syllabes proéminentes afin de promouvoir un certain rythme, de même qu'il est capable de choisir le mot qu'il va mettre en valeur, sans que cela soit imposé par la structure syntaxique et grammaticale de son énoncé.

L'originalité de la théorie de Bolinger réside dans le fait que le locuteur est central, indépendant et libre de gérer les propriétés rythmiques et prosodiques de sa langue maternelle. Dès lors, la dichotomisation traditionnelle des langues vole en éclats, car cette liberté rythmique du locuteur ouvre les portes sur toutes sortes de possibilités de rythmisation de la langue. Par exemple, on peut aisément imaginer que la recherche inconsciente de régularité va pouvoir se porter à tous les niveaux de l'organisation rythmique. Une telle conception permet de transcender les limites des unités métriques traditionnelles (syllabes et pieds métriques), pour chercher la rythmicité au sein des autres unités linguistiques. C'est le choix méthodologique que Couper-Kuhlen (1993) a fait lors de son étude sur la conversation en anglais.

Pour cela, elle se situe à contre-courant des théories métriques, dont elle emprunte pourtant sinon la méthodologie, du moins les constituants prosodiques de base. Elle déplore ainsi que Selkirk (1984) écarte la possibilité d'une organisation rythmique à des niveaux de constituance différents en établissant que le rythme de l'anglais ne peut se porter qu'au niveau du battement de base (mot phonologique), et le rythme du français qu'au niveau du semi-battement (syllabe).

Or, Couper-Kuhlen (1993) montre que le phénomène d'isochronie en anglais conversationnel peut se retrouver aussi bien au niveau syllabique, qu'aux niveaux du pied, du mot phonologique, du syntagme phonologique etc. Les principes de la Gestalt auditive permettent en effet de rendre compte de la régularité de formes issues de catégories prosodiques plutôt que morphosyntaxiques. Ce principe gestaltiste permet de transcender les niveaux de représentation métrique, et de se dédouaner ainsi du *texte* au profit du *message*.

Enfin, dans une séquence particulière, la rythmicité intervient à un seul niveau, i.e. qu'il ne peut y avoir de changement de niveau à l'intérieur d'une seule et même structure. Une fois qu'un niveau de rythmicité a été établi, seules les proéminences à un même niveau sont des candidats potentiels pour des battements futurs. Cela pourrait bien être, à un moment de la production, la syllabe pour l'anglais et le pied accentuel pour le français.

Dans la conversation, la limite de l'isochronie peut parfois s'étendre au delà du tour de parole. Couper-Kuhlen (1993) y voit une preuve de l'origine physiologique ou biologique du rythme, mais surtout de son origine psychologique et sociale. Ainsi, le rythme serait avant tout interactionnel.

2.2.3. *Caractéristiques phonétiques distinguant les langues*

Fant et al. (1991a et b), dans la lignée des travaux de Dauer (1983, 1987), interprètent la distinction traditionnelle entre langues à rythmicité accentuelle et langue à rythmicité syllabique par des distinctions d'ordre phonétique. Ainsi, la rythmicité accentuelle n'est pas due selon eux à une isochronie physique des intervalles interaccentuels, mais à une dominance perceptive des syllabes fortes ('heavy'), dont la succession est ressentie comme quasi périodique. Par cette analyse, ils rejoignent la conception de Lehiste (1973) selon laquelle l'isochronie n'a de réalité que dans la perception. Une langue serait donc ressentie comme étant à rythmicité syllabique lorsque les indices accentuels (incluant le contraste de complexité syllabique entre syllabes inaccentuées et accentuées, et la précision de l'articulation sous l'accent) sont réduits. En français, le rythme syllabique dominerait perceptivement, car il est moins 'masqué' par l'accentuation lexicale. La tendance à classer le français parmi les langues à rythmicité syllabique serait selon eux liée aux caractéristiques acoustico-phonétiques de l'accent en français, moins prégnantes que celles de l'accent anglais, notamment eu égard à l'établissement retardé du maximum de f0 sur les syllabes accentuées. Ce constat avait déjà été établi par Bally en 1926:

« En français, les valeurs rythmiques se sont à tel point volatilisées, et l'accent de mot est si faible, que seul le compte des syllabes s'impose impérieusement à l'oreille (...) » . (Bally, 1926 : 262).

Cette remarque est assez chargée d'a priori théoriques, et devrait sans doute être révisée à la lumière de l'évolution, aujourd'hui reconnue, de l'organisation accentuelle du français (cf. § « L'accentuation initiale : évolution du système accentuel en français », p. 80).

D'autre part, l'observation de Bally correspond à certaines propositions actuelles sur le français, que Wenk & Wioland (1982) attribuent au prisme conceptuel imposé par les descriptions prédominantes consacrées à l'accentuation de l'anglais.

« To the extent that terms such as « stress » and « accent » tend to heighten one's expectations of increased force, pitch-jump and a peculiar type of vowel reduction, those who fail to detect French accents may in some sense be victims of their own conceptual framework. » (Wenk & Wioland, 1982 : 201).

C'est dans la même optique que Dauer (1983) achève de remettre en question la dichotomisation des langues initiée par Pike dans un article où elle montre que les intervalles entre les syllabes accentuées ne sont pas plus isochrones en anglais, langue à rythmicité accentuelle, qu'en espagnol, langue à rythmicité syllabique. Selon elle, la tendance à l'isochronie est une propriété universelle des langues, et s'il existe effectivement des différences entre les langues dites à rythmicité syllabique et les langues à rythmicité accentuelle, ce sont plus des différences au niveau de la structure syllabique propre à chaque langue, des différences de réduction vocalique, de réalisation phonétique de l'accent et de l'influence qu'il exerce sur le système linguistique. En 1987, Dauer propose de systématiser ces facteurs phonologiques et propose un catalogue de traits permettant de spécifier le caractère plus ou moins 'syllabique' ou plus ou moins 'accentuel' du rythme d'une langue. Il apparaîtrait que, sur la base d'une classification en termes de traits phonologiques, le français occupe une position particulièrement non prototypique,

intermédiaire sur un continuum de rythmicité syllabique et accentuelle (Auer, 1990).

Ce point de vue est repris par Bertinetto (1989). Dans un article où il fait une très bonne synthèse des controverses soulevées par la dichotomie langues à rythmicité syllabique/ langues à rythmicité accentuelle, il propose une orientation scalaire de cette dichotomie. Selon lui, il n'existe qu'un seul type de langues au niveau rythmique mais avec des orientations polaires, c'est à dire que les différents types de langues varient en termes de déviations scalaires par rapport au "prototype idéal d'isochronie pure". Il justifie son point de vue par le fait que selon lui, en termes de contrôle central (production), il existe une base commune pour le rythme dans toutes les langues. Le rythme de la parole est contrôlé par un même phénomène de base au niveau de la production, quelle que soit la langue.

« Iso-accentual and iso-syllabic languages do not contrast because of some radically different feature they are supposed to possess (or, alternatively, not to possess), but rather because of the varying degree to which they seem to exhibit the various features, phonetical and phonological, which are supposed to contribute to orient a language more towards one or towards the other pole of rhythmical behaviour. » (Bertinetto, 1989 : 111).

Enfin, selon Bolinger (1981), la différence de rythmicité entre les langues serait à imputer essentiellement à la qualité vocalique. L'anglais présente une tendance à la réduction vocalique en syllabes inaccentuées, créant un certain type de rythme. En revanche, la qualité vocalique pleine crée en anglais une autre sorte de rythme, qui donne l'impression d'un battement syllabique. Cette sensation de battement est généralement la conséquence de la succession de monosyllabes, mais comme le fait remarquer Bolinger, ce n'est pas tant le caractère monosyllabique qui est responsable de ce type de rythme, mais la non réduction des voyelles. Il donne l'exemple de l'anglais Hawaïen où la non réduction vocalique est une caractéristique dialectale, donnant à l'anglais parlé à Hawaii une rythmicité syllabique.

Dans cette perspective, le rythme serait directement influencé et prendrait racine dans la structure segmentale et les caractéristiques phonémiques d'une langue.

2.2.4. Caractéristiques physiologiques et acoustiques de l'accentuation : incidence sur le rythme

Au vu des paragraphes précédents, on remarque que le rythme de la parole est paradoxalement loin d'être régulier, et ce quelle que soit la langue. Benguerel & D'Arcy (1986) expliquent cette distorsion rythmique dans le temps par des contraintes articulatoires et linguistiques. Ils considèrent que ce qui est régulier au niveau de pré-production d'un énoncé, subit une distorsion qui tend vers la décélération. Cette décélération est le résultat d'un allongement syllabique, marqué plus fortement en fin d'énoncé ou en fin de groupe respiratoire.

Également, l'effort articulatoire est constitutif du rythme, au même titre que la constitution phonématique des syllabes à laquelle il est étroitement lié (Llorca, 1987). Il se reflète sur le plan acoustique par des modifications caractéristiques de la courbe de fréquence fondamentale, de la courbe de pression sonore et du spectre de fréquence, ainsi que de l'allongement de la syllabe accentuée. En français, même si l'intensité reste faible sous l'accent, on le perçoit grâce à l'effort articulatoire qui se reflète clairement dans les

changements de timbre, dans l'allongement de la syllabe accentuée et dans le mouvement tonal (Fónagy, 1980).

Wenk & Wioland (1982) et Wenk (1983) proposent de distinguer les deux 'prototypes' de langues (anglais et français) non pas en termes d'isochronie ou d'isosyllabité, mais plutôt en faisant référence à la régulation des groupes rythmiques. Celle-ci s'effectuerait pour le français par le marquage droit des groupes rythmiques (langue codachrone), et en anglais par le marquage gauche des groupes rythmiques (langue capochrone). Cette constatation s'appuie sur des considérations perceptives, d'une part, et physiologiques, d'autre part.

Sur le plan perceptif, des études montrent que la position perçue de l'élément accentué varie en fonction de la nature physique de l'accent. Si l'accent est produit par une augmentation de l'intensité ou de la hauteur, il sera perçu comme débutant le groupement ; en revanche, si l'élément accentué est simplement plus long que les autres, il sera perçu comme apparaissant à la fin du groupe. Ces résultats sont valides pour des auditeurs de groupes linguistiques variés (Frasse, 1974).

Sur le plan physiologique, les caractéristiques rythmiques d'une langue peuvent s'expliquer par les contraintes articulatoires liées à l'activité langagière. Les cycles de tension et de relaxation musculaire caractéristiques du rythme de la parole déterminent la place et le type d'accent d'une langue donnée (Wenk, 1983). Cela peut se visualiser par une courbe rythmique dont la pente montante précédant l'accent demanderait une plus grande tension articulatoire que la pente descendante post-accentuelle. Il semble naturel de placer l'accent rythmique au sommet de la courbe car, quels que soient les différents paramètres qui accompagnent les syllabes accentuées, la sur-articulation qui implique un maximum d'énergie articulatoire, accompagne invariablement les syllabes accentuées (cf. Figure 1).

Par conséquent, les langues "trailer-timed" (codachrones) comme le français voient leurs syllabes inaccentuées situées au début de la tension musculaire, alors que les langues "leader-timed" (capochrones) comme l'anglais placent au contraire leurs syllabes inaccentuées dans la partie de relâchement de la tension musculaire; les syllabes accentuées des deux types de langues correspondant à la tension musculaire maximale. La place des syllabes inaccentuées des deux types de langues sur la "courbe" de tension musculaire explique le phénomène de centralisation des voyelles de syllabes inaccentuées en anglais, puisque les mouvements articulatoires sont à ce moment là moins amples que pour les voyelles accentuées. Inversement, en français, les syllabes inaccentuées sont situées en plein effort articulatoire et sont donc moins centralisées.

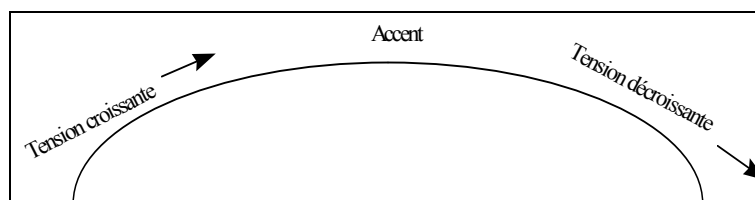


Figure 1 : Courbe rythmique (tiré de Wenk, 1983 : 157).

Les différences entre les deux langues représentant typiquement les archétypes des

langues à rythmicité accentuelle (anglais) et à rythmicité syllabique (français) s'expliquent en fait par des considérations purement phonétiques, physiologiques, perceptives etc. (cf. Tableau 1).

Tableau 1 : Principales caractéristiques des syllabes accentuées et inaccentuées en codachronie et en capochronie. (tiré de Wenk, 1983:156)

	<i>Codachronie (trailer-timing)</i>	<i>Capochronie (leader-timing)</i>
Syllabes accentuées		
1. Position	Fin de groupe	Début de groupe
2. Durée	++ allongement	± allongement *
3. Intensité	En diminution	En augmentation
4. Hauteur¹	Changements retardés	Changements brusques
5. Articulations	Explicites	Explicites
Syllabes inaccentuées		
1. Tension	Assez tendues	Assez relâchées
2. Centralisation	Voyelle peu centralisée	Voyelle fortement centralisée
* Bien que les syllabes accentuées de l'anglais aient tendance à être plus longues que les inaccentuées, il existe de nombreux cas où la durée de l'inaccentuée dépasse celle de l'accentuée (cf. Abercrombie 1964).		
¹ Nous préférons le terme « hauteur » à la désignation originelle d'« intonation » choisie par Wenk, car nous considérons l'intonation comme un système indépendant du rythme, tous deux constitutifs de la prosodie d'une langue. « Hauteur » réfère à un paramètre de l'accentuation, au même titre que « durée » et « intensité ».		

2.2.5. Conclusion sur la rythmicité des langues

Il apparaît à la lumière des paragraphes précédents que les recherches sur le rythme linguistique sont parvenues à faire voler en éclats la caractérisation binaire des langues en termes de langues à rythmicités accentuelle ou syllabique. La prise en considération d'unités descriptives dépassant le cadre métrique traditionnel (limité à la syllabe et au pied accentuel) a permis de mettre en évidence des niveaux de rythmicité multiples, communs à diverses langues. De plus, les considérations d'ordre plus général (et pas seulement rythmique ou métrique) telles que les aspects phonétiques et physiologiques dans différentes langues ont permis de mettre en lumière toute une série de composantes essentielles à la caractérisation des langues sur le plan prosodique. Se faisant, les langues se trouvent situées sur une échelle de rythmicité caractérisée par divers facteurs phonétiques et phonologiques, parmi lesquels l'isochronie accentuelle et syllabique n'est qu'une composante.

Malgré cette évolution conceptuelle du rythme linguistique, le penchant vers la dichotomisation semble tenace.

Ainsi, Fraise (1967) propose une dichotomisation non plus en termes de *structure* (syllabique ou accentuelle) mais en termes de *périodicité*. La distinction entre les langues pourrait s'opérer essentiellement en fonction de leur façon de marquer la *périodicité*. Dans certaines langues (anglais, italien ...), la périodicité est marquée par l'accentuation, alors que dans d'autres langues, elle serait marquée par les coupures temporelles. Cette proposition n'est pas sans rappeler la dichotomie proposée par Beckman (1992) et Vaissière (1991), où les langues 'accentuelles' (comme l'anglais) s'opposeraient aux langues 'de

frontière' (français), notions lourdes de présupposés théoriques que nous discuterons plus amplement ultérieurement (cf. § « Discussion sur la pertinence phonologique de l'accentuation du français », p. 75).

Les considérations d'ordre physiologique et perceptif, qui conduisent Wenk & Wioland (1982) et Wenk (1983) à catégoriser le français comme une langue codachrone (trailer-timed) et l'anglais comme une langue capochrone (leader-timed) sont également à l'origine d'une dichotomisation des langues. Leur terminologie coïncident avec celle de Hirst & Di Cristo (1984) selon laquelle les groupes rythmiques du français sont caractérisés par une structure de "tête à droite" (gabarit tonal BH), et ceux de l'anglais par une structure de "tête à gauche" (gabarit tonal HB).

Cependant, il nous semble moins dangereux d'opérer une éventuelle dichotomie entre les langues sur la base d'une qualification du pied accentuel, qui prend sa source dans la phonologie métrique. Cette caractérisation binaire n'interdit pas de considérer plusieurs niveaux de rythmicité, et la coexistence possible des différents types de rythmicité, fondés sur des considérations phonologiques ou purement segmentales et phonétiques.

C'est d'ailleurs sur la base d'une approche métrique que nous présenterons une analyse des niveaux de rythmicité en français (cf. Chapitre 7 - L'organisation temporelle de la parole - , p. 259).

L'analyse du rythme parolier suppose la définition d'une unité d'observation de base. Nous avons choisi la syllabe comme point de rencontre privilégié entre les phénomènes segmentaux et prosodiques.

2.3. La syllabe, unité rythmique

La syllabe est sans doute l'unité suprasegmentale la plus contestée et paradoxalement la plus unanimement utilisée. Son existence en tant qu'unité phonologique est notamment remise en cause par les générativistes (pour une discussion, voir Hyman, 1975). Ces derniers rejettent en fait l'idée de la syllabe en tant qu'unité de compétence, i.e. appartenant à la représentation profonde de la langue. Par contre, ils reconnaissent son existence en tant qu'unité de performance, i.e. de production et de perception des énoncés.

L'existence de la syllabe en tant que *constituant* est parfois remise en question. Pourtant, la syllabe est le domaine de nombreuses règles phonologiques et la frontière syllabique par exemple est souvent un locus où s'appliquent des règles phonologiques. D'autre part, les syllabes constituent des structures cibles pour les jeux langagiers ou les erreurs de langage. Enfin, elles sont reconnues par les intuitions linguistiques des natifs d'une langue. S'il est difficile de fournir la preuve phonologique de leur existence, les travaux en psycholinguistiques semblent unanimes sur sa pertinence en tant qu'unité d'encodage et de décodage.

2.3.1. La syllabe, une unité d'encodage et de décodage

La théorie motrice propose le phonème comme unité de perception de base. Or, différentes études ont montré les processus de compensation de durée des phonèmes entre eux à l'intérieur du cadre syllabique. Ces résultats nous conduisent à penser, à l'instar de Massaro (1975), que la syllabe constitue une unité de pré-programmation. Les

psycholinguistes apportent également des preuves expérimentales du rôle prépondérant de la syllabe dans l'accès lexical, notamment en comparant les stratégies de décodage utilisées par des locuteurs de langues différentes (Mehler & Christophe, 1992). Il apparaît que les auditeurs adoptent des stratégies de 'segmentation métrique' lors du processus de segmentation du flot continu de parole en unités lexicales (Cutler, 1992). Plus précisément, les auditeurs anglophones utilisent leurs connaissances implicites des propriétés rythmiques de leur langue (les syllabes métriquement fortes se trouvent dans une grande proportion à l'initiale des mots lexicaux) et segmentent la parole à l'attaque des syllabes fortes. Selon Cutler (1992), cette forme de segmentation métrique est hautement performante, car elle permettrait de localiser correctement plus de 90 % des mots lexicaux.

Bacri & Banel (1993) tirent des conclusions similaires pour le français. Les auditeurs francophones utilisent semble-t-il le rythme iambique propre à la langue française pour segmenter le flux de parole continue en mots.

Il semblerait donc que la syllabe, en tant qu'interface entre les niveaux segmental, lexical et rythmique, soit une unité de programmation idéale, tant du point de vue de l'encodage (performance) que du décodage.

Cependant, Pompino-Marschall (1990) propose plutôt que la perception des phénomènes rythmiques se fonde davantage sur l'occurrence « d'événements » initiés par des changements abrupts des paramètres psychoacoustiques, tels qu'ils sont produits par les modulations articulatoires du signal de parole. Si la syllabe est une unité qui dépend d'un processus cognitif d'ordre supérieur, la perception des événements temporels notamment se produirait plutôt sur des 'segments' de niveau inférieur, tels les noyaux syllabiques, les intervalles intersyllabiques ou les intervalles interaccentuels.

2.3.2. *Structure syllabique*

Blevins (1994) expose les différents modèles rendant compte de la structure interne de la syllabe. Ces théories fondent leurs hypothèses sur les contraintes phonotactiques basées sur le principe de sonorité. Depuis les années 45, les théories proposant soit des structures plates (les constituants sont les segments eux-mêmes), soit des structures à branchements binaires (Attaque + Rime {N+C}, ou Corps {A+N} + Coda), soit des structures ternaires (Attaque + Noyau + Coda), coexistent. Ce qui distingue les trois derniers modèles est l'existence de niveaux de constituance intermédiaires (Corps ou Rime), ainsi que la cohésion entre les trois constituants Attaque, Noyau et Coda.

Selon Blevins, deux arguments majeurs fondés sur les valeurs de sonorité vont dans le sens de l'existence phonologique des Attaques, Noyaux et Codas :

- tout d'abord, alors que les groupes consonantiques C initiaux et finaux dans ##C₀V₁C₀## peuvent montrer une adhérence interne rigide aux échelles de sonorité, la valeur de sonorité des consonnes Cs pré-vocaliques et post-vocaliques n'est pas déterminée en fonction de la valeur de sonorité des voyelles Vs adjacentes.

- deuxièmement, pour de nombreuses langues, les contraintes de classement de sonorité qui ont lieu à l'intérieur des séquences C pré-vocaliques ne sont pas forcément l'image miroir de celles qui contraignent les séquences C post-vocaliques.

On peut avancer un autre argument, d'ordre phonétique et perceptif, en faveur de la

constituance syllabique en Attaque, Noyau et Coda. Selon Massaro (1975), la division syllabique en Attaque, Noyau et Coda proposée dans la littérature phonologique prendrait sa source dans le concept fondamental de la voyelle comme Noyau de la syllabe, sans laquelle la syllabe n'existerait pas. Les voyelles fonctionnent en effet comme des unités perceptives de base. Les syllabes V, VC et CV sont des unités indivisibles de la perception de la parole. En effet, les voyelles isolées sont identifiables alors que les phonèmes consonantiques ne sont pas fonctionnels sur le plan perceptif à moins d'être placés en contexte vocalique.

Blevins avance ensuite une argumentation en faveur de l'existence de la Rime (N + C) comme constituant primordial de la syllabe. La preuve la plus robuste est selon elle la sensibilité au poids syllabique dans de nombreuses langues, définie sans faire référence à la portion pré-vocalique de la syllabe (également, Durand, 1990).

Certains jeux langagiers utilisent la rime comme support de transformation ; Blevins donne ainsi l'exemple du jeu 'oppen-gloppen' qui, en anglais transforme la phrase 'you are mad' en [y-op-u op-aœ m-op-aed]. Les substitutions respectent la construction de la Rime, alors que l'Attaque est touchée (également, Durand, 1990).

Enfin, la difficulté à rendre compte en phonologie de l'existence de l'Attaque et de la Coda comme constituants syllabiques à part entière, conduit Blevins à proposer un modèle de structure syllabique où les éléments d'Attaque sont immédiatement dominés par σ , et les éléments de Coda sont ceux immédiatement dominés par la Rime (cf. Figure 2).

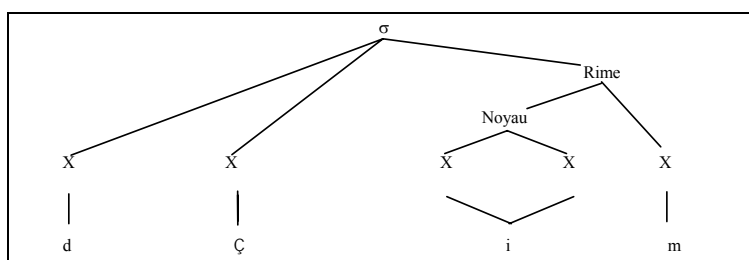


Figure 2 : Structure interne de la syllabe, tiré de J. Blevins, 1994, p. 216.

Nous avons choisi pour notre part d'adopter la structure syllabique interne en Attaque, Noyau et Coda (cf. Figure 13, p. 153), afin de rendre compte pour le français des phénomènes d'allongement différentiel affectant les syllabes différemment selon leurs fonctions et leurs places dans l'énoncé, comme l'ont montré avant nous Fant et al (1991a et b), et Campbell (1992 et suivantes) pour d'autres langues. Il semblerait apparemment que, s'il est difficile de fournir la preuve de l'existence phonologique des constituants Attaques et Codas sur la base des critères de sonorités, ces deux constituants syllabiques soient le lieu de règles d'ajustement temporel dans le cadre de la syllabe, notamment dans la distinction qui s'opère entre proéminences initiales et finales.

L'agencement temporel des syllabes constitue la base du rythme. Afin de mieux cerner le rythme, phénomène biologique universel, nous proposons maintenant d'envisager une comparaison entre le rythme parolier et musical.

2.4. La performance rythmique comme expression de l'esthétique humaine

La performance rythmique est essentiellement caractérisée par la tendance à la distorsion temporelle par rapport au 'mètre'. Lors du passage de la représentation sous-jacente à la performance, toutes sortes de distorsions temporelles s'opèrent, qui ont trait aux contraintes de production, à l'état du locuteur, à la situation de communication etc. On peut également imaginer que ces altérations temporelles ne sont pas uniquement issues de contraintes, mais qu'elles ont une fonction quelconque dans la communication. La tendance à la distorsion temporelle semble caractériser la plupart des activités humaines, de la parole aux activités artistiques. On peut donc supposer un parallèle entre les performances linguistique et musicale par exemple, où le rythme de surface représenterait l'écart par rapport au mètre, permettant de situer le locuteur/musicien dans son individualité et dans la situation de communication/expression.

On a vu que le rythme du français se caractérise par une alternance rythmique aussi bien entre les syllabes accentuées et inaccentuées, qu'entre les syllabes inaccentuées elles-mêmes (Touati, 1987). Ainsi, Duez & Nishinuma (1985) ont-ils montré que les syllabes inaccentuées ne sont pas de durée équivalente en français, mais tendent à produire un patron rythmique d'alternance long/bref. Cependant, cette tendance d'alternance est inférieure au seuil différentiel de durée tel qu'il a été déterminé par Rossi (1972). En effet, une étude perceptive indique que ces variations binaires long/bref ne relèvent que du domaine de la production, et ne sont pas perceptibles (Nishinuma & Duez, 1987). En fait, la stabilité durationnelle des syllabes inaccentuées constatée au niveau psychoacoustique constitue une certaine isochronie atone, et tend à conférer à la syllabe inaccentuée un statut d'unité rythmique de base en français parlé (Duez & Nishinuma, 1985 ; Nishinuma & Duez, 1987).

Ce phénomène d'alternance long/bref est également observé en musique. Gabrielsson (1982) montre que le rythme produit en musique subit une variation systématique par rapport à la norme écrite. Dans la notation musicale conventionnelle, la durée de chaque événement sonore est donnée par la valeur de la note (noire, croche, demi-croche etc.) ainsi que par le tempo. Les événements rythmiques vont ensuite se dérouler dans un cadre rythmique donné, en fonction de la mesure initiale du morceau (3/4, 6/8 etc.). Or, Gabrielsson remarque que, lors de la performance musicale, une certaine déviation plus ou moins systématique par rapport à ces valeurs données, s'opère. Les notes intrinsèquement courtes (doubles croches) vont être allongées, et les notes intrinsèquement plus longues (noires, croches) vont se trouver raccourcies. L'archétype rythmique d'une performance musicale va donc osciller autour des valeurs données, dans un schéma rythmique 'en dents de scie'.

On pourrait faire un parallèle ici entre le rythme musical et le rythme parolier. Lors de la mise en parole du message, il est possible que la représentation mentale 'métrique' du rythme sous-jacent, par définition figée et mécanique, subisse des variations systématiques, à l'instar de la performance rythmique musicale. Dans une analogie avec la musique, la situation de communication (conversation, interview etc., fonction également de l'enjeu communicatif, des interactants, et des relations entre les interactants etc.) déterminerait la mesure et la tonalité générales du message (représentation sous-jacente, métrique). Les déviations systématiques par rapport à cette représentation mentale interviendraient ensuite lors de la performance du message, pour des raisons purement rythmiques (alternance des

longues et des brèves autour d'une 'valeur moyenne donnée', équilibrage rythmique), mais également en fonction de l'état émotionnel du locuteur, de la portée informative du message (mise en valeur pragmatique de certains éléments locaux de son message), et de l'évolution de l'interaction.

Les variations systématiques observées lors du passage à la performance musicale et langagière s'expliqueraient donc non pas comme une simple contrainte physiologique (Nishinuma & Duez, 1987), mais comme le résultat de contraintes rythmiques ('rythme' dans le sens de 'performance'), pragmatiques, individuelles et inter-individuelles, situationnelles etc. Toutes ces variations rythmiques observées par rapport aux représentations métriques sous-jacentes ne sont peut-être pas identifiables (perceptibles) en tant que telles, mais sont néanmoins nécessaires à la communication. Elles ont également pour fonction de situer le message, quel qu'il soit (musical ou langagier) dans le domaine de l'expressivité, propre à toute communication ; elles sont même le résultat de cette expressivité. Il a en effet été montré par le recours à la modélisation et à la synthèse, aussi bien en musique qu'en parole, que le respect strict de l'organisation métrique induit un manque de naturel et nuit à la communication. Si les variations rythmiques locales ne sont pas forcément perçues en tant que telles, elles sont pourtant une condition formelle sine qua non de l'optimalisation de la communication.

Ce constat va dans le sens des recherches en psycholinguistiques (notamment Fraise, 1956 et 1974, mais aussi Benguerel & D'Arcy, 1986), selon lesquelles les distorsions temporelles observées lors de la performance linguistique ou non linguistique dues à des contraintes physiologiques et environnementales, seraient 'réparées' au niveau perceptif par la régularisation des déviations rythmiques, imposée par l'auditeur. L'auditeur s'est-il adapté aux contraintes physiologiques et linguistiques du locuteur, ou le locuteur tient-il compte de cette tendance perceptive à imposer une régularisation et structuration aux événements rythmiques ? En d'autres termes, les variations systématiques observées dans le domaine temporel sont-elles contraintes par l'étape de production ou de perception ?

Il semblerait en fait que les productions des locuteurs respectent inconsciemment les préférences perceptives des auditeurs. C'est ce que propose Fraise, notamment.

En effet, une des particularités du français selon Fraise (1967) tient dans la tendance pour cette langue à créer des structures qui rompent avec les automatismes périodiques, notamment en introduisant des structures impaires (le 'jeu des impairs'). Cette tendance se refléterait, dans la structure de surface, par le déplacement d'accent, la syncope etc.

Dans sa dichotomie entre 'structure' et 'périodicité', où la première est en quelque sorte l'expression 'rythmique' de la seconde ('métrique'), Fraise (1967) situe ainsi la particularité du rythme humain.

« Lorsque la structure l'emporte sur le périodique, qu'il s'agisse de poésie ou de musique, l'homme se trouve emporté par le mouvement rythmique qui l'oblige à rompre avec la monotonie périodique. Cet aspect des arts du rythme nous fournit alors des satisfactions plus intellectualisées et plus spiritualisées. Ces formes structurées, ces ruptures avec la régularité sont véritablement une conquête de l'homme et le sommet de ce que l'on peut appeler les rythmes humains. » (Fraise, 1967 : 33).

Les variations rythmiques systématiques seraient donc liées à une inclination de l'être humain pour une certaine esthétique, dont les fondements doivent probablement être cherchés dans le rythme biologique.

2.5. Perception du rythme

La perception est un processus de schématisation du signal acoustique (Fraïsse, 1956). La perception du rythme peut être influencée par deux facteurs: la tendance, du côté du locuteur, à augmenter la durée syllabique au cours de l'énoncé (tendance naturelle au ralentissement du débit vers la fin des énoncés) ; et la tendance, pour l'auditeur, à plaquer une certaine rythmicité sur les séquences de parole. (Benguerel & D'Arcy, 1986).

Fraïsse (1956) a d'ailleurs montré que l'accentuation objective ou subjective entraîne toujours une surestimation de l'intervalle qui la suit. Il y aurait donc non seulement distorsion de la rythmicité au moment de la production de l'énoncé, mais aussi lors du décodage. Il faut aussi être conscient des variations inter-individuelles dans la perception de la structuration rythmique, variations dépendant de facteurs socioculturels, de l'âge, du sexe et des origines géographiques de l'auditeur, tout ceci constituant une série de filtres qui influencent différemment la perception (Paseloup, 1990).

Il a été montré lors de nombreuses expériences perceptives que la régularité des phénomènes accentuels est une illusion perceptive. La régularité existe bien, mais à un niveau profond, au niveau de la programmation du rythme; mais elle disparaît au moment de la production (performance), à cause de contraintes sur la durée des phonèmes et des syllabes, imposées par les mécanismes de production de la parole. Le système de perception de la parole "corrige" les distorsions de la performance et restaure la régularité sous-jacente (Lehiste, 1973). Scott et al. (1985) précisent cependant que le phénomène de régularisation perceptive n'est pas la preuve de l'isochronie sous-jacente d'une langue. Ils ajoutent que l'illusion de régularité n'est pas spécifique aux langues dites à rythmicité accentuelle.

En effet, la régularisation *perceptive* des intervalles interaccentuels est aussi naturelle en français qu'en anglais par exemple (Scott et al., 1985). Les auteurs remarquent également, à l'instar de Lehiste (1973), que la tendance à la régularisation des intervalles interaccentuels de nature linguistique est plus forte que pour les stimuli non linguistiques (clicks). Alors que Lehiste conclut de ces résultats que l'isochronie perceptive est le propre de la parole, Scott et al. déduisent de leurs observations (dont les stimuli comportaient de la parole, des clicks et de la parole distordue) que la régularisation des intervalles est l'apanage des stimuli complexes (parole et parole distordue). En ce cas, la tendance des auditeurs à imposer une régularisation perceptive ne serait pas à imputer à la seule nature linguistique des stimuli. D'ailleurs, le même phénomène perceptif est observé en musique (Gabrielsson, 1982).

En fait, les croyances sur l'appartenance du français à la catégorie des langues isosyllabiques peuvent trouver leur source dans les phénomènes perceptifs. Il est probable que l'allongement des syllabes finales de groupes en français passe inaperçu car il serait confondu avec une pause subjective. Duez (1987) montre en effet que la pause en français fait partie de l'allongement final, c'est à dire que la pause non sonore et l'allongement final relèvent selon elle d'un même phénomène phonotactique sous-jacent. D'autre part, Fraïsse (1974) montre que les auditeurs sont incapables de distinguer plus de 2 ou 3 intervalles de

temps différents. Cela s'explique par le fait que la durée est un paramètre perceptivement peu saillant. De fait, des éléments de durées différentes (plus ou moins longs) vont avoir tendance à converger perceptivement. Les différences de durées entre les syllabes accentuées finales et les syllabes inaccentuées seraient minimisées, et conduiraient à l'impression perceptive d'une certaine isosyllabité.

Enfin, la réalité perceptive de l'isochronie conduit Cutler (1980) à envisager l'existence d'un rythme régulier sous-jacent. C'est ce rythme sous-jacent que l'auditeur perçoit malgré les perturbations multiples résultant des variations segmentales tendant à le masquer sur le plan acoustique. Malgré les contraintes de production, la poussée du rythme sous-jacent conduit les locuteurs soit à rajouter des syllabes inaccentuées entre deux accents (cas des déplacements d'accents, démontrés par Bolinger, 1985 entre autre) soit à 'omettre' certaines syllabes pour rééquilibrer leur énoncé rythmiquement (Cutler, 1980). Au vu de cette constatation, on pourrait arguer que cette tendance plaide en faveur d'une programmation syllabique, et non pas uniquement accentuelle, du rythme de l'anglais. Or, Cutler estime que cela plaide plutôt en faveur de l'organisation en pied métrique, mais que la possibilité de réduction des syllabes inaccentuées dans les pieds métriques longs étant limitée, l'omission syllabique supplante la réduction là où elle n'est plus physiquement possible. En tout cas, cette tendance est selon elle une preuve concrète de l'existence d'une isochronie, au moins au niveau de la représentation sous-jacente.

Pour conclure, il semblerait qu'une voie de recherche prometteuse concernant le phénomène d'isochronie réside dans l'unité de perception des p-centers (Pompino-Marschall, 1989). Cette unité permettrait de dépasser les débats liés aux unités de rythmicité et d'appréhender de manière unifiée le rythme subjectif des langues (cf. § «Perception des phénomènes accentuels», p. 65).

3. RESUME

Nous tentons, dans ce chapitre, de situer le rythme dans le domaine de la prosodie. Pour cela, nous proposons de considérer le rythme comme « structure noyau prééminente à l'intonation ». De même, notre définition du rythme comme « organisation temporelle des proéminences » nous permet d'éviter les conceptions duales du rythme et de dépasser la dichotomie entre 'ordre temporel' et 'ordre accentuel'.

Nous nous situons dans une approche alternative, qui propose d'intégrer les concepts issus du courant post générativiste (Lieberman & Prince, 1977 ; Di Cristo & Hirst, 1997, notamment) et du courant comportementaliste (Bolinger, 1972 ; 1985 ; 1989 ; Guaïtella, 1991 etc.). En effet, nous partageons le point de vue des comportementalistes selon lequel il est indispensable d'étudier la parole en situation, comme 'comportement humain' à part entière. De fait, nous admettons, à l'instar de ces théoriciens, l'indépendance du rythme par rapport au texte. Nous concevons le rythme dans son aspect musical et métrique. Cependant, nous partageons avec les tenants de l'approche métrique la conception selon laquelle il existe un système linguistique noyau (invariants) qui sous-tend les réalisations alloformes induites par les différentes situations de communication (variabilité). Par conséquent, si nous acceptons le postulat de Bolinger (1972) qui propose que l'accent n'est pas prévisible à partir du texte, nous acceptons également l'existence d'une structure rythmique sous-jacente, présente dans tout style de parole mais dont la réalisation de surface est contrainte par le processus d'encodage.

Nous adoptons les principes de la Gestalt comme expressions des rythmes (linguistiques, musicaux, poétiques etc.). Cependant, nous pensons que le rythme a également une existence au niveau de représentation sous-jacent (Cutler, 1980). Nous proposerons d'ailleurs de mettre en évidence les traces objectives de l'isochronie, à plusieurs niveaux d'actualisation rythmique (Couper-Kuhlen, 1993). Dans cette optique, nous argumentons dans ce chapitre en faveur de l'existence conjointe du rythme syllabique et du rythme accentuel en français, dont nous nous proposons de rendre compte dans les chapitres expérimentaux.

- CHAPITRE 2 -

Les Paramètres du Rythme

Nous avons défini le rythme comme l'organisation temporelle des proéminences. En ce sens, nous considérons l'accentuation comme phénomène central du rythme. Cependant, nous envisageons également le timbre, le débit et les pauses comme participant de la structuration rythmique. Nous nous proposons donc de présenter dans le présent chapitre ces différents paramètres, en insistant particulièrement sur l'accentuation qui constitue l'essentiel de notre investigation.

1. L'ACCENTUATION

Nous distinguons, à l'instar de Di Cristo & Hirst (1997), l'accentuation non emphatique de l'accentuation emphatique. Le premier type fait référence essentiellement à l'accent final de groupe du français, qui a donné lieu à une abondante littérature dans les descriptions traditionnelles. On fera également largement référence à l'accent initial de mot ou de groupe (accent secondaire rythmique) et nous discuterons du flou définitionnel qui entoure ce type d'accentuation, souvent confondue avec l'accent emphatique initial. Nous présenterons ensuite plus particulièrement le cadre théorique de notre étude de l'accentuation du français, qui s'inscrit dans la tradition des théories métriques.

Dans un premier temps, nous présentons les paramètres acoustico-phonétiques de l'accentuation dans un cadre descriptif élargi à de nombreuses langues, en insistant particulièrement sur l'accentuation du français. On notera que les comparaisons avec d'autres langues concernent dans une grande mesure l'anglais, langue abondamment décrite, et souvent présentée comme le prototype prosodique diamétralement opposé à celui du français.

1.1. Corrélat phonétiques de l'accentuation non emphatique

La combinaison des paramètres de l'accentuation (durée, hauteur et intensité) ainsi que la relation entre les événements segmentaux et suprasegmentaux, sont sans doute plus pertinents dans la distinction entre les langues que chaque paramètre pris individuellement (Delattre, 1966b). L'influence qu'ont les différents paramètres accentuels les uns par rapport aux autres ne peut être négligée. C'est pourquoi certains phonéticiens tentent d'avoir une approche intégrée de l'accentuation (Vaissière, 1983), en ne concentrant pas leurs analyses sur un seul paramètre de l'accentuation, comme cela a souvent été le cas depuis notamment Fry (1958), lorsque la tendance était à la hiérarchisation des paramètres pertinents de l'accentuation (Rossi, 1967a).

Nous présentons cependant rapidement les caractéristiques du français par paramètre acoustico-phonétique, afin de mieux comprendre ce qui caractérise l'accentuation du français sur le plan de la substance. Ceci constitue une étape préliminaire avant d'aborder les notions de caractérisation phonologique du français.

1.1.1. Paramètres d'intensité et de hauteur

Si le paramètre d'intensité joue un rôle important dans la réalisation accentuelle de l'anglais, les travaux sur le français semblent conclure que l'intensité n'est pas un paramètre déterminant de l'accentuation, tout au moins en ce qui concerne l'accentuation non emphatique. Delattre (1966a) montre en effet que la voyelle d'une syllabe finale

(accentuée) est en moyenne légèrement moins intense que celle d'une syllabe non finale (inaccentuée). La différence d'intensité vocalique entre syllabes accentuées et syllabes inaccentuées est plus forte en anglais (environ 4,4 dB contre 0,5 dB en français, en faveur de la syllabe inaccentuée).

Pour ce qui est du paramètre de hauteur, Vaissière (1990) propose que la principale distinction entre l'anglais et le français réside dans le rôle que joue la f_0 dans l'accentuation. Les variations de f_0 en français seraient selon cet auteur directement liées aux frontières des mots et non pas aux syllabes accentuées comme en anglais. Les frontières les plus importantes en français sont en effet marquées acoustiquement par un changement de direction de f_0 d'un registre élevé à un registre plus bas, accompagné par un ralentissement local du tempo. De plus, alors que les variations de f_0 se situent sur une seule syllabe en anglais (syllabe accentuée), elles s'étalent sur plusieurs syllabes en français. Nous verrons que cette argumentation phonétique qui conduit Vaissière à classer le français comme une langue de 'frontière', ne résiste pas à une analyse phonologique qui restitue au français son statut de langue accentuelle, au même titre que l'anglais.

La paramètre de durée est important dans la caractérisation de l'accentuation du français. C'est pourquoi nous lui accordons une place prépondérante dans notre description.

1.1.2. La durée, paramètre central de l'accentuation du français

Deux hypothèses sont possibles :

- la durée syllabique en français est une simple conséquence de la présence de l'accent
- la durée est une des composantes possibles de l'accent.

Un argument en faveur de la seconde hypothèse serait de comparer le système accentuel du français à celui de l'anglais. L'accent en anglais se réalise en effet sans augmentation notable de durée, ce qui montre bien que la composante de durée n'est pas essentielle à l'impression d'accent. Si on trouve une augmentation de durée sous l'accent en français, c'est donc qu'elle en est une composante, et même en réalité une composante majeure. Fraisse (1974) estime d'ailleurs qu'il est possible de créer l'accent en français par une légère augmentation de la durée d'un élément.

Cependant, l'accent de groupe en français est congruent aux frontières prosodiques. Or, l'allongement devant frontière prosodique est un phénomène universel, y compris dans les langues à accentuation lexicale comme l'anglais (Fant et al., 1991a et b ; Delattre, 1966a et b etc.). Il est donc légitime de se demander si l'allongement sous l'accent en français n'est pas simplement lié à la présence d'une frontière. Il est intéressant pour cela de rechercher ce qui différencie un allongement, composante de l'accent, d'un allongement corollaire de la proximité d'une frontière. Un élément de réponse pourrait se trouver dans les gestes articulatoires qui sont différents suivant qu'ils sous-tendent la réalisation d'un accent de syntagme (allongement corollaire d'un accent lexical) ou la réalisation d'un allongement final à une frontière prosodique majeure (Beckman & Edwards, 1992). Une telle analyse pour le français permettrait sans doute de mettre en valeur l'existence des accents situés à des frontières prosodiques comme phénomènes accentuels à part entière. Ces hypothèses restent cependant à vérifier.

Un autre argument en faveur de la seconde hypothèse réside dans la comparaison phonétique de l'allongement accentuel entre différentes langues. Dans une comparaison avec l'anglais, l'espagnol et l'allemand, Delattre (1962 ;1966b) montre que le français présente le rapport de différence de durée entre syllabes inaccentuées et accentuées le plus important (de 1 à 2). L'allongement plus important des syllabes accentuées en français est expliqué d'une part par son poids métrique, et d'autre part par sa position finale, favorisant naturellement l'allongement. Puisque la prééminence accentuelle en français ne repose pas sur un surcroît d'intensité (-10% d'intensité sur les syllabes accentuées), comme c'est le cas pour les autres langues, la durée doit seule se charger de la prééminence accentuelle.

La prééminence de la durée est également remarquable par rapport aux paramètres physiologiques (pression sous-glottique et débit d'air) qui se révèlent inadéquats pour mesurer l'accentuation non emphatique en français (Benguerel, 1973).

A partir de ces différents constats, on peut naturellement se poser les questions suivantes :

Quelles sont les contraintes agissant sur la durée, et quelles fonctions peut-on attribuer à la durée ?

1.1.2.1. Diverses contraintes influant sur la durée

L'observation de la très grande variabilité de durée inter- et intralocuteur conduit Klatt (1976) à analyser les différents facteurs influant sur l'organisation temporelle des énoncés. La question est de savoir dans quelle mesure les effets de durée (et par conséquent la variabilité des phénomènes de durée) portent une information linguistique. Il distingue ainsi :

- les facteurs *extralinguistiques* (variables psychologiques et physiques) et *sémantiques* (notamment la nouveauté sémantique marquée par l'emphase), ces deux facteurs influant principalement sur le débit de parole ;
- les facteurs *syntactiques* qui déterminent en partie les patrons accentuels syntagmatiques d'un énoncé, et la localisation des frontières de propositions et de syntagmes où les segments sont allongés ;
- les facteurs *lexicaux* notamment liés à la position des segments à l'intérieur du mot ;
- les facteurs *phonologiques et phonétiques*, concernant la durée intrinsèque et co-intrinsèque des segments, ainsi que les effets de l'accent et des contraintes rythmiques ;
- et enfin les facteurs *physiologiques* liés aux contraintes articulatoires.

Klatt relève que la durée phonémique peut varier en fonction de tous ces facteurs d'un rapport de 1 à 8. Cependant, les voyelles et les consonnes subissent des variations spécifiques de durée. Dans un but de modélisation, il faudrait prendre en compte pour les voyelles les contrastes de durées intrinsèques, de position dans le syntagme, d'influence de la consonne post-vocalique, de l'accentuation et de l'emphase. Pour les consonnes en revanche, il faudrait implémenter des règles de raccourcissement des consonnes non

initiales de mot et en syllabes inaccentuées, ainsi que l’allongement des consonnes finales de syntagme.

A ces facteurs d’influence, il convient d’ajouter un facteur relativement négligé dans la plupart des études : la contrainte *pragmatique*, comprenant la situation de communication et le type de discours (divers types de parole lue et spontanée), ainsi que les particularités psycho- ou sociolinguistiques (Santi, 1992).

Dans le même ordre d’idée, Campbell (1992) propose d’étudier 6 facteurs susceptibles d’avoir une influence sur la variabilité des durées syllabiques :

- le nombre de phonèmes dans la syllabe ;
- le type de voyelle ;
- la position de la syllabe dans le groupe tonal ;
- le type de pied métrique (normal ou avec anacrouse) ;
- l’accent ;
- la classe du mot (mot plein ou mot outil).

La combinaison des 6 facteurs explique 76% de la variabilité des durées, le facteur le plus important semblant être le type de voyelle.

Plus tard, Campbell (sous presse) essaie de modéliser les variations de durées imputables uniquement au contexte prosodique. Pour cela, il distingue les contraintes de *bas niveau*, et les contraintes de *haut niveau*. Les contraintes de *haut niveau* vont affecter le rythme de la parole ; elles sont spécifiques au locuteur et au contexte d’énonciation, de sorte que la même séquence de sons produite par 2 locuteurs différents (ou par le même locuteur à deux occasions différentes) aura fort probablement des caractéristiques rythmiques différentes. Par contre, les contraintes de *bas niveau* assurent une certaine similarité dans les productions puisque le conduit vocal et les mécanismes articulatoires utilisés pour la production des sons individuels sont fondamentalement identiques pour les différents locuteurs de chaque langue. C’est ainsi qu’il justifie l’utilité de la normalisation des durées segmentales, afin de mettre en évidence les fonctions purement linguistiques des variations de durée (cf. § « Normalisation des durées segmentales », p. 175).

1.1.2.2. Fonctions des variations de durée

Klatt (1976) montre que les indices de durée portent des informations linguistiques considérables en parole connectée. Les changements de durées sont réguliers (donc prévisibles) et les auditeurs en connaissent les règles. Ainsi, les indices de durées pertinents linguistiquement sont les variations de durées phonologiques intrinsèques, l’allongement à une frontière de syntagme et les effets de l’accent et de l’emphase.

A l’instar de Campbell, Klatt reconnaît que la durée intervient essentiellement à deux niveaux :

- en tant qu’indice global de l’organisation temporelle, la structuration des durées segmentales contribue à la perception des structures de constituants et des patrons

accentuels (matériel emphatique notamment). La durée participe ainsi à la structuration du discours au même titre que les variations de f0 (Campbell, 1993).

– Au niveau segmental, les contrastes de durée contribuent à l'identification phonétique des divers types de segments.

Outre les variations de bas niveau (différences intrinsèques et co-intrinsèques des durées segmentales), la durée agit donc également comme indice perceptif majeur permettant de décoder la structure de surface des phrases parlées :

« Since the durational increase at a phrase boundary is usually quite large, we suggest the possibility that the marking of phrase boundaries by increasing the duration of the phrase-final syllable can serve as a primary perceptual cue to the decoding of surface structure in spoken sentence. »
(Klatt, 1976 : 1219).

Il est néanmoins difficile de déterminer dans quelle mesure les règles de durée sont motivées du point de vue de la communication, ou si elles ne sont que le résultat des contraintes imposées par les mécanismes de production. On peut simplement émettre l'hypothèse que si un effet de durée régulier est trop petit pour servir d'indice perceptif principal, alors son origine est sans doute dans le domaine de la production. Klatt (1976) situe les différences de durée pertinentes perceptivement autour de 20% ('Just Noticeable Difference') (cf. également Rossi, 1972).

Nous verrons dans le chapitre d'analyse empirique consacré aux durées que l'effet des contraintes prosodiques sur les durées segmentales en français est bien trop régulier pour ne pas être motivé linguistiquement, dans un but de communication (cf. Chapitre 5 - Durées Segmentales - , p. 165).

1.2. Place et fonction de l'accentuation non emphatique

L'accent remplit un certain nombre de fonctions, dont la plupart sont communes aux langues dites à accent libre ou à accent fixe.

Une des fonctions de l'accent les plus souvent reconnues est celle de signalisation de la structure informationnelle hiérarchique du message parlé. En ce sens, les prééminences permettraient de révéler la structure thématique dépendant de la syntaxe, et la structure informationnelle dépendant de la prosodie (Vihanta, 1993). L'auteur insiste cependant sur la multiplicité des fonctions de l'accent :

« Il serait cependant idéaliste d'imaginer que l'unique fonction des prééminences soit d'indiquer le nouveau, ou l'apport d'information, ou de ne participer qu'à la hiérarchisation de l'information. Dans des discours réels les prééminences ont plusieurs autres fonctions, totalement indépendantes de la structure informationnelle, ayant des rapports avec l'organisation linéaire de l'énoncé, la cohésion et la démarcation. Elles peuvent aussi remplir les fonctions d'intensificateur et de contraste, ainsi que, selon le genre de discours, des fonctions communicatives, expressives, et emphatiques. »
(Vihanta, 1993 : 259).

Outre la fonction *hiérarchique* de l'accent, permettant d'opérer une distinction entre les

niveaux de proéminences de l'énoncé, l'accent remplit également des fonctions *démarcatives, contrastives, rythmiques et culminatives*.

L'accent est en effet *démarcatif* dans les systèmes où il indique les frontières de mots ou de syntagmes. Il sera de fait plus naturel de trouver l'accent soit à l'initiale soit en finale de mot ou de groupes (Hyman, 1975). En ce sens, Hyman rejoint les psycholinguistes (Fraisie, 1967 et suivantes) ainsi que les linguistes pragmatiques tels Bolinger (1989) et Fónagy (1980), qui décrivent la tendance centrifuge de l'accentuation des langues.

Parallèlement à sa fonction démarcative d'unités, l'accent assume une fonction de *contraste* sur le plan syntagmatique à l'intérieur de l'unité accentuelle, en ce sens qu'il met en valeur les unités accentuables (syllabes) entre elles (Garde, 1968). Cet auteur considère le mot comme unité accentuelle potentielle. En aucun cas selon lui, le morphème peut-il être le domaine de l'accentuation en français. Or, si l'on considère l'accent initial rythmique en français, on constate que ce type d'accent peut délimiter le début d'un morphème lexical ne coïncidant pas forcément au début du mot lexical. C'est le cas notamment dans les mots longs composés d'un préfixe (Pasdeloup, 1990).

L'accent peut être *rythmique* dans les systèmes où les syllabes accentuées et inaccentuées alternent afin d'éviter les clashes accentuels (Kager, 1994).

Dans les langues dites 'accentuelles' enfin, l'accent est *culminatif*, en ce sens que chaque mot lexical contient au moins une syllabe accentuée.

Traditionnellement, on décrit le français comme ayant une accentuation finale (oxytonique) frappant la dernière syllabe du syntagme (ou du groupe de mots). Nous verrons ultérieurement, lorsque nous décrirons plus en détail l'accentuation du français, que l'on peut considérer qu'il existe également un accent initial rythmique marquant la première syllabe d'un mot ou d'un groupe de mots. A la suite de Fónagy (1980), Di Cristo & Hirst (1993) suggèrent que ces deux accents rythmiques consistent à relier par un 'arc accentuel' le début et la fin d'un mot ou d'un syntagme. Ils auraient donc tous deux une fonction intégrative d'unités rythmiques au sein d'une unité de niveau supérieur : l'unité intonative. D'autre part, ces deux types d'accent ont une fonction *démarcative* de groupes rythmiques. Cependant, certains auteurs s'accordent à penser que l'accent initial a essentiellement une fonction démarcative d'unités lexicales (Llorca, 1987 ; Pasdeloup, 1990, notamment).

1.2.1. Cas particulier : l'allongement final, un phénomène universel

L'allongement final est un cas particulier d'accentuation : s'il représente en français une marque accentuelle, congruente avec le marquage de frontière, il semble que ce phénomène soit universel. Klatt (1976) s'interroge sur le caractère physique ou linguistique de ce phénomène : on a du mal à savoir si ce type d'allongement possède ou non une fonction dans la communication. Selon cet auteur, comme cet allongement s'étend sur plusieurs syllabes, il est probablement lié à la décélération de l'activité motrice générale à la fin des actes de parole. Il diffère d'ailleurs de l'allongement final de syntagme à l'intérieur des énoncés qui sont généralement localisés sur la syllabe finale du syntagme.

« It is not known whether a speaker learns to lengthen segments at the ends of phrase boundaries in order to help the listener decode the message, or if

there is simply a natural tendency to slow down at the ends of all motor sequences or planning units. » (Klatt, 1976 : 1212).

Si la *quantité* d'allongement final a bien une fonction d'identification perceptive des frontières prosodiques, la *distribution* de cet allongement sur la Rime n'assumerait pas de valeur communicative. Ce serait un phénomène physiologique de production, corrélé à la proximité d'une frontière prosodique (Cambier-Langeveld et al., 1997).

Le phénomène d'allongement final a été particulièrement bien décrit par Lindblom (1978), qui note que ce phénomène est observé aussi bien en musique que dans la parole.

« A unit of music or of spoken language has a longer duration when, within a larger unit, it occurs finally (before a pause or a boundary). » (Lindblom, 1978 : 85).

S'agit-il d'un phénomène universel ou appris ? Lindblom (1978) donne trois interprétations possibles à ce phénomène :

- l'allongement final est appris et appartient à la grammaire de la langue ;
- il est la conséquence de contraintes de production de la parole, et peut s'expliquer selon plusieurs modèles, dont notamment le modèle de dépendance à la f0 (l'allongement final est la conséquence du temps d'exécution nécessaire à l'exécution d'un contour de f0 important), et le modèle de la planification (l'allongement final, lieu de ralentissement de l'exécution, est également le lieu qui favorise la planification des éléments subséquents).
- il est motivé d'un point de vue perceptif, et signale la structure des constituants aux auditeurs. En tant que marqueur prosodique, il se réalise non seulement devant les pauses objectives, mais également à des frontières syntaxiques (ou prosodiques) mineures, et contribue de ce fait à la perception d'une pause subjective. Certains auteurs proposent d'ailleurs que l'allongement final et la pause relèvent du même phénomène sous-jacent (Duez, 1987). L'allongement final doit être considéré comme une modalité de la pause à valeur démarcative, même si la pause n'est pas représentée par un silence (Rossi, 1987).

Lindblom conclut que l'allongement final est un phénomène phonétiquement naturel, mais néanmoins appris et spécifique à la langue. En ce sens, le degré de l'allongement final varie suivant la langue. Il existe dans les émissions sonores des nouveau-nés, mais est absent dans le langage des sourds muets.

Konopczynski (1991) adhère complètement au constat de Lindblom, mais elle ajoute que l'acquisition de l'allongement final est facilité ou ralenti suivant le système linguistique.

« La mise en place de l'allongement final, quoique fait très fréquent dans de nombreuses langues, est donc un phénomène acquis, et non inné, ni contraint physiologiquement, contrairement à ce que pensent de nombreux chercheurs. S'il n'est pas inné il nous semble en revanche intéressant de considérer l'allongement final, à la suite de Lindblom 1978, comme un processus naturel qu'on retrouve en musique, en danse, etc. D'où son acquisition

précoce et son utilisation dans de nombreuses langues. Par ailleurs, l'étude contrastive du langage émergent français, anglais, hongrois, portugais, nous amène à poser une seconde hypothèse : il semblerait que, dans les langues où la prééminence a une place stable (français, portugais), l'enfant l'acquière relativement rapidement, car il est en présence d'un modèle présentant peu de variabilité ; en revanche, dans les langues où la prééminence accentuelle est répartie à des places très variables, l'enfant, qui ne dispose pas d'un modèle stable, a plus de difficultés. ». (Konopczynski, 1991 : 336).

En français, où le marquage de la frontière de mot n'est pas en compétition avec le marquage de l'accentuation, l'allongement final est renforcé par la présence de la syllabe accentuée. Nord et al. (1990) proposent une explication métrique à cela. La structure métrique du français favorisant les pieds iambiques, l'allongement final se trouve correspondre avec la syllabe métriquement forte. Par conséquent, l'élément fort du iambe est plus marqué que l'élément fort de la trochée, et l'élément faible du iambe est moins marqué que l'élément faible de la trochée (ce dernier étant situé en position finale de pied métrique). Le contraste entre syllabes inaccentuées et accentuées est donc plus important pour les langues à pieds iambiques (comme le français) que pour les langues à pieds trochaïques (typiquement, l'anglais). En postulant la force métrique des syllabes de pieds iambiques, ce constat relativise le point de vue traditionnel selon lequel le français présenterait un contraste accentué/inaccentué moins important que l'anglais. D'un point de vue métrique en effet, ceci est fortement contestable.

Repp (1992) considère également que l'allongement final est un phénomène musical à part entière, et se manifeste dans la performance des musiciens en tant que déviations d'un battement régulier dans la structure musicale sous-jacente. Comme pour la parole, l'allongement final a une fonction première de démarcation des frontières structurelles. Cependant, il n'intervient pas uniquement aux frontières majeures (où le compositeur a prévu un 'ritardando') mais également à la fin des sous-sections et phrases mélodiques individuelles. Le degré d'allongement tend à être proportionnel à la force structurelle de la frontière (phénomène identique en parole : cf. Swerts, 1993 ; 1995). Dans une structure hiérarchique des niveaux de patrons rythmiques, l'allongement est en effet plus important aux endroits où les niveaux de frontière se superposent (Levitt, 1992).

Enfin, que ce soit du point de vue de l'interprète ou de l'auditeur, si le *degré* d'allongement varie selon les individus, l'*emplacement* attendu de l'allongement est relativement commun et varie peu à travers les individus. Il apparaît que, d'un point de vue de l'auditeur, les variations temporelles musicalement cohérentes paraissent régulières, alors que les variations temporelles non appropriées à la structure musicale sont perçues comme des distorsions. Des expériences perceptives indiquent que les auditeurs ne perçoivent pas de variations temporelles (allongement) lorsqu'elles sont situées à des frontières de phrases musicales ou à des positions métriques fortes (l'allongement étant attendu, et l'ensemble paraissant isochrone), alors qu'ils perçoivent de petites variations temporelles à des endroits incongrus. Ces résultats sont tout à fait comparables à ce qui a été trouvé pour des événements linguistiques et non linguistiques (Benguerel & D'Arcy, 1986).

On voit donc que l'analogie entre musique et parole se situe particulièrement dans l'organisation temporelle. Les phénomènes d'allongement final et de force de frontières

relative à la profondeur de la hiérarchie métrique, ainsi que les effets de compensation entre allongement final et pause subséquente (relevant aussi bien en musique qu'en parole d'un même phénomène sous-jacent), sont communs aux deux types de 'performance'.

1.3. Corrélat phonétiques de l'accentuation emphatique : nature, place et fonction

On distingue traditionnellement deux types d'accents dits emphatiques: l'accent focal (ou d'insistance) et l'accent de contraste. L'accent de contraste met en relief un élément par rapport à un autre sur l'axe paradigmatique. L'accent d'insistance est le résultat d'un choix syntagmatique; il assure la mise en relief éventuelle des unités syntaxiquement ou sémantiquement importantes de l'énoncé. Les accents emphatiques peuvent se manifester en français sur n'importe quelle syllabe sous la forme d'un contraste de hauteur, de durée ou d'intensité (Martin, 1987), le plus souvent une combinaison des trois paramètres. Cependant, on note en français une prédilection de l'emphase sur les premières et deuxièmes syllabes du mot.

1.3.1. Nature de l'accent emphatique

A l'inverse de l'accent final, les paramètres physiologiques et acoustiques tels que l'augmentation de la pression sous glottique et la variation de f_0 permettent de caractériser l'accent emphatique. Etant donné que l'intensité objective découle d'une augmentation de la pression sous glottique (Lecuit & Demolin, 1998), on peut par conséquent dire que le paramètre d'intensité, ou plus précisément la sonie (intégration temporelle de l'intensité) est un facteur déterminant pour les accents emphatiques.

En ce qui concerne le paramètre de durée, il faut noter que la durée de la syllabe augmente sous emphase, ainsi que la durée globale du mot. L'augmentation de la durée globale est souvent accompagnée d'un ralentissement du débit syllabique (Di Cristo, 1985). Le contraste durationnel entre la partie emphatisée et le contexte est amplifié par une baisse de l'intensité des éléments contextuels (Dahan & Bernard, 1996).

Pour Séguinot (1976), c'est essentiellement la hauteur et la rupture mélodique qui caractérisent ce type d'accent. Mais on observe également un allongement important et un effort articulatoire intense sur l'élément consonantique. Les expériences perceptives ayant pour objet la parole spontanée indiquent que les auditeurs se fient à l'allongement de la tenue des consonnes initiales dans une grande mesure, le trait 'durée vocalique' étant récessif. La présence d'un coup de glotte avant l'accent d'insistance est signe, selon Séguinot, d'une évolution des habitudes articulatoires en français, et participe grandement à la perception d'un accent d'insistance, surtout lorsqu'il est combiné avec le paramètre de hauteur. Son analyse instrumentale révèle que les accents d'insistance en parole spontanée sont marqués essentiellement par la hauteur, l'augmentation de l'amplitude ne représentant que 53% des cas, et ce avec une variation réduite. La longueur accompagne l'accent d'insistance au même titre que les autres syllabes accentuées, mais frappe préférentiellement les consonnes initiales des syllabes emphatiques.

Il semblerait que la hauteur de f_0 ainsi que les modifications du patron de f_0 sur le mot emphatisé (montée de f_0 inhabituelle sur la première syllabe) jouent un rôle fondamental dans la reconnaissance par l'auditeur de la fonction emphatique de la prééminence

(Séguinot, 1976 ; Dahan & Bernard, 1996).

Enfin, l'insistance peut être renforcée en français par la présence d'une pause qui la précède, et même être isolée par une pause des deux côtés. Le ralentissement de la vitesse d'articulation observé sur les parties emphatisées est fortement corrélée à l'allongement de la pause précédente. La pause fait donc partie des paramètres de la focalisation au même titre que la f_0 , la durée et l'intensité.

1.3.2. Fonctions de l'accent emphatique

Quelles fonctions remplissent les accents emphatiques ? On trouve une grande diversité terminologique dans la littérature concernant l'accent emphatique en tant que procédé de mise en relief : 'affectif', 'intellectuel', 'expressif', 'émotionnel', 'rhétorique' etc. Dans son modèle, Rossi (1980) propose de distinguer l'accent d'insistance de l'accent final traditionnel. Il qualifie l'accent d'insistance « d'expressif » ou « d'ictus endosémantique », qui peut être soit 'affectif ou émotionnel', soit 'intellectif, intellectuel ou distinctif, oratoire ou logique'. Il est donc défini par sa fonction, et non par sa nature. En 1985(b), Rossi lui reconnaît un rôle de hiérarchisation de l'information, dont le rôle est de rhématiser et de focaliser un élément de l'énoncé, et qui se manifeste en principe sur la première syllabe du mot. Il met ainsi en avant sa valeur sémantique ('accent énonciatif').

Pour Séguinot (1976), l'insistance remplit des fonctions identificatrices (émotions) et impressives (démarcatives, contrastives et oppositionnelles). L'aspect impressif, volontaire de l'accent d'insistance, est généralement présent dans toutes les études sur ce phénomène, alors que sa fonction indicielle (ou identificatrice) est souvent ignorée au profit de la fonction de signal linguistique. Cette limitation est directement liée, selon Séguinot, au fait que les descriptions de l'accent d'insistance sont issues de textes littéraires et sont soumises à la conscience linguistique des auteurs. L'analyse de parole spontanée en situation est impérative selon lui, pour explorer la véritable nature de l'accentuation emphatique.

Il apparaît en effet que l'accentuation emphatique est plus ou moins utilisée selon les styles de parole, et peut constituer la marque d'un phonostyle particulier (par exemple, discours journalistique, à visée communicative). En ce sens, ce type d'accentuation emphatique est semblable à ce que Lucci (1983) nomme 'accent didactique', ce dernier constituant alors une variante de l'accent emphatique.

1.3.3. Distribution et domaine de l'accent emphatique

Selon Boudreault (1968), ce type d'accent est considéré comme un accident rythmique, sans grande influence sur le rythme et la mélodie. Séguinot (1976) ne partage pas du tout ce point de vue. Il remarque que le domaine de l'accent d'insistance (ce que Séguinot appelle le 'syntagme expressif') est plus petit que le syntagme référentiel. Par conséquent, l'emphase 'détruit' ou en tout cas interrompt le groupement rythmique. Dans le cas de l'accent focal de contraste, Di Cristo (sous presse) indique cependant que l'organisation métrique est conservée dans les séquences post-focales. Si la mélodie disparaît, les *Mots Prosodiques* tendent à conserver leur accentuation finale marquée par l'allongement syllabique.

Il apparaît d'autre part que le domaine de l'accent emphatique est plus petit que le mot accentuel (le mot prosodique), mais ne correspond pas pour autant au mot orthographique.

Séguinot (1976) trouve cependant 37% de cas dans son corpus de parole spontanée, où syntagme expressif et référentiel sont congruents.

« (...) l'insistance serait simplement une rupture dans le groupe rythmique attendu. Peut-on dire alors qu'en affectant le rythme du groupe, l'insistance cherche à redonner au monème l'individualité qu'il perd habituellement dans la chaîne parlée ? Tel semble bien être le cas quand un mot est vraiment détaché de son contexte, mis entre guillemets en quelque sorte, voire précédé et suivi d'une pause. Mais cela est loin d'être l'actualisation la plus fréquente de l'insistance. Il n'en est pas moins vrai que l'insistance semble bien constituer une perte d'accentuation rythmique au profit de l'accentuation lexicale. » (Séguinot, 1976 : 7).

Il semblerait en effet que l'accent d'insistance ait une fonction démarcative lexicale, active en français car elle est le résultat d'un choix du locuteur et non pas d'une contrainte morphosyntaxique. Il est possible que les procédés d'insistance ne constituent qu'un trait de la fonction expressive et qu'ils doivent se combiner avec d'autres - lexicaux, sémantiques et syntaxiques - pour former un faisceau et remplir convenablement leurs fonctions impressive et identificatrice.

Séguinot (1976) remarque également que la fréquence d'utilisation de ce type d'accent varie considérablement d'un locuteur à l'autre, et ne semble pas liée au débit de parole. Il s'agirait donc bien d'un accent expressif, dont l'apparition est probablement moins contrainte par le rythme ou en tout cas moins liée à l'organisation rythmique.

Les mots pleins (lexicaux), de par leur caractère accentogène, reçoivent dans une grande mesure l'accent d'insistance. A l'intérieur des mots lexicaux, les adjectifs et les adverbes sont largement emphatisés, par rapport à leurs fréquences d'emploi dans la langue.

« les adjectifs qualificatifs et les adverbes de manière, déjà expressifs en langue, augmentent, sous l'accent d'insistance, la redondance expressive du discours » (Séguinot, 1976 : 31).

Bien que non accentogènes, les mots grammaticaux représentent une place non négligeable parmi les phénomènes d'insistance.

Enfin, les accents emphatiques portant à 95 % des mots de une à trois syllabes, marquent généralement la première (dans 79% des cas) ou la dernière syllabe. Séguinot rapproche cette tendance des contraintes morphologiques de la langue.

« Il est toujours possible, du point de vue articulatoire, d'accentuer n'importe quelle syllabe. Il n'en est pas moins vrai, cependant, que les sujets parlants sont soumis à certaines contraintes linguistiques. L'insistance est soumise, elle aussi, à une certaine contingence du système. Nous n'avons jamais trouvé par exemple, de mot de quatre syllabes accentué sur la troisième, comme "c'est un élément fondamental". » (Séguinot, 1976 : 34).

Nous pensons, à l'instar de Bolinger (1987) que cette tendance relève de contraintes rythmiques, et de l'attraction des pôles naturels de début ou de fin. Pour Bolinger en effet, un accent dont la fonction est d'attirer l'attention sera placé en début d'énoncé, et un accent de 'puissance' est placé en fin d'énoncé.

Pasdeloup (1990) note une tendance pour les accents initiaux, qu'ils soient rythmiques ou emphatiques, à marquer le début des morphèmes dans les mots plurimorphémiques, ce qui explique selon elle que si certains accents secondaires ou emphatiques ne frappent pas systématiquement l'initiale du mot lexical, ils frappent l'initiale morphémique. Cette tendance est combinée avec le marquage préférentiel des syllabes CV par les accents initiaux (ex : indispensable).

1.3.4. Mise en doute de l'emphase comme procédé accentuel

On a souvent refusé au français la possibilité d'avoir recours à un procédé essentiellement prosodique pour marquer l'emphase, arguant que le français utilise la structure syntaxique pour le faire (Garde, 1968).

« Le français ignore complètement la mise en relief d'un mot à l'intérieur d'une phrase par un renforcement de l'intensité normale de la syllabe accentuée : nous avons dû rendre par des moyens grammaticaux français le procédé d'insistance propre à l'anglais. » (Garde, 1968 : 44-45).

Garde (1968) reconnaît cependant l'existence de deux types de procédés d'insistance : l'accent d'insistance 'affectif', qui se manifeste par un allongement de la première consonne d'un mot ayant une valeur affective ou de réprobation, et l'accent d'insistance 'intellectuel', qui frappe la première syllabe du syntagme mis en relief, dont la fonction est essentiellement contrastive sur l'axe syntagmatique. Chacun des deux procédés touche la première syllabe du mot, mais l'accent 'affectif' frappe la deuxième syllabe du mot si la première commence par une voyelle. Selon Garde, aucun des deux procédés ne peut cependant être assimilé à un accent. L'accent 'affectif', en effet, n'affecte pas l'unité accentuable (la syllabe en français) mais un phonème consonantique, et il n'est pas une propriété de l'unité accentuelle (le mot) car il n'est pas obligatoire. Garde conclut donc qu'en se surajoutant au mot, il est du domaine de l'intonation. L'accent 'intellectuel' enfin, s'il affecte bien l'unité accentuable (la syllabe), ne se substitue pas à l'accent de mot normal, le mot conservant généralement son accentuation finale. De plus, le domaine de l'accent 'intellectuel' peut dépasser le cadre de l'unité accentuelle, pour atteindre n'importe quelle unité significative (morphème, groupe de morphèmes). Il conclut que ce procédé, propriété de la phrase, ne doit pas recevoir le nom 'd'accent', réservé aux traits qui sont des propriétés du mot.

C'est donc parce que l'accent d'insistance ne prend pas appui sur l'unité accentuelle, ou n'est pas congruent à l'unité accentuelle du français (le mot pour Garde), comme c'est le cas en anglais, que Garde lui refuse sa fonction de procédé accentuel.

Nous verrons que s'il est difficile de rendre compte de l'actualisation de l'accent emphatique d'un point de vue métrique (Di Cristo, sous presse), on peut néanmoins en proposer une interprétation en termes de production langagière en adoptant le point de vue « adaptationniste » de Lindblom (1990) (Di Cristo, sous presse). De fait, l'accent emphatique est, par le biais de l'accent initial rythmique dont il serait une version 'Hyper', potentiellement présent dans la structure sous-jacente du mot. De même, nous verrons dans le chapitre expérimental sur les durées segmentales que le domaine de l'accent emphatique d'intensification (initial) correspond bien à l'unité accentuable syllabique, car l'allongement différentiel propre à ce type d'accent affecte de manière constante les

constituants syllabiques, y compris le Noyau vocalique.

1.4. Perception des phénomènes accentuels

Quelle est l'unité perceptive de base pour l'accentuation ? Peut-on donner une définition perceptive de la notion d'accent ? Quels sont les paramètres acoustiques pertinents dans la perception des accents ? Quelles sont les particularités de l'accent en français sur le plan perceptif ?

Nous avons vu dans les chapitres précédents que d'un point de vue phonologique et psycholinguistique, la syllabe s'impose comme unité naturelle du rythme et de l'accentuation. Cependant, des recherches assez récentes ont montré qu'une unité intermédiaire entre les syllabes et les intervalles intervocaliques représenterait davantage la réalité perceptive du rythme et de l'accentuation de la parole. Pompino-Marschall (1989) notamment, s'est fait le porte-parole de cette nouvelle unité : les P-centers ('Perceptual centers' = centres perceptifs).

Les séquences de syllabes, de tons ou de bruits impulsifs séparés par des intervalles temporels isochrones ne sont pas perçus subjectivement comme ayant un rythme régulier. Il apparaît en fait que les débuts psychologiques d'événements acoustiques ne correspondent pas aux attaques acoustiques objectives (Pompino-Marschall, 1989). Dans des expériences d'ajustement des battements (tapping) à des syllabes accentuées, on s'aperçoit que le placement du battement par le sujet précède le début de la voyelle en proportion de la durée de la consonne initiale de la syllabe. De même, la durée de la rime influe sur la localisation du p-center, mais dans une moindre mesure que la consonne initiale; plus la rime est longue, plus le p-center est tardif dans la syllabe.

L'avantage de cette unité sur la syllabe ou sur l'intervalle intervocalique est qu'elle prend en compte la constitution segmentale, dont on voit qu'elle a une certaine influence sur la perception des accents. Cependant, l'effet des contraintes discursives dans la détermination de ces unités perceptives n'a pas été analysé. Il reste encore à tester leur validité en parole connectée.

C'est pourquoi nous privilégions pour notre part la syllabe en tant qu'unité accentuable. Sur le plan perceptif, la syllabe accentuée correspond à la notion de 'proéminence' :

« Prominence is the property by which linguistic units are perceived as standing out from their environment » (Terken, 1991b : 1768).

Cependant, la notion de proéminence est étroitement liée avec la notion de frontière prosodique, et par conséquent avec la notion de constituance, notamment en français (cf. § « Niveaux de constituance », p. 87 et § « Fonctions de la prosodie », p. 70).

Pour Portele & Heuft (1997) par exemple, le terme de proéminence qualifie aussi bien les frontières (en tant que degré de frontière) que les syllabes (degré d'accents). Leur perception est graduelle, et constante à travers les sujets. Le choix d'une même définition des phénomènes accentuels et de frontière a l'avantage de permettre l'utilisation d'un même formalisme et la quantification d'un phénomène continu.

« Prominence is a quantitative parameter of a syllable or a boundary that describes markedness relative to surrounding syllables and boundaries,

respectively. » (Portele & Heuft, 1997 : 63)

Nous émettons l'hypothèse que la perception des accents de distribution différente (initiaux/finaux) est catégorielle, alors que la perception des accents de même distribution est continue. La distinction entre un accent initial rythmique et un accent initial emphatique par exemple se fera de manière continue, alors qu'un accent initial rythmique sera distingué de manière catégorielle d'un accent final rythmique. Pour les mots monosyllabiques, la proéminence peut marquer soit la fin d'un groupe de mot (accent final de Mot Prosodique), soit le début d'un autre (accent initial générateur d'arc accentuel). Dans ce cas, des critères phonétiques comme les pentes de f_0 (Astésano et al., 1997b ; 1998), et la distance du point cible par rapport à l'attaque vocalique (Di Cristo, sous presse) participeraient de la distinction fonctionnelle et distributionnelle (pour une discussion, cf. § « Le débat sur l'analyse en configurations vs. l'analyse en niveaux », p. 238).

Globalement, il faut savoir que les facteurs lexicaux et phonologiques sous-tendant l'organisation hiérarchique des proéminences n'affleurent pas à la conscience des auditeurs. Selon Terken (1991a) en effet, les différences de proéminences auraient une fonction essentiellement démarcative sur le plan perceptif et non pas une fonction de signalisation des différences d'importance relative. Les mécanismes d'accès à l'information linguistique seraient ainsi non ressentis, non perçus comme tels.

« L'auditeur est certes capable d'extraire du signal, grâce à des mécanismes perceptifs complexes, les indices porteurs d'information, mais ces mécanismes n'affleurent pas à la conscience et on ne peut les activer artificiellement en dehors de leur contexte de fonctionnement. » (Rossi, 1980 : 16).

D'autre part, les indices accentuels perceptivement pertinents varient selon les langues et sont plus ou moins prégnants.

Dans ses expériences sur l'influence des paramètres acoustiques (durée, intensité et f_0) sur la perception de l'accent en anglais, Fry (1958) rapporte que la f_0 semble être le paramètre le plus important, avant la durée puis l'intensité respectivement. Cependant, ces résultats suggèrent que l'amplitude de variation de f_0 n'est pas aussi important dans la perception des accents, que la seule réalisation d'un changement de f_0 . En d'autres termes, les changements de f_0 ont un effet très important sur la perception de la présence/absence d'accent, mais la magnitude du changement n'est qu'un indice de second ordre.

Des résultats plus récents prenant en compte le contexte prosodique et accentuel, remettent en cause les remarques de Fry, en établissant une distinction principalement entre l'accentuation emphatique et non emphatique. Il apparaît en effet que la perception de l'accent dans l'énoncé dépend du *type* d'accent. Ladd & Morton (1997) rapportent que, dans un contexte prosodique 'normal' (non emphatique), la perception de l'accent lexical ou nucléaire en anglais est le résultat d'une évaluation contextuelle syntagmatique de la dynamique tonale, alors que la perception de l'accent emphatique implique une évaluation locale de la dynamique tonale ; plus précisément, la réduction de la dynamique tonale d'un accent est équivalente perceptivement à l'augmentation de la dynamique tonale d'un accent (emphatique) adjacent. Par conséquent, ces résultats apportent un éclairage supplémentaire au postulat de Fry (1958) selon lequel la f_0 induit un type de perception binaire (absence/présence d'un accent). Il semblerait plutôt que la dynamique tonale agisse comme

indicateur de la présence d'un accent non emphatique (perception catégorielle), mais qu'elle permette la perception graduelle ou continue des accents de type emphatique. Ladd & Morton attribuent cette différence de comportement à une distinction phonologique entre accents emphatiques et 'normaux'.

Si l'on prend en compte l'ensemble des corrélats acoustiques permettant la perception des prééminences accentuelles et des frontières, les résultats semblent indiquer plutôt une prééminence des facteurs temporels (durée syllabique et pausale) sur les facteurs de f_0 .

Portele & Heuft (1997) rapportent ainsi des résultats sur la perception du *degré* de prééminence et de frontière prosodique en allemand. S'il est vrai que la présence d'une frontière prosodique est généralement indiquée par un mouvement de f_0 montant ou descendant, seuls les contours descendants donnent lieu à une corrélation entre la hauteur du dernier pic et le degré de prééminence perçue. Pour les mouvements ascendants, ni la hauteur du pic final ni la raideur de la montée ne semblent influencer la prééminence perçue. La prééminence est ici entendue comme degré de frontière. Il semble donc que l'influence de la f_0 soit moins claire que celle du paramètre de durée dans la perception de la prééminence accentuelle ou de frontière. Ainsi, les mêmes auteurs constatent que le degré de frontière perçue est corrélé avec la durée de la syllabe finale de syntagme (pour les frontières de degré 0 à 3 perçues sur une échelle de 9), et avec la durée de la pause qui les suit pour les frontières de 4 à 9. En d'autres termes, le facteur de durée est prééminent dans la perception des frontières mineures, et le facteur de durée de pause est plus important pour la perception des frontières majeures. Avec ces résultats, les auteurs apportent des précisions supplémentaires sur l'importance relative des corrélats acoustiques des prééminences et des frontières perceptives (Swerts, 1995).

En ce qui concerne le français, on note une grande disparité dans la perception des prééminences par les auditeurs. Fónagy (1980 : 32) attribue cette diversité à « *l'ambiguïté inhérente à l'accent phrastique du français moderne* », due à la spécificité acoustico-phonétique des accents, et à la combinaison parfois étrange des paramètres acoustiques (ce que Fónagy appelle le '*chiasme acoustique*'). Il s'agit essentiellement d'une dispersion de l'effort accentuel sur plusieurs syllabes adjacentes, en termes d'intensité, de f_0 , de durée et d'autres phénomènes articulatoires comme les coups de glotte, notamment dans le style prosodique journalistique.

Allen (1975) indique que le paramètre de durée est le moins prégnant perceptivement. Or l'accent rythmique en français est essentiellement marqué par l'allongement de durée. Il semble donc raisonnable de penser que cette caractéristique de l'accentuation du français soit la cause de la difficulté pour les natifs du français à percevoir l'accent dans leur langue.

Hyman (1975) propose une explication où phonologie et corrélat phonétique de f_0 sont imbriqués. Un contour descendant est le contour sous-jacent à tout accent ; par conséquent, l'assignation de l'accent sur la pénultième est plus logique que sur la finale, car une hauteur élevée sur la pénultième suivie d'une chute de f_0 sur la finale maximise la chute, phonologiquement caractéristique de l'accent. Elle est donc plus naturelle que la réalisation d'un contour descendant sur la seule syllabe finale. Cette tendance pourrait donc expliquer pourquoi l'accent final en français est si faible et peu prégnant perceptivement, même pour les natifs (House et al., 1997). La tendance en français à placer l'accent sur la dernière syllabe conduit également au syncrétisme de l'accentuation avec l'intonation, ce qui

contribue à la confusion selon laquelle le français serait une langue sans accent (Rossi, 1980).

1.5. Approches phonologiques de l'accentuation

Afin de mieux cerner le flou théorique qui entoure la notion d'accent en français, et qui met en cause par là même l'existence d'une organisation accentuelle dans cette langue, il est nécessaire de définir la notion d'accent. Il est également souhaitable de revenir sur les notions de base concernant la phonologie de la prosodie.

1.5.1. Définitions de l'accent

Suivant à quel niveau d'observation on se situe, le terme 'accent' recouvre des réalités différentes. La distinction traditionnellement opérée en anglais entre les termes 'accent' et 'stress' en est l'illustration.

Dans un premier courant de pensée, la distinction s'opère sur les base du domaine accentuel : 'accent' correspondrait au domaine de l'énoncé, et 'stress' au domaine lexical. Cette distinction met en avant les fonctions hiérarchique et contrastive de la prosodie.

« Stress and accent each concern the relative prominence of one syllable in comparison with others ; but as defined here, stress is a property of words, accent of sentences (or utterances). » (Cutler, 1984 : 77).

Dans cette optique, les deux notions ne sont pas indépendantes l'une de l'autre ; en effet, l'accent est généralement réalisé sur une syllabe marquée dans le lexique par 'stress'. L'accent d'énoncé, propriété de la phrase parlée, a donc ici une acception d'accent focal ou nucléaire, en ce sens qu'il est dépendant d'un choix du locuteur. Au contraire de beaucoup de descriptions linguistiques, ce modèle de l'assignation des accents ne s'appuie pas sur la structure syntaxique de la phrase : l'accent est assigné principalement sur la base de facteurs sémantiques et pragmatiques. Cutler étaye son argumentation par des preuves issues des erreurs de langage. Selon elle, les indices de focalisation prosodique sont prééminents aux phénomènes de focalisation syntaxique (topicalisation etc.).

Dans une deuxième approche, la distinction entre 'stress' et 'accent' s'opère en fonction du niveau d'analyse auquel on se situe : le niveau abstrait ou le niveau concret. Pour Bolinger, 1989, 'stress' correspond à l'accent lexical potentiel (représentation phonologique abstraite), et 'accent' à la réalisation effective de cet accent (réalisation acoustico-phonétique). Mais cette réalisation concrète, phonétique, se fait essentiellement sur la base de la f0. Pour Bolinger (1985, 1989) en effet, la f0 est certainement l'indice accentuel le plus prégnant, car c'est un paramètre beaucoup moins sensibles aux contraintes extérieures, et plus révélateur des véritables intentions du locuteur. 'Accent' pour Bolinger est donc non seulement une manifestation phonétique mais également la manifestation d'une intention communicationnelle du locuteur.

« It [pitch] is probably the most efficient cue. Length and loudness can be varied in only one dimension, more vs. less ; pitch can in addition adopt a variety of shapes, including skips, glides, arrests, and combinations of these. Loudness may be distorted by wind currents, interfering objects, distance, and direction of transmission ; pitch patterns are almost immune to such

distortion. Length is affected by breathing, fatigue, and the phonetic influence of individual speech sounds to a greater degree than pitch. The ear is more sensitive to minute changes in length or loudness. (...) For these reasons, ACCENT as used in this book will normally refer to accent cued by pitch. » (Bolinger, 1985 : 22).

Enfin, l'accent peut être défini en fonction de ses caractéristiques perceptives, i.e. de sa fonction de proéminence perceptive contrastive.

« [la proéminence est] le phénomène perceptuel qui consiste à mettre en relief un ou plusieurs éléments d'un énoncé aux dépens des autres, indépendamment de sa fonction et de la manière par laquelle elle est réalisée. » (Vihanta, 1993 : 258).

Ce choix terminologique est également pour certains auteurs un moyen de désigner par un même terme les deux faces de la réalité accentuelle, de réunir l'aspect concret et abstrait. Di Cristo (sous presse) choisit ainsi le terme de 'proéminence' pour définir l'accent, car il fait à la fois référence à un phénomène perceptif (concret) et à une entité abstraite constitutive des représentations mentales linguistiques.

« D'une façon générale, nous concevons l'accent comme un fait local de proéminence (assurant la promotion d'une unité de la chaîne, principalement la syllabe) qui participe à la structuration et à la hiérarchisation des unités de la langue et du discours telles que les mots, les syntagmes et des unités de rang supérieur. » (Di Cristo, sous presse).

Lors de l'activité de hiérarchisation des proéminences, un choix s'opère sur l'axe syntagmatique pour mettre en relief l'unité la plus importante de l'énoncé. Ce choix dépend de la structure informationnelle du message et permet l'affectation de l'accent nucléaire dans un domaine de focalisation large (Di Cristo, sous presse).

1.5.1.1. L'accent nucléaire

L'accent nucléaire signale la distribution focale de la phrase. Mais deux approches linguistiques différentes expliquent son assignation. Dans le courant de pensée de la linguistique générativiste (Chomsky & Halle, 1968), la position du nucleus est déterminée par les choix lexico-syntaxiques du locuteur. C'est dans cette mouvance que se situe Crystal (1980), pour lequel la grande majorité des accents nucléaires actualisés par les variations de f0 n'apporte aucune information sémantique, car leur occurrence est syntaxiquement prévisible. Dans le courant de pensée des linguistes pragmaticiens (Bolinger, 1972), aucune règle ne peut prédire la place du nucleus car le locuteur est entièrement libre de choisir le mot auquel il va donner une importance particulière dans l'énoncé. Gussenhoven (1987) propose une vision intermédiaire, où il admet certes que l'accent est probabilitaire mais soumis à une structure linguistique sous-jacente (primitives linguistiques). Il distingue ainsi le Background (connaissance partagée) de la Variable (matériau sémantique auquel le locuteur applique une série de manipulations en fonction de la connaissance partagée). Le focus est donc la catégorie linguistique qui spécifie la taille de la Variable. Le nucleus est le moyen principal de signaler le marquage du focus. Bolinger (1987) ne rejette pas la notion de 'domaine accentuel', mais il rejette l'idée de localiser ces domaines dans la syntaxe. La théorie de Bolinger ne restreint pas l'assignation

de l'accent nucléaire aux entités lexicales et grammaticales ; elle permet les connections avec ces entités, mais leur pertinence est statistique plutôt que régulière.

En 1985, Bolinger donne de l'accent nucléaire la définition suivante :

« From the hearer's standpoint there is impact by virtue of final position, on the psychological principle of last-heard-best-noted. (...) This explains in part why the last accent in an utterance - the one that tends to gravitate toward final position - is usually regarded as the most important one. It is sometimes called the 'sentence accent', 'sentence stress', or 'nucleus'. »
(Bolinger, 1985 : 49).

L'intention d'optimiser sa communication va pousser le locuteur à produire un accent nucléaire sur la fin de l'énoncé, afin de faciliter l'accès à l'information. Il s'agit, selon Bolinger, d'un phénomène psychologique universel : la tendance, dans un syntagme intonatif, à donner une prééminence plus importante au premier et au dernier accent (accent thématique vs. accent rhématique).

Les fonctions hiérarchique et informationnelle de l'accentuation nous amènent tout naturellement à poser la question du lien qu'entretient la prosodie avec la syntaxe et la sémantique.

1.5.2. Fonctions de la prosodie : le débat sur le lien avec la syntaxe

Lorsqu'on aborde les fonctions de la prosodie, on englobe les notions d'accent, de pause, de débit etc. Globalement, la combinaison de ces paramètres assume une fonction démarcative. Dès lors, on aborde la notion de 'constituance', dont rend compte la théorie métrique dans sa version la plus récente (cf. § « La grille métrique parenthésée (Halle & Vergnaud, 1987 ; Halle et Idsardi, 1994) », p.88). Cependant, le lien entre les unités prosodiques et les unités de types syntaxiques ou sémantiques ne s'établit pas facilement, preuve en sont les nombreux débats qui résultent de cette problématique. Ces débats opposent d'une part les théoriciens de mouvance générativiste (à la suite de Chomsky & Halle, 1968) qui postulent que l'accentuation est déterminée par la structure syntaxique, et d'autre part les tenants d'une approche pragmatique pour lesquels l'accentuation révèle la structure informationnelle du message (notamment Bolinger, 1972 ; Cutler, 1984 ; etc.).

1.5.2.1. Prosodie et syntaxe

L'organisation accentuelle est contrainte par divers facteurs : cognitifs, linguistiques (syntaxiques, sémantiques...) et pragmatiques. De fait, on ne peut nier une certaine correspondance entre le rythme d'une langue, les groupes syntaxiques et les groupes de sens.

Cependant, certains auteurs vont jusqu'à postuler la prééminence de la syntaxe sur l'organisation accentuelle (Martin, 1980 ; 1981a ; Dell, 1984 ; Mertens, 1993). Di Cristo (1981) affirme même que l'intonation est toujours congruente à la syntaxe à un niveau de représentation *sous-jacent*, alors que Rossi (1981b) propose que le rôle de révélateur de la structure syntaxique ne s'applique qu'au niveau *concret* de réalisation de l'intonation (hiérarchisation des frontières syntaxiques).

D'après Rossi (1987) en effet, les contraintes énonciatives et syntaxiques interagissent

sur l'organisation prosodique. L'intonation, dans son rôle démarcatif et de hiérarchisation, est indépendante de la syntaxe, mais fortement contrainte par elle.

« L'intonation n'a pas de fonction syntaxique essentielle ; elle joue un rôle démarcatif qui délimite en tout état de cause des constituants bien formés ; la fonction de hiérarchisation qu'elle assume par ailleurs ressortit au niveau de l'énonciation. Ainsi, la prosodie a pour rôle de réorganiser l'énoncé, malgré la rigidité du cadre syntaxique, pour les besoins du message dans l'acte de communication. Mais tout n'est pas possible. La syntaxe impose des contraintes. C'est l'interaction de ces deux facteurs, énonciatif et syntaxique, qui est sous-jacente dans les règles proposées. », (Rossi, 1987 : 41).

Mais les véritables corrélations existant entre l'organisation accentuelle et la syntaxe sont mal connues. De nombreuses études fondent en effet leurs allégations sur des tests de désambiguïsation syntaxique de phrases homophones à l'aide de la prosodie. Or, cela ne constitue pas selon nous une preuve de la prééminence de la syntaxe sur la prosodie. Ces cas sont en effet très rares dans la parole réelle. Il nous semble donc extrêmement réducteur et même dangereux de fonder des modèles prosodiques sur des preuves aussi ténues. C'est d'ailleurs pourquoi certains auteurs dans la même mouvance théorique tentent de fonder leurs thèses sur des cas de parole un peu plus réaliste, à défaut d'être 'naturelle'.

Ainsi, Delais (1994) propose d'analyser la part de variabilité dans la distribution des accents imputable aux contraintes syntaxico-sémantique.

« Nous émettons l'hypothèse que les réalisations prosodiques possibles pour un énoncé [i.e. la variabilité] s'expliquent par la façon dont sont gérés les conflits et les interactions entre des contraintes de nature biologique et cognitive et des contraintes de nature linguistique. » (Delais, 1994 : 380).

La recherche des relations existant entre les contraintes rythmiques, prosodiques et linguistiques permettrait, selon Delais, d'introduire dans les systèmes de synthèse, la part de variabilité prosodique, participant de la qualité et du naturel de la synthèse vocale. Pour évaluer ces interactions, Delais s'appuie sur des phrases porteuses de type : 'Pierre a dit que demain, [le (adj.) (nom)], sans doute, irait mieux', dans lesquelles le nombre de syllabes du syntagme nominal en milieu de phrase varie de 3 à 9 syllabes. Elle distingue deux types d'accents : les accents régulateurs rythmiques et les accents linguistiques démarcatifs d'unités prosodiques. Les indices de régularité de nature rythmique qu'elle cherche à mettre en valeur sont la *fréquence des accents*, les *écarts entre accents démarcatifs* et la *recherche d'un équilibre syllabique entre les constituants prosodiques*.

La fréquence des accents est la même pour les deux types d'accents, et en règle générale, on trouve un accent toutes les deux ou trois syllabes, ce qui correspond au 'Principe du moindre nombre de syllabes (Least syllable principle)' de Wenk & Wioland (1982). Cependant, les accents démarcatifs se retrouvent toutes les 3 ou 4 syllabes, et lorsque les groupes rythmiques contiennent plus de 3 syllabes, un accent régulateur rythmique est très souvent réalisé. Son étude révèle que la recherche d'un équilibre syllabique des groupes rythmiques à l'intérieur d'un même énoncé induit des découpages qui ne respectent pas forcément la hiérarchie syntaxique. Selon Delais par conséquent, les contraintes d'alignement entre constituants prosodiques et syntaxiques (notamment les règles de construction des groupes de sens fondées sur des bases syntaxiques, selon Nespor

& Vogel, 1983) ne sont pas si simples, et la nécessité d'étudier les interactions entre les diverses contraintes de niveaux différents est impérative.

Il semblerait donc, à la lumière des résultats menés sur un corpus très contrôlé sur le plan syntaxique, que les contraintes rythmiques soient prééminentes sur l'organisation syntaxique. Nous verrons (cf. Chapitre 7 - L'organisation temporelle de la parole - , p. 259 et suivantes) que les contraintes d'organisation temporelle des prééminences que nous avons pour notre part observées, ne correspondent pas forcément avec les résultats trouvés par Delais (1994). Ceci est sans doute dû à la nature différente des objets d'étude, notre corpus comportant des énoncés de parole continue, alors que le corpus de Delais est constitué uniquement de phrases porteuses isolées.

A l'heure actuelle, il semble donc assez unanimement reconnu que la structuration prosodique est relativement indépendante de la structuration syntaxique. Ce constat est également partagé par les tenants d'une analyse métrique, ou acoustico-phonétique de phrases lues (Lehiste, 1972 ; Jassem et al., 1984 ; Monnin & Grosjean, 1993 ; Delais, 1994 ;).

« The temporal organisation of spoken language seems to take place in terms of speech production units which are fairly independent of the morphological or syntactic structure of the utterance. » (Lehiste, 1972 : 2024).

Si ce constat émane d'auteurs issus du courant générativiste, il porte néanmoins en lui une certaine contradiction. En effet, les théories et modèles prosodiques tels que ceux de Selkirk (1980) et Nespor & Vogel (1983), proposent de fonder la hiérarchie des constituants prosodiques essentiellement sur les bases de la structure syntaxiques. Le Syntagme Phonologique (situé dans la hiérarchie au dessus du Mot Prosodique, du Pied et de la Syllabe) est, dans le modèle de Selkirk (1980), directement associé à la syntaxe puisque sa structure interne est déterminée et définie par la structure syntaxique. Le Syntagme Intonatif (niveau immédiatement supérieur), qui est le domaine du contour intonatif, ne correspond par contre pas nécessairement à un constituant de la structure syntaxique. Par conséquent, si à certains niveaux de la hiérarchie prosodique il n'y a pas isomorphisme entre les structures prosodique et syntaxique d'un point de vue formel, Selkirk constate tout de même une certaine correspondance entre les deux, la prosodie reflétant souvent la structure syntaxique.

Il existe une volonté affichée de la part de ces auteurs de se détacher du formalisme strict de Chomsky & Halle, tout en demeurant très proches de la philosophie de base, à mille lieues de la parole réelle, en situation.

Un bon compromis réside, selon nous, dans la conception selon laquelle la structure syntaxique contribue à la structuration prosodique, sans forcément la déterminer (Shattuck-Hufnagel, sous presse).

1.5.2.2. Prosodie et structure informationnelle

Si la plupart des auteurs issus de la théorie générativiste trouvent une adéquation parfaite entre la structure syntaxique et la structure prosodique de la phrase, c'est parce qu'ils utilisent pour leurs démonstrations des groupes de mots très prévisibles, qui constituent une unité sémantique évidente. Or, Bolinger (1972) propose un bon nombre

d'exemples, issus de corpus de parole en situation, où les groupements de mots sont sémantiquement peu liés, et pour lesquels aucune prédiction de l'accent nucléaire ne peut être faite. Il apparaît que l'item portant l'accent le plus fort dans un patron intonatif est fonction de l'information sur laquelle le locuteur veut attirer l'attention, et non pas fonction de la structure de l'énoncé. Par conséquent, l'accent ne peut pas être prédit par la seule structure syntaxique, car le locuteur régule le placement de l'accent focal à partir d'une évaluation continue de son discours.

Il semblerait donc que le rythme soit essentiellement une aide à l'accès au sens, à l'information. D'un point de vue psycholinguistique, le rythme permettrait l'anticipation du flot de parole en promouvant la prédiction temporelle des séquences parlées (Allen, 1975) et faciliterait ainsi le décodage de la parole en unités de sens (Darwin, 1975). Le rythme serait donc essentiellement motivé par l'aspect de décodage de l'information. Le locuteur utiliserait le rythme comme vecteur et organisateur d'information.

La tendance psychologique à imposer une organisation rythmique sur ce que l'on dit ou entend est si forte, qu'en l'absence de toute organisation rythmique audible, l'auditeur sera tenté d'en plaquer une qui lui est propre ; pour cela, il s'aidera de la syntaxe. Ce n'est que dans ces cas extrêmes où la communication est en danger que Bolinger admet le lien entre rythme et syntaxe.

« In fact, in the total absence of all phonological and visual cues, the psychological tendency to impose an accent is so strong that it will be done as a 'best guess' from the syntax. (This is of course what we do when we interpret a written text.) » (Bolinger, 1985 : 17).

Dans toute autre circonstance, le lien s'établit essentiellement entre le rythme et la structure informationnelle du message. Selon Bolinger (1985), l'énoncé doit éveiller l'intérêt de l'auditeur : l'accent initial est censé éveiller l'attention, puis le déplacement de l'accent vers la droite crée la tension nécessaire pour tenir l'attention en éveil (Bolinger, 1985). C'est ainsi que le rôle de l'accent final est de graver dans la mémoire le contenu de l'énoncé (Fónagy, 1989).

« [Bolinger] démontre par cette analyse que l'intonation reflète fidèlement la genèse d'une pensée. » (Fónagy, 1989 : 218).

Cependant, alors que le locuteur a l'entière liberté d'assigner un accent focal ('power accent') où il le désire, il n'utilisera ce procédé qu'avec parcimonie. Cela répond en fait à un principe de maximisation de l'information, ou de pertinence : si trop d'éléments dans l'énoncé sont accentués, rien n'est mis en valeur et l'emphase particulière est détruite.

Ces phénomènes de déplacement d'accent en anglais sont courants, et répondent donc à un principe de pertinence de l'information. Ils correspondent également à la recherche inconsciente du rythme (régularité, isochronie). En ce sens, le rythme est non seulement indépendant mais également prééminent sur la grammaire de la langue (Bolinger, 1981). Pour Bolinger (1989), le seul lien que l'on puisse établir entre la prosodie et la grammaire est d'ordre pragmatique, et non linguistique.

Cette relative souplesse dans la réalisation de l'accent lexical en anglais laisse penser que ce dernier n'est pas essentiel dans l'accès au sens. En tout cas, on peut émettre un doute quant à l'explication phonologique réduite selon laquelle l'accent lexical en anglais

constitue un trait phonologique distinctif, entre des lexèmes de type : /'permit/ et /per'mit/, distinguant entre le substantif et le verbe. Suivant la variété dialectale également, certains lexèmes se verront même accentués sur des syllabes différentes.

En réalité, si un déplacement d'accent induit par des contraintes rythmiques, ou une erreur de placement d'accent, n'altèrent pas la structure segmentale sous-jacente du mot, les locuteurs ne réparent pas cette erreur (Cutler, 1983). Il semblerait que pour l'anglais, le maintien de la forme canonique du mot en termes de voyelles pleines et réduites est ce qui importe par dessus tout dans la reconnaissance des mots par les auditeurs. Les placements d'accents erronés ne sont pas corrigés s'ils n'ont pas d'implication *pragmatique* et n'handicapent pas l'accès au sens de l'énoncé. Ce constat va dans le sens du modèle proposé par Bolinger (1981) selon lequel le rythme de l'anglais est dual, à la fois syllabique (alternance des voyelles réduites et pleines) et accentuel (patrons de f0), le rythme accentuel étant plus variable et moins prévisibles que le rythme syllabique.

L'accentuation du français présente également cette tendance au déplacement d'accent. Si l'on admet le caractère 'mouvant' ou 'instable' de l'accent en français, c'est parce qu'il dépend de facteurs distributionnels (place de la syllabe dans le mot, la place du mot dans l'énoncé), de facteurs phonétiques (structure phonétique de la syllabe), de facteurs grammaticaux et syntaxiques (catégorie grammaticale et fonction syntaxique du mot), mais également de facteurs informationnels (sémantiques), individuels (idiosyncrasies), émotionnels et contextuels (situationnels). C'est en référence à ces derniers facteurs, dépendant pleinement de la situation de communication, que Fónagy (1980) caractérise l'accent en français d'*accent probabilitaire*.

Les paramètres prosodiques sont les indices concrets du rôle informationnel du rythme. Parmi eux, les modulations de f0 (continuation et conclusion) marqueraient les étapes de conceptualisation du locuteur. Dans cette optique, le rythme remplirait des fonctions essentiellement pragmatiques (expression du doute, du désir de passage du tour de parole à l'interlocuteur, et indication de l'achèvement de la réflexion) (Guaïtella, 1991).

« Ainsi, peu importe le 'niveau syntaxique', ou même 'sémantique' dans le sens d'une relation directe signe/sens, l'intonation montante ou descendante exprimerait plutôt la situation conceptuelle du locuteur par rapport au propos qu'il tient. » (Guaïtella, 1991 : 206).

Guaïtella note cependant que le phénomène de conceptualisation est souvent congruent avec la structuration syntaxique et sémantique en lecture, de par la nature même de ce type de support, alors que cette tendance serait moins nette en parole spontanée.

D'autres paramètres prosodiques reflètent la structure informationnelle du message. En participant à la démarcation des frontières du discours, la mélodie et les pauses sont des indices de la structure informationnelle (Swerts & Geluykens, 1994). Il apparaît également que le locuteur indique le début d'un topique par un accent mélodique haut (Brown, 1983 ; Swerts & Geluykens, 1994).

1.5.3. Discussion sur la pertinence phonologique de l'accentuation du français

Vihanta (1993), constatant l'absence de fondement phonétique de l'accentuation finale en français, que ce soit au niveau de l'auditeur ou du locuteur, qualifie l'interprétation 'accentuée' de la syllabe finale d'énoncé « *d'imagination phonologique* ». En ce sens, l'auteur apporte son crédit à la théorie selon laquelle le français serait une 'langue sans accent' (Rossi, 1980).

Rossi fonde son argumentation essentiellement sur deux points : l'absence d'accentuation lexicale en français et le syncrétisme de l'intonation et de l'accentuation.

Le français ne possède pas d'accentuation lexicale distinctive comme l'anglais par exemple ; on peut considérer dès lors que l'accent constitue un trait redondant (Garde, 1968). Cependant, les lexèmes possèdent une propriété accentuelle en ce sens qu'ils sont générateurs d'accent (accentogènes). En d'autres termes, Rossi postule l'existence du trait /accent/ dans la structure profonde des mots lexicaux en français. Or, le domaine de réalisation de l'accent en français est le groupe intonatif, où accent et intonème sont syncrétisés. Etant donné que l'intonème domine l'accent, Rossi postule que les lexèmes sont en réalité intonogènes. A partir de ce constat, Rossi postule que le français est une langue sans accent :

« *En raison du syncrétisme de l'accent et de l'intonation dans une unité, le syntagme, où seul la modulation - au sens glossématicien du terme - assume une fonction linguistique, on ne peut pas identifier l'accent comme une unité indépendante. Il semble donc que la désaccentuation du mot dans le groupe intonatif soit un corollaire de l'état de langue sans accent.* » (Rossi, 1980 : 13).

Malgré cette prise de position, Rossi (1980) reconnaît l'existence de deux types d'accents en français :

- l'accent final (ou logique), qui est un accent *interne* car il a le monème (lexème ou morphème) comme domaine sur un plan abstrait ; il est générateur d'intonation
- l'accent d'insistance, que l'auteur qualifie d'accent *externe*.

En 1985 (b), Rossi renomme l'accent externe en *accent énonciatif*, lui donnant par là même une place majeure dans la structure informationnelle de l'énoncé. Il précise que cet accent a pour fonction de rhématiser et focaliser un élément de l'énoncé et se manifeste généralement sur la première syllabe d'un mot.

Rossi se démarque ainsi de l'analyse de Garde (1968) selon lequel les morphèmes ne possèdent pas de propriétés accentuelles dans les langues à accent fixe comme le français. Il se démarque également de ce qu'il avait écrit en 1967b lorsqu'il considérait que le morphème n'est accentuable que dans la structure de *surface* dans l'unité accentuelle. Or, nous l'avons vu, si l'on élargit le cadre accentuel du français et qu'on reconnaît l'existence d'un accent secondaire rythmique, il existe certains cas d'accentuation initiale de *morphème* lexical ne coïncidant pas forcément avec le début du mot lexical si celui-ci contient un préfixe. Cette tendance accentuelle se retrouve particulièrement dans le cas de mots lexicaux longs plurimorphémiques (Pasdeloup, 1990).

Sur la base d'une analyse de discrimination de phrases homophones, Rossi (1980) conclut que l'intonation est un actualisateur de la hiérarchie syntaxique, et non l'accent. L'accent et l'intonation ne constituent en effet selon lui, ni par leur nature ni par leur fonction, deux unités distinctes. Si les morphèmes et les lexèmes ont des propriétés accentuelles, l'accentème qui les caractérise n'est qu'un générateur d'intonation, et ne peut être un générateur d'accent de mot. Rossi ne réfute cependant pas l'existence de l'accentème dans la structure profonde, mais en tant que générateur d'intonation.

Cette vision de la prosodie du français ne nous satisfait pas. Si l'on adopte un point de vue métrique de l'accentuation, on est à même de rétablir l'existence d'une organisation accentuelle en français. Un fervent défenseur de l'existence phonologique de l'accent en français se fonde d'ailleurs sur des arguments puisés dans une analyse métrique de l'accentuation. Il s'agit de Ladd (1996, plus particulièrement).

L'argumentation de base de Ladd pour réhabiliter l'accent en français est fondée sur l'indépendance entre phonétique et phonologie. Selon lui, il n'y a aucune raison de penser qu'il existe une corrélation entre les paramètres phonétiques et une catégorie phonologique quelle qu'elle soit. L'acoustique peut ne pas correspondre aux catégories perçues, notamment si on se fie aux principes perceptifs mis en évidence par les psycholinguistes (Verluyten, 1983). Par conséquent, poids métrique et prégnance perceptive ou acoustique ne vont pas forcément de pair.

Ladd (1996) argumente sur la difficulté à définir la notion d'accent. Selon lui, cette notion renvoie à une problématique purement théorique et non pas empirique. La phonologie métrique, combinée avec la notion d'association 'tonal-texte', permet de résoudre le problème théorique, d'autant plus qu'elle apporte un éclaircissement sur les trois facettes de l'accent ('stress'), souvent mélangées dans les études :

- les relations de proéminence abstraites (la 'force métrique'),
- la proéminence acoustique concrète ('accent', i.e. 'stress' en anglais),
- et la localisation des événements intonatifs liés à la proéminence (l'accent de 'pitch').

Ladd se fonde sur la constituance métrique pour apporter des preuves de l'existence d'un système accentuel à part entière en français. Certains auteurs réfutent le caractère 'accentué' de certaines syllabes, sous prétexte qu'elles sont peu marquées phonétiquement par rapport aux autres syllabes du constituant. C'est notamment le cas pour les syllabes finales d'énoncés en français, qui sont les syllabes les plus fortes de leur constituant sur le plan métrique, mais qui sont phonétiquement peu marquées, comparées à d'autres types d'accents (comme « l'accent d'insistance » par exemple). Cette remarque a conduit Vaissière (1991) et Beckman (1992) à considérer que le français ne possède pas d'accent au sens 'stress' de l'anglais (ou au sens d'accent d'intensité selon Vaissière), mais simplement des indices démarcatifs de syntagmes ou de mots phonologiques. Ce point de vue reflète une fois de plus le prisme conceptuel imposé par les descriptions phonétiques de l'accentuation en anglais. Bien que ces auteurs ne soutiennent pas ouvertement la position de Rossi (1980), la notion de 'langue de frontière' nous paraît dangereuse car elle peut trop facilement s'interpréter dans le sens du français 'langue sans accent'.

Ladd (1996) s'oppose également à cette notion et reprend l'argumentation de Dell (1984) pour le démontrer. Dell montre que l'association du contour tonal s'effectue avec la dernière syllabe pleine d'un mot phonologique, mais jamais avec une syllabe contenant un schwa. Par conséquent, il ne sera pas forcément associé avec la frontière immédiate du constituant. Ce constat trouve également sa justification dans les théories métriques de l'accentuation du français, selon lesquelles l'accent en français est assigné à la syllabe finale à moins qu'elle ne contienne un schwa (Anderson, 1984). Ladd déduit de cette observation que le contour tonal marque la syllabe métriquement forte, ce qui lui confère son statut d'accent. Que cette syllabe soit phonétiquement proéminente ou non n'est pas pertinent dans cette analyse. S'il est vrai que généralement, les syllabes métriquement fortes sont également associées directement avec un accent mélodique ('pitch accent') et sont accentuées ('stressed'), la théorie n'interdit pas la possibilité de dissociation entre le niveau métrique et le niveau mélodique et perceptif. Une syllabe peut être forte métriquement ou proéminente sans être forcément 'accentuée' ('stressed', en anglais).

« By and large, syllables that are prominent in the abstract metrical structure are also phonetically stressed, but the theory does not preclude the occurrence of systematic dissociations. A syllable can be metrically strong or prominent without necessarily being stressed. Such a dissociation is comparable to those found in music : for instance ; dissociations between rhythmical prominence (defined by the underlying beat) and melodic or dynamic prominence (defined by the harmonic changes, note durations, added loudness, etc.). (Ladd, 1996 : 59).

En d'autres termes, les accents mélodiques sont un indice indirect de la proéminence d'une syllabe, mais ils ne constituent pas eux-mêmes la proéminence de la syllabe. Si dans certaines langues comme l'anglais, les syllabes métriquement proéminentes associées avec des accents mélodiques sont également accentuées ('stressed'), ce n'est pas forcément le cas pour le français par exemple, où les accents mélodiques peuvent être associés avec des syllabes qui ne sont pas forcément accentuées (au sens où les corrélats acoustiques ne sont pas assez prégnants pour permettre la perception nette de ces syllabes). C'est pourquoi la prégnance phonétique est moindre et la perception des accents du français plus difficile pour des naïfs ou des non natifs.

Pour étayer le point de vue de Ladd, on peut également faire référence à des travaux portant sur la perception de l'accent. Rossi (1972) propose de tester l'influence de la langue maternelle sur la perception des seuils de durée. Pour cela, il compare les performances d'auditeurs français et tchécoslovaques. Le présumé sous-tendant cette expérience était que les tchécoslovaques devaient être capables de discerner des variations de durée plus fines, car leur langue contient une opposition phonologique entre voyelles brèves et longues. Ses résultats révèlent que la langue d'origine n'influence pas le seuil de perception. Par contre, le seuil de perception est influencé par l'ordre d'apparition des stimuli, quelle que soit la langue. Ainsi, le seuil est plus fin lorsque le stimulus-étalon précède la variable. Rossi en déduit donc que nous sommes habitués dans le langage à identifier une durée par rapport à un stimulus qui précède et qui tient lieu d'étalon.

Si cela est le cas, on pourrait formuler l'hypothèse que la métrique du français (pieds iambiques) favorise la perception de la proéminence marquée essentiellement par la durée. En effet, si le seuil de perception est universellement plus fin dans la condition où l'étalon

(syllabe inaccentuée) précède la variable (syllabe accentuée), on peut considérer que le poids métrique de la syllabe est renforcé en position finale, à une frontière prosodique.

A la lumière de ces arguments phonétiques et phonologiques, les phénomènes d'accentuation et de frontière prosodique peuvent être dissociés : s'il y a bien congruence entre l'accentuation et les frontières, les deux phénomènes sont néanmoins indépendants.

Dès lors, il est possible d'envisager l'existence d'une organisation accentuelle à part entière en français.

1.6. L'organisation accentuelle en français

1.6.1. Introduction

Traditionnellement, le français est considéré comme une langue à accentuation fixe située sur la syllabe finale des mots et des groupes de mots (Delattre, 1966a ; Garde, 1968 ; Dell, 1984 ; Mertens, 1992 ...). L'accent en français a donc essentiellement une fonction démarcative de groupe (syntagme, groupe rythmique, groupe accentuel, mot phonologique ou mot prosodique, suivant le courant de pensée dans lequel on se situe). La place finale de l'accent en français expliquerait pourquoi les corrélats acoustiques de l'accent final sont moins prégnants que ceux de l'accent anglais. Les autres types d'accents, dont la distribution n'est pas finale mais le plus souvent initiale, ont massivement été rangés sous l'étiquette d'*accents d'insistance*. On leur reconnaît pour domaine le mot, et ils auraient une fonction essentiellement de mise en relief lexicale et sémantique. Comme on l'a vu, on a qualifié ce type d'accent de « rhétorique, expressif, émotionnel, intellectuel, etc. ». Par conséquent, les définitions de ce type d'accent le confinent au code extra-linguistique, l'image du système accentuel du français demeurant par la même occasion 'pure' (Vihanta, 1993). Malgré sa description précoce (Passy, 1892, cité par Vihanta, 1993), ce type d'accentuation est traditionnellement considéré comme dénotant une appartenance sociolectale peu académique, voire comme une réalisation linguistique fautive. Certains linguistiques utilisent à son encontre un vocable alarmant, insinuant la propagation d'une maladie dangereuse, à caractère épidémique.

« l'accent d'insistance devient fréquent, si fréquent qu'il peut paraître normal [...] Le langage de la société cultivée maintient l'accent sur la finale. Mais pendant combien de temps le français familier résistera-t-il à la poussée du français populaire ? Et une fois le français familier atteint, qu'advient-il du français plus soigné ? » (Fouché, 1956 : 62 (LXII), cité par Séguinot, 1976).

Cependant, l'article de Fónagy (1980) a permis de mettre un terme à ce débat tendancieux et a ouvert la porte à des descriptions plus appropriées de l'accentuation du français. Fónagy remarque que la stricte oxytonie, où l'accent frappe régulièrement la dernière syllabe des groupes syntaxiques ou rythmiques, est rare. L'unité du groupe serait marquée par une polarisation accentuelle, où la première et dernière syllabe d'un groupe seraient marquées par un accent. Ce type d'accentuation initiale coexistant avec l'accentuation finale traditionnelle serait cependant plus fréquente dans certains types de discours comme les conférences ou les informations télévisuelles et radiophoniques. La généralisation de ce type d'accentuation à d'autres types de discours (conversation

notamment) serait cependant la preuve d'une évolution diachronique de l'accentuation du français (Fónagy, 1980). Cette accentuation initiale, qualifiée de *didactique* par Lucci (1983), se retrouverait dans tout type de discours à caractère informatif et ne se limiterait donc pas, selon Vihanta (1993) au discours didactique. En ce sens, Vihanta rejoint quelque peu la proposition terminologique de Rossi (1985b) qui, admettant que le français peut utiliser des procédés accentuels (et non plus seulement des procédés syntaxiques) pour hiérarchiser l'information, qualifie ce type d'accentuation « *d'accent énonciatif* ». Notons que Rossi (1985b) ajoute à son modèle de l'intonation (1980) un troisième type d'accent : *l'ictus mélodique*, à fonction rythmique, qui a pour rôle d'organiser rythmiquement une portion d'énoncé de plus de 5 syllabes « *dénuées d'accent interne ou d'intonème* ». Se faisant, l'auteur se démarque selon nous encore un peu plus de son postulat de 'français, langue sans accent'.

Pour certains auteurs, cette (nouvelle) tendance dans l'organisation accentuelle du français est source de confusion.

« (...) *it is very difficult to make clear statements on French prosody, since there are typically at least two different prosodies.* » (Vaissière, 1991 : 261).

Nous verrons plus loin que si le français est une langue en pleine évolution, les 'deux' tendances accentuelles ne sont pas contradictoires. Elles sont complémentaires et peuvent être intégrées dans un même modèle de la prosodie du français.

Avant de présenter le modèle sur lequel notre étude est fondée, nous proposons de présenter quelques caractéristiques de l'organisation accentuelle du français.

1.6.2. Le phénomène de désaccentuation en français

Ce phénomène découle directement de la structuration syntagmatique du français. En effet, l'accent rythmique en français se trouve sur la dernière syllabe de tout mot isolé. Dès que les mots isolés sont intégrés dans un groupe rythmique, la désaccentuation est la règle pour les syllabes finales des mots lexicaux à l'intérieur du groupe.

Garde (1968) propose de qualifier ces groupes rythmiques d' « *unités accentuelles virtuelles* », dont les limites varient d'un énoncé à l'autre. Cette terminologie permet de les distinguer des unités accentuelles définissables grammaticalement (les mots) et auxquelles elles peuvent être supérieures. Selon Garde, la fonction des unités accentuelles virtuelles est d'éviter la succession de deux accents (clash accentuel). Cela crée une tendance à l'élasticité de l'unité accentuelle virtuelle.

Si l'on parle de "désaccentuation", c'est parce que l'accent de mot survit en partie dans le groupe, plus ou moins suivant la constitution consonantique de la syllabe désaccentuée (consonnes allongeantes ou pas), suivant le débit, suivant le rapport syntaxique entre le mot désaccentué et le mot suivant etc. La désaccentuation incomplète est donc la règle générale (Delattre, 1966a).

La désaccentuation est une stratégie rythmique permettant de marquer ou de renforcer la cohésion à l'intérieur d'une unité de parole (Bruce et al., 1992). Ce phénomène de désaccentuation est parfois accompagné par l'assignation, sur l'initiale de l'unité de parole, d'un accent initial permettant la formation d'un arc accentuel. Di Cristo (sous presse) propose d'ailleurs de distinguer deux types d'arcs accentuels : les arcs purement

rythmiques, qui permettent d'éviter un clash accentuel (règles d'eurythmie) ; et les arcs accentuels dont la vocation est la mise en valeur d'une unité sémantique par la cohésion accentuelle. Dans ce dernier cas, une trace de l'accent final subsiste sur la syllabe finale du premier mot lexical de l'unité (cf. § « Le modèle actuel », p. 95, pour une discussion).

1.6.3. L'accentuation initiale : évolution du système accentuel en français

L'organisation temporelle est prédominante dans le groupement rythmique du français, mais il existe également une accentuation initiale manifestée par une proéminence de hauteur, ce qui complexifie le schéma prosodique du français actuel. Cette forme accentuelle se rapproche de l'accentuation propre à l'anglais, où les proéminences de hauteur sont dominantes et l'allongement récessif (Vaissière, 1990). La question que l'on peut se poser face à ce phénomène particulier concerne le rôle de l'accent initial : s'agit-il d'un accent purement lexical, dont le domaine des propriétés est le mot, ou d'un accent rythmique dont le domaine est une séquence temporelle, un groupe intonatif ou un groupe de sens? (Padeloup, 1987).

Une revue de la littérature concernant les descriptions de ce phénomène montre que l'accent initial a également longtemps été assimilé à un phénomène d'emphase. C'est d'ailleurs encore souvent le cas dans certaines descriptions actuelles. Son utilisation a longtemps été mal considérée, ou réduite à une fonction d'expressivité (colère) :

« Dans une langue même très évoluée, le rôle de l'accent de mot redevient prédominant et significatif quand la pensée s'échauffe et s'émotive, et il y a là parfois le germe d'un nouveau régime accentuel. C'est ainsi qu'en français moderne les mots porteurs d'affectivité sont frappés d'un accent intense sur la première syllabe commençant par une consonne ; ainsi des appellations injurieuses telles que coquin, filou, canaille, imbécile ne peuvent plus être accentués sur la finale ; il en résulte un conflit avec l'accentuation traditionnelle, et il arrive qu'un même mot est oxytoné ou barytoné selon qu'il exprime une idée pure ou une idée teintée d'affectivité (...). » (Bally, 1926 : 261).

Nous postulons l'existence de différentes formes d'accents initiaux en français, que l'on peut distinguer sur un plan fonctionnel. Nous pensons que l'on peut essentiellement distinguer entre les accents initiaux rythmiques (accents dits 'secondaires') et les accents emphatiques, ces derniers ayant toutes sortes de fonctions différentes (cf. § « Corrélats phonétiques de l'accentuation emphatique : nature, place et fonction », p. 61).

L'évolution du système accentuel en français a été observée très tôt par les phonéticiens français. Certains auteurs proposent même de chercher sa source dans l'influence des langues celtes sur le français (Di Cristo, sous presse). Mais Fónagy (1980) considère que la grande mobilité de l'accent en français est plus particulièrement un phénomène moderne, de par son extension à toutes les couches de la société et à toutes les situations de communication.

1.6.3.1. Fonctions des accents initiaux rythmiques et emphatiques

L'accent initial aurait deux fonctions principales : une fonction rythmique et une fonction d'emphase.

L'accent initial serait le résultat d'une contrainte physiologique et rythmique (Rossi, 1985b : 139) et aurait avant tout une fonction démarcative lexicale (Vihanta, 1991 ; 1993) :

« Pour nous, l'interprétation de l'accent initial repose sur sa fonction linguistique et énonciative qui semble être de signaler l'identité et l'importance des mots sémantiquement pleins, dans les cas où il n'y a pas de raison informative de les focaliser dans le cadre de l'énoncé. Selon nous, il s'agit donc d'accent de mot neutre à fonction démarcative, sans insistance ou d'autres fonctions expressives ou émotionnelles. » (Vihanta, 1993 : 284).

D'après Llorca (1987), l'accent initial est purement rythmique car il aiderait à organiser et à hiérarchiser les unités de la parole. Elle estime qu'il est plus facile d'organiser de longues séquences de parole dans une configuration mélodique (c'est à dire à l'aide de variations de f₀) que dans un schéma uniquement basé sur des variations de durées ou d'intensité. Padeloup (1990) précise que l'accent initial se réalise sur la première syllabe d'un mot débutant par une consonne et permet à la production accentuelle d'être conforme aux normes physiologiques et psychologiques de production du rythme. Elle poursuit qu'il n'est pas probabilitaire mais conditionné phonorhythmiquement suivant le nombre de syllabes du mot, la position du mot dans la phrase, le contexte accentuel, la nature phonétique de la première unité segmentale du mot et la stratégie accentuelle du locuteur.

Di Cristo & Hirst (1993) notent que l'accent initial étant davantage soumis à des contraintes phonotactiques, il est réalisé moins systématiquement que l'accent final; de plus, pour éviter une « collision accentuelle », il a tendance à s'effacer au profit de l'accent final. Ce constat amène ces auteurs à penser que la planification prosodique d'un accent initial est dépendante de celle de l'accent final. C'est pourquoi ils utilisent la terminologie d'accent primaire et d'accent secondaire pour désigner respectivement l'accent final et l'accent initial.

L'accent initial assumerait également une double fonction : une fonction d'indication de la hiérarchie prosodique (Fougeron & Keating, 1997) et une fonction de démarcation d'unités lexicales (Fougeron, 1998). En effet, si l'allongement final est proportionnel au degré de la frontière prosodique et par conséquent est un indicateur de la hiérarchie des constituants, il semblerait que le renforcement articulaire observé à l'initiale des constituants ait également une fonction linguistique et contribue à la segmentation prosodique. De même, l'accent initial serait une conséquence prosodique de la tendance diachronique au marquage lexical :

« Il est connu qu'au niveau du Mot, les consonnes initiales présentent, tant du point de vue diachronique que synchronique, une résistance particulière aux phénomènes de réduction phonétique et phonologique. Par exemple, les consonnes initiales de Mot se sont maintenues du latin au français alors que les consonnes non-initiales ont été réduites jusqu'à parfois disparaître. D'autre part, les processus d'assimilation, de lénition ou de coarticulation affectent moins fréquemment les segments en début de Mot. On peut alors

faire l'hypothèse qu'un renforcement articulatoire progressif des consonnes initiales à différents niveaux prosodiques, pourrait aussi avoir une fonction 'préservatrice'. Les consonnes 'renforcées' aux frontières clefs dans l'organisation prosodique de l'énoncé seraient alors plus résistantes aux influences réductrices de la parole. » (Fougeron, 1998. : 294)

Ce type d'accentuation initiale a été décrit par Fónagy (1980) comme participant de la formation d'*arcs accentuels* qui relient le début et la fin des unités lexicales ou syntaxiques (ex : 'la majeure partie'). Fónagy attribue cette forme accentuelle à la tendance centrifuge qui pousse « *l'accent du premier membre d'un syntagme vers la première syllabe et celui du deuxième membre vers la dernière syllabe d'un syntagme* », afin de « *faire mieux ressortir son unité syntactique et sémantique* » (Fónagy, 1980 : 143).

Il s'agit donc, selon cette définition, d'un cas de déplacement d'accent de la position finale à la position initiale. Il apparaît que cette définition recouvre les deux types d'accents initiaux proposés par Garde (1968) : l'accent secondaire et l'écho d'accent.

Le déplacement d'accent, dans les théories métriques, est formalisé de façon à rendre compte de la tendance à éviter les *clashes* accentuels. Il nous semble que si c'était réellement le cas, il y aurait dans la substance phonétique de cet accent déplacé à l'initiale, une trace de son appartenance phonologique (allongement par exemple). Or, les accents initiaux ont des caractéristiques phonétiques spécifiques, qui les distinguent significativement des accents finaux, tant sur le plan des paramètres de durées que des paramètres de f0 (cf. Chapitre 5 - Durées Segmentales - , p. 165, et Chapitre 6 - Configurations Tonales - , p. 235). Nous pensons qu'il s'agit plutôt d'un accent phonologique à part entière, appartenant à la structure sous-jacente du mot (Di Cristo, sous presse). Nous proposons que ce type d'accent, par la propension qu'il a à créer des arcs accentuels favorisant la constitution de groupements, contribue à l'émergence d'un niveau de constituance métrique indépendant : le Mot Prosodique complexe (cf. § « Niveau du Mot Prosodique », p. 265).

Nous observons également que cette fonction de mise en valeur de groupements peut également être doublée d'une fonction de hiérarchisation informationnelle lorsque l'accent initial est réalisé dans sa version emphatique. C'est cette fonction de hiérarchisation de la structure informationnelle que Lucci (1983) attribue essentiellement à ce qu'il nomme l'« *accent didactique* ». A l'instar des autres types d'accents en français (accent final rythmique par exemple), l'accent didactique pourrait en effet jouer un rôle de hiérarchisation de la structure informationnelle du message. Ainsi, suivant le type de discours, on pourrait envisager que toutes les unités d'un énoncé soient accentuables à l'initiale, car elles ont toutes un caractère informationnel.

L'auteur note cependant que certains mots dans l'énoncé, de par leur valeur sémantique, sont plus à même de se trouver accentués à l'initiale (*important* est plus à même d'être marqué par un accent initial que le mot 'douceur') (Lucci, 1983 : 81). Ce constat amène Lucci à considérer que l'accent didactique a plutôt une valeur 'expressive'. Il y aurait dans ce cas un rapport étroit entre signifiant et signifié (Fónagy, 1980).

Les travaux de Lucci (1983) sur « l'accent didactique » résument bien le problème de définition de ce 'nouveau' type d'accent auxquels les différents auteurs sont confrontés.

1.6.3.2. Confusion fonctionnelle des accents initiaux

Le statut de l'accent initial n'apparaît pas clairement à travers la myriade de dénominations dont il fait l'objet. Traditionnellement, cet accent est plutôt décrit comme un accent de type emphatique, qu'il soit à valeur *expressive*, *impressive* ou à fonction *énonciative de focalisation* (Rossi, 1985b). À l'évidence, cet accent revêt une certaine *polyvalence fonctionnelle* (Rossi, 1987), qui peut englober sa fonction *didactique* dans certains types de discours (Lucci, 1983).

En règle générale, ce type d'accentuation initiale en français est considéré comme un procédé de 'mise en valeur intellectuelle', dont la distribution à l'initiale d'un mot ou d'un syntagme, permet la mise en relief d'un mot ou d'un groupe de mots sémantiquement plein. Dans la lignée de Fónagy (1980) et des travaux de Léon (1993 et antérieurs), Lucci (1983) note que ce type d'accent est particulièrement présent dans les discours adressés à un public (Journaux télévisés ou radiophoniques, conférences, débats politiques etc.) et qu'il constituerait une 'marque identificatrice' socioprofessionnelle. En effet, l'accent didactique ou d'insistance se retrouve particulièrement dans des couches socioprofessionnelles et socioculturelles particulières, ayant l'habitude de s'adresser à un auditoire (enseignants, hommes politiques, intellectuels etc.). Cependant, il est aussi conditionné par des variables situationnelles, un homme politique dans l'exercice de ses fonctions l'employant beaucoup plus que le même homme politique en situation de conversation informelle.

L'accent didactique tel qu'il est défini par Lucci (1983), et telle qu'en est apparemment l'acception selon de nombreux auteurs (Mertens, 1992 notamment), correspondrait à ce que nous qualifions d'accent emphatique (d'insistance), dont la fonction est de détacher d'un fond (organisation accentuelle rythmique) un mot ou un groupe de mots sémantiquement plein par un procédé de focalisation.

Malgré la confusion assez répandue entre ces deux types d'accents, certaines descriptions distinguent clairement les deux accents initiaux, et ce depuis le début de leur identification.

1.6.3.3. Deux accents initiaux distincts

Séguinot (1976) rapporte que Passy, à la fin du siècle dernier, avait déjà repéré les divers types d'accents initiaux en français, qu'il considérait comme un déplacement d'accent par rapport à l'accent tonique traditionnel. Il avait déjà distingué deux sortes de déplacements d'accent :

- le déplacement rythmique dû à la recherche inconsciente du rythme (repris dans la terminologie des années 80 pour exprimer la recherche d'équilibrage rythmique, selon les règles dites *d'eurythmie* : Padeloup, 1990 ; Delais, 1994 ; Di Cristo, sous presse, etc.) ;
- le déplacement emphatique dû à l'importance logique ou émotionnelle des mots.

Cette dichotomie proposée par l'un des précurseurs de l'analyse prosodique du français contemporain, est à rapprocher de la dichotomie proposée par Di Cristo (sous presse) qui explique la réalisation de l'accent emphatique.

Selon Di Cristo (sous presse) les mots lexicaux seraient marqués dans la structure phonologique profonde par un accent sur l'initiale et sur la finale du mot. Lors de l'actualisation à un niveau structurel de surface, l'apparition de ces deux types d'accents dépendrait de l'application de règles d'eurythmie et de règles sémantico-pragmatiques.

– Les règles d'eurythmie permettent l'équilibrage des groupes rythmiques par l'insertion d'un accent initial *rythmique*, ou par le phénomène de désaccentuation, donnant lieu dans certains cas à la formation d'arcs accentuels. Ce type d'accent est traditionnellement appelé accent secondaire (Garde, 1968). Cependant, Garde distingue également un autre type d'accent rythmique, destiné à éviter les clachs accentuels, très faiblement marqués sur le plan phonétique. Il s'agirait d'un *écho d'accent*, et plus précisément de l'accent final : Ex : 'ca-fé-'noir, et ca-'fé-au-'lait.

– Les règles sémantico-pragmatiques sont responsables de l'assignation des accents nucléaires et emphatiques, et dépendent de facteurs individuels, contextuels et situationnels. Selon Di Cristo, l'accent emphatique prendrait appui sur un accent rythmique, et serait en quelque sorte la version 'Hyper' de l'accent initial rythmique. Di Cristo reprend ici la terminologie de la théorie « H&H » de Lindblom (1990). Selon Lindblom, le locuteur va orienter son discours selon deux tendances opposées lors de la production : si la production du discours ne nécessite pas d'effort de communication particulier, le locuteur va suivre le 'principe d'économie' et produire des mouvements articulatoires moins coûteux (parole 'hypo') ; par contre, si son discours est orienté pour un auditoire, il l'adaptera de manière à optimiser le caractère discriminatoire de sa production (parole 'Hyper'). Le locuteur adapte son discours sur cet axe continu (hypo/Hyper) suivant le besoin et l'évolution de la situation de communication.

La relation étroite existant entre ces deux types d'accents initiaux avait déjà été relevée par Rossi (1981a).

« L'ictus mélodique est le point potentiel de réalisation de l'accent externe (Rossi, 1979), appelé de façon impressive 'accent d'insistance' » (Rossi, 1981a : 196).

Vihanta (1991) distingue également en français, les accents initiaux des accents initiaux renforcés (focalisation permettant d'introduire le thème de l'énoncé). L'accent focal est caractérisé par une augmentation abrupte de la f0 et de l'intensité. L'accent initial, qu'il compare à l'ictus mélodique de Rossi (1985b) mais qu'il qualifie de non proéminent sur le plan perceptif, se distingue de l'accent focal par le *degré* d'augmentation de la f0 et de l'intensité. Les deux types d'accents initiaux se distinguent par contre très bien des accents finaux. (cf. nos propres résultats : Chapitre 6 - Configurations Tonales - , p. 235).

A part ces quelques références, il semblerait que les descriptions ont pour la plupart ignoré jusqu'à ce jour cette accentuation initiale rythmique (voir cependant Hirst & Di Cristo, 1984 ; Padeloup, 1990 ; Llorca, 1987). Nous nous proposerons dans la suite de ce travail de rendre compte des distinctions entre accents initiaux rythmique et emphatique sur un plan expérimental.

Nous sommes cependant consciente de la difficulté de la tâche, que même un cadre méthodologique précis, définissant les types d'accents sur un plan fonctionnel, ne suffit pas forcément à désambiguïser. En ce sens, nous partageons la prudence de Vaissière (1997) :

« Même si un accent, initial ou final, est perçu unanimement par un ensemble d'auditeurs, son interprétation sur le plan fonctionnel n'est pas toujours constante. La lecture des écrits de Fónagy invite à la plus grande méfiance devant une interprétation trop rigoureuse des événements prosodiques. Par exemple, un événement simple comme la montée initiale de f_0 souvent observée dans la parole lue de façon neutre (et qui ne donne pas forcément lieu à la perception d'une proéminence) (Vaissière, 74, 75), peut recevoir un certain nombre d'interprétations, parmi lesquelles il est difficile de trancher : 1) une explication d'origine rythmique (réalisation d'un écho de l'accent final) ; 2) un procédé d'origine stylistique, caractéristique des corpus lus ou d'une lecture professionnelle, un accent didactique ; 3) marque régionale ; 4) une caractéristique intrinsèque des mots longs, ou une marque de contraste entre mots lexicaux et mots grammaticaux généralement prononcés dans le registre inférieur ; 5) un indice de frontière gauche marquant le début d'un mot ou d'un groupe de mots ou même d'une phrase (fonction linguistique démarcative) ; 6) une insistance particulière sur le mot (fonction hiérarchisante, sémantique, relative au sens de l'énoncé ou pragmatique) ; 7) ou même une 'phonologisation' de la tension naturelle en début d'unité linguistique observable dans plusieurs langues (Vaissière, 1995). » (Vaissière, 1997 : 482-483).

Avant de présenter le modèle de l'accentuation sur lequel nous avons choisi de baser notre étude, nous proposons de le situer par rapport au courant théorique dont il est issu: la phonologie métrique.

1.6.4. Approches métriques de l'accentuation

1.6.4.1. Les origines de la théorie métrique

Kager (1994) propose un rappel historique de la théorie métrique. Elle a été fondée à l'origine par Liberman (1975), puis élaborée par Liberman & Prince (1977) et Halle & Vergnaud (1978). Elle s'inscrivait à la fin des années 70 dans la lignée de la phonologie non linéaire (à l'instar de la phonologie autosegmentale), et avait pour but de développer une alternative aux propositions de la phonologie linéaire (Chomsky & Halle, 1968). Des opérations locales étaient formalisées pour rendre compte des relations hiérarchiques entre les différentes proéminences lexicales.

L'idée centrale de la théorie métrique réside donc dans la représentation et la formalisation des relations de hiérarchie accentuelle. Pour cela, deux types de représentation ont été proposés : la représentation en arbre métrique, et la représentation en grille métrique.

1.6.4.1.1. Les arbres métriques

L'accent y est représenté comme une hiérarchie de structures à branchement binaire, chacun étant étiqueté Strong-Weak ou Weak-strong. L'accent tel qu'il est représenté dans un arbre métrique est une propriété *relationnelle*. Si au niveau de la syllabe la relation est SW, elle peut être au niveau supérieur WS et représentera la proéminence relative des syllabes (cf. Figure 3).

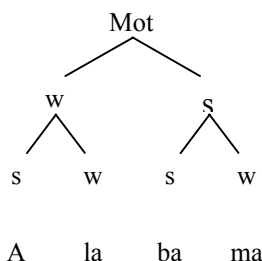


Figure 3 : Représentation en arbre métrique du mot 'Alabama'

1.6.4.1.2. Les grilles métriques

L'arbre représente les proéminences relatives mais est incapable de représenter l'*alternance* rythmique entre les syllabes fortes et les syllabes faibles, ainsi que les cas de *clashes* accentuels. En introduisant la représentation en grille métrique, où la hauteur de la colonne de la grille indique le degré de proéminence, Liberman (1975) propose de tenir compte de l'aspect de structuration rythmique. La grille métrique représente donc l'accentuation comme un procédé *hiérarchique* plutôt que *relationnel* (cf. Figure 4).

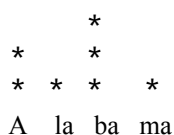


Figure 4 : Représentation en grille métrique du mot 'Alabama'

1.6.4.2. Evolution de la représentation métrique

1.6.4.2.1. Inclusion du niveau du pied métrique

Une évolution centrale dans la représentation métrique est l'inclusion du niveau du *pied accentuel* par Halle & Vergnaud (1978) et Selkirk (1980). Ce niveau a permis d'éliminer le trait accentuel segmental de $[\pm \text{Stress}]$ qui avait été introduit par Liberman & Prince (1977) pour distinguer entre les arbres de même composition. Chaque pied métrique en effet suppose la présence d'une syllabe forte et unique (la *tête*) coïncidant avec une syllabe accentuée, et il comprend les syllabes faibles associées à la tête. Le principe de cette nouvelle représentation suppose également le principe de hiérarchie stricte et fermée, i.e. que toute catégorie de niveau n doit être dominée par une catégorie de niveau $n+1$. Enfin, une condition d'exhaustivité exige que chaque syllabe soit incluse dans la structure rythmique. La pertinence du niveau du pied métrique en tant qu'unité de représentation de la hiérarchie accentuelle est appuyée par la découverte de sa fonction de domaine pour les règles segmentales (Selkirk, 1980). Le niveau du pied induit par conséquent un paramètre de construction des arbres ou des grilles de *dominance*, qui détermine le côté du pied où la *tête* va être placée. La dominance droite ou gauche du pied métrique va déterminer la *direction* de construction du pied.

Typiquement par exemple, le pied accentuel en français a une *dominance* droite et une *direction* de construction de droite à gauche. Si les règles accentuelles dans la théorie

métrique sont relativement indépendantes du squelette segmental des constituants syllabiques, elles ont tout de même trait à la complexité interne de ces constituants. Ainsi, pour le français, si l'accent est assigné à la syllabe finale, il est assigné à la pénultième dans le cas où la syllabe finale contient un schwa (Anderson, 1984).

1.6.4.2.2. *Les opérations dans les grilles métriques*

Nous avons vu que les chercheurs ont trouvé un consensus dans la représentation en grille métrique, fondé sur la nécessité d'une représentation hiérarchique de l'accentuation, comblant ainsi les lacunes de la représentation en arbre métrique. La grille métrique prend appui sur la nature rythmique de l'accentuation. Elle permet de formaliser les phénomènes de désaccentuation et de déplacement rythmique de l'accent. Les configurations 'mal formées' de la grille, telles les clashes accentuels ou les vides métriques, sont transformées en configurations métriquement acceptable par des opérations telles que la suppression, l'insertion et le déplacement des marques métriques (x). Ces opérations se situent au niveau du *ped métrique*.

- Supprimer x est une opération locale induite par le paramètre de *dominance* afin d'éviter les phénomènes de clashes accentuels ;
- Insérer x , également induite par le paramètre de *dominance*, permet un équilibrage rythmique des éléments accentuels dans des séquences trop longues de syllabes métriquement faibles.
- Déplacer x peut se décomposer en 'supprimer x et insérer x ', et remplit donc les mêmes fonctions que les deux opérations susmentionnées.

1.6.4.3. Niveaux de constituance

Une caractéristique fondamentale des modèles métriques de l'accentuation est leur nature hiérarchique. Les différentes règles qui interviennent dans la construction des grilles ou des arbres métriques ne s'appliquent correctement que si la notion d'alternance rythmique est définie hiérarchiquement. Selkirk (1980) propose 3 niveaux entre celui de la syllabe et celui du syntagme intonatif, pour rendre compte des accents secondaires, des relations syntagmatiques entre les accents et de l'assignation de l'accent nucléaire. Certains modèles proposent jusqu'à 6 niveaux intermédiaires. Couper-Kuhlen (1993), par exemple, ajoute un niveau intermédiaire (le groupe clitique) à la hiérarchie proposée par Selkirk (1980) et Nespor & Vogel (1983). Hirst (1983) propose un seul niveau intermédiaire entre la syllabe et le syntagme intonatif : le Pied.

Dans la tradition de l'école anglaise, les énoncés sont composés de groupes tonals majeurs, eux-mêmes composés de groupes tonals mineurs. En ce sens, cette hiérarchisation des niveaux de constituance respecte l'hypothèse de hiérarchie stricte (Selkirk, 1984). Ladd & Campbell (1991) remettent en question ce point de vue, où les domaines prosodiques sont forcément en nombre limité. Ils proposent par contre l'existence de 'super domaines', ou 'domaines prosodiques composés', du même type que les domaines qu'ils dominent. Ils se fondent sur des études phonétiques sur les corrélats acoustiques des frontières prosodiques qui indiquent une très grande profondeur des structures prosodiques, reflétée par la durée segmentale et les propriétés de f_0 dans le voisinage des frontières. Ladd & Campbell (1991) essaient de montrer qu'un approfondissement de la structure prosodique

parvient à mieux rendre compte de la variabilité des durées syllabiques. Ainsi, le modèle à quatre niveaux de constituance (opposé aux modèles traditionnels à deux niveaux) rend compte des différences de durées des syllabes *finales* de syntagmes. Par contre, le modèle à deux niveaux suffit pour rendre compte des différences de durée des syllabes initiales et médianes.

La nature hiérarchique du rythme est par ailleurs reconnue par des courants de recherches non rattachés à la représentation métrique de la parole. Le rythme, comme activité polymorphe et comme principe organisateur universel dans des activités humaines diverses, est intrinsèquement multi-niveau.

" (...) speech rhythms emerge from diverse rhythmic levels ; there is the rhythm of the words, phrases, and sentences, the rhythm portraying emotions, the rhythm used for emphasis, the rhythm unique to a regional dialect, and so on. As has been found for musical rhythms, the way one speech rhythmic level operates depends on the value of the other levels. The emergent rhythm is multifaceted. There is the beat or meter, the pace or tempo, the accent rhythm, the timing rhythm, and the melodic or intonation rhythm (Ladd, 1980...) These aspects are not independent in much the same way that the rhythmic levels of a basketball game are not independent. They support each other and allow each to emerge. » (Handel, 1989: 383).

En ce sens, la définition de Handel se rapproche également de la description des deux rythmes de l'anglais par Bolinger (1981), et fait le pont entre les différentes théories du rythme.

1.6.4.4. La grille métrique parenthésée (Halle & Vergnaud, 1987 ; Halle et Idsardi, 1994)

Cette nouvelle version de représentation métrique de l'accentuation repose sur le concept fondamental de prééminence de la notion de *constituance* sur la notion de *proéminence*. L'accent serait le reflet des groupements que les locuteurs imposent sur les séquences d'éléments linguistiques ; une fois ces groupements établis, une proéminence plus importante est donnée à certains éléments dans le groupe. La proéminence est donc un effet secondaire du groupement des éléments en constituants.

Les différentes étapes de construction de la grille métrique sont représentées par des lignes parallèles (la ligne des phonèmes et les lignes de marqueurs abstraits) constituant un 'plan' métrique. Les parenthèses servent à délimiter les groupements qui désigneront les éléments à gauche et à droite comme têtes de constituants. L'ensemble des *x* (marqueurs abstraits) et des parenthèses (groupements) ainsi construits constituent les grilles métriques. Elles contiennent des informations concernant le groupement des éléments en constituants, les têtes des constituants, et les différents degrés de proéminence des éléments de la chaîne. Ainsi, des marques *x* adjacentes sur une même ligne sont regroupées en constituants, dont la tête unique est alignée verticalement à une marque de niveau supérieur.

Sur la ligne 0, les marqueurs *x* représentent les unités accentuables, et les parenthèses correspondent aux frontières de pieds. La ligne 1 contient les têtes des constituants de la ligne 0. La ligne 2 contient les éléments de têtes des constituants de la ligne 1 etc. La grille métrique est donc le résultat final de l'application de principes universaux de *Projection des*

éléments métriquement forts de la ligne 0 aux lignes supérieures, de *Marquage des frontières* et de *Localisation des Têtes*. Des procédés morphologiques et phonologiques propres à chaque langue permettent de définir la structure des pieds représentatifs du système accentuel de chaque langue. Suivant les langues par exemple, les constituants métriques regroupent des projections d'unités particulières susceptibles de porter l'accent, pouvant être soit des séquences de phonèmes, des syllabes, des mores etc. Di Cristo (sous presse) propose la représentation suivante pour le français, pour laquelle la proéminence la plus à droite sur la ligne (L1) est investie de la proéminence maximale dans la séquence et constitue ainsi la borne d'un constituant métrique de rang plus élevé. Par conséquent, cette proéminence reçoit une parenthèse sur cette ligne et est projetée comme tête sur la ligne supérieure (L2) (cf. Figure 5). Di Cristo s'inspire de la représentation en grille métrique parenthésée de Halle & Idsardi (1994), qui proposent d'indiquer les niveaux de constituance par une seule parenthèse, alors que le modèle antérieur de Halle & Vergnaud (1987) proposait un double parenthésage droite/gauche des constituants métriques :

						x	(L2)
		x				x)	(L1)
x	x	x)	x	x	x	x)	(L0)
Les	a-	mis	de	mes	voi-	sins	

Figure 5 : Représentation en grille métrique parenthésée de l'énoncé français : « les amis de mes voisins » (tiré de Di Cristo, sous presse).

La grille parenthésée permet donc la représentation conjuguée des proéminences, du rythme et de la constituance. Cependant, elle permet aussi de représenter la constituance sans la proéminence, et la proéminence sans la constituance. Elle permet l'application des opérations de suppression, insertion et déplacement sur les marques *x*, ainsi que des opérations sur les parenthèses.

La construction de la grille métrique repose donc uniquement sur deux mécanismes formels, fondés sur des contraintes universelles : le placement des marqueurs *x* et le placement des parenthèses.

Tous les paramètres n'entrent pas en ligne de compte suivant les langues : le poids syllabiques par exemple n'est pas toujours un facteur d'accentabilité. De même, on choisira de projeter la frontière droite ou la frontière gauche de la syllabe, suivant la langue.

Enfin, il faut noter que Halle & Vergnaud (1987) et Halle & Idsardi (1994) se démarquent du principe de hiérarchie d'alignement strict proposé par Selkirk (1984). Ils proposent de représenter les structures accentuelles et syllabiques sur des plans différents, afin de rendre compte notamment du comportement des têtes syllabiques d'éléments comme le Schwa, qui dans certaines langues peuvent être métriquement faibles dans certains cas et métriquement fort dans d'autres cas.

1.6.5. Limites de la représentation métrique

Une limitation majeure de la représentation métrique est qu'elle ne permet pas de rendre compte de l'organisation rythmique lors du passage à la performance. Et ceci pour une raison principale : elle se fonde sur une représentation morphosyntaxique des relations de proéminences, et par là même, elle s'interdit la représentation des structures prosodiques

dépassant le cadre de la phrase telle qu'elle est définie en syntaxe (Couper-Kuhlen, 1993). Par conséquent, elle est peu adaptée à la description de la parole spontanée, qui, comme le proposent Blanche-Benveniste & Jeanjean (1987), possède une syntaxe propre. De même, comment peut-elle rendre compte de l'organisation prosodique du message dépendant du choix du locuteur, lié à des considérations pragmatiques ? (Bolinger, 1972). Cependant, cette critique peut s'étendre à d'autres types de modèles. On peut supposer que cette limitation s'estompe si le cadre métrique est appliqué non pas au *texte*, mais au *message*, supposant alors le passage par une phase de délimitation *perceptive* des prééminences et des groupements.

Nous nous proposons maintenant de présenter le modèle de l'accentuation du français sur lequel nous fondons notre étude.

1.6.6. Un modèle de l'accentuation du français

Le modèle que nous utilisons a donné lieu à un certain nombre de publications (Hirst & Di Cristo, 1984; Di Cristo & Hirst, 1993; Di Cristo & Hirst, 1996; Hirst & Di Cristo, 1996; Di Cristo & Hirst, 1997; Hirst et al., sous presse). Nous le désignons de manière générique comme le 'modèle de Di Cristo & Hirst', mais nous ferons référence aux différentes publications au fur et à mesure de la présentation de l'évolution du modèle. Nous renvoyons les lecteurs aux articles de fond sur l'accentuation du français de Di Cristo (sous presse), qui proposent l'historique détaillé de l'évolution du modèle en le situant par rapport aux diverses descriptions de l'accentuation en français.

1.6.6.1. Fondements théoriques

Le modèle de description de Di Cristo & Hirst s'inscrit dans la tradition de la grammaire générative et de l'approche cognitive, opposée à l'approche éthologique qui a pour but l'étude du comportement (Hirst, 1988). Le but de l'approche cognitive est de formaliser les représentations mentales de l'énoncé, qui font que la communication peut s'établir. Une fois les représentations mentales formalisées, on suppose que les variations observées ont également un but dans la communication et sont le reflet de contraintes isolables.

« Nous pouvons supposer que d'un sujet à l'autre, pour un même énoncé, nous trouverions des variations importantes dans la nature physique de cette représentation. Il semble, cependant, légitime de penser que ces représentations mentales, bien que physiquement différentes, ont des propriétés abstraites équivalentes d'un sujet à l'autre. Ce sont donc ces propriétés abstraites que l'approche cognitive se donne comme but de déterminer, en faisant abstraction (légitime) des variations individuelles. » (Hirst, 1988 : 211).

Cette quête des éléments abstraits constitutifs des systèmes linguistiques prend sa source dans la recherche des universaux linguistiques, chers à la tradition générativiste. Si l'on prend deux langues opposées sur le plan rythmique et intonatif, que sont l'anglais et le français par exemple, on peut supposer que leurs systèmes intonatifs possèdent les mêmes éléments abstraits. Les règles de dérivations appliquées à ces éléments dans le processus de linéarisation créent des gabarits tonals spécifiques à la langue. Cette quête des universaux

justifie l'existence de l'Unité Tonale dans le modèle de Di Cristo & Hirst (pour une discussion, cf. § « Evolution du modèle », p. 92).

Les modèles phonologiques qui s'attachent à décrire les systèmes linguistiques peuvent être scindés en deux courants principaux : les modèles issus de la phonologie de l'intonation (Pierrehumbert, 1980) et ceux issus de la phonologie métrique (Halle & Vergnaud, 1987, notamment).

Le modèle de Di Cristo & Hirst s'intègre dans le courant de la théorie métrique et postule de fait la prééminence du rythme dans l'organisation prosodique comme « *principe organisateur de premier ordre* » (Di Cristo, sous presse). Le rythme est donc non seulement indépendant de l'intonation mais est de surcroît prépondérant dans une description des systèmes linguistiques. Di Cristo & Hirst (1993) constatent en effet que le rythme, dont les fondements biologiques et psychoacoustiques sont universels, joue un rôle fondamental dans l'identification des langues.

De fait, une telle approche permet de distinguer entre les événements de f0 locaux liés aux phénomènes de proéminence, des événements de f0 globaux liés à l'intonation. Ainsi, les relations de proéminences et la constituance de surface (accentuation et groupement) sont représentés dans l'arbre métrique, et l'intonation est représentée dans la chaîne tonale (Ladd, 1992). Sans un cadre théorique cohérent, les phénomènes accentuels et intonatifs peuvent être aisément confondus, comme cela a souvent été le cas en français particulièrement. C'est pourquoi le modèle de Di Cristo & Hirst a évolué de façon à rendre compte des différents niveaux de constituance, permettant de définir les entités rythmiques (groupes accentuels) indépendamment des groupes ou syntagmes intonatifs (cf. § « Evolution du modèle », p. 92).

Un principe fondamental du modèle est le principe de l'indépendance des niveaux d'analyse. Les auteurs proposent quatre niveaux de représentation indépendants les uns des autres, chacun des niveaux étant cependant interprétable aux niveaux adjacents (immédiatement supérieurs et inférieurs) (Hirst, 1983 ; Hirst et al., sous presse). La représentation des différents niveaux permet d'éviter le danger d'une interprétation directe entre les niveaux acoustico-phonétiques et les catégories linguistiques fonctionnelles. En effet, les propriétés acoustiques observables, comme la f0, la durée et l'intensité, ne sont pas des corrélats directs de catégories fonctionnelles (le focus par exemple) ni la réalisation directe des structures morphosyntaxiques. Elles n'expriment qu'indirectement le focus ou les frontières de syntagme, mais elles sont le corrélat de la structure et des catégories phonologiques. En ce sens, l'approche métrique autosegmentale permet de rendre compte des relations entre les phénomènes prosodiques et les formes et fonctions prosodiques. Une approche phonologique de la prosodie permet donc d'éviter d'établir des liens directs entre les formes phonétiques et la fonction communicative (Ladd, 1992).

En partant du plus concret au plus abstrait, les quatre niveaux sont les suivants :

- le niveau physique, domaine des paramètres acoustiques et physiologiques de l'accent;
- le niveau phonétique, auquel les données physiques ont été isolées : à ce niveau peuvent intervenir des procédures de modélisation et de normalisation des données

acoustiques et physiologiques brutes. C'est également le niveau de représentation perceptive des accents ;

- le niveau phonologique de surface, où les données continues sont discrétisées, est le niveau de représentation des structures prosodiques observables ;
- le niveau phonologique profond, qui représente la structure sous-jacente du système de la langue, est le niveau de représentation des gabarits métriques abstraits.

1.6.6.2. Evolution du modèle

Bien que les tenants de la phonologie syntagmatique s'accordent sur la nature hiérarchique de la structure prosodique, les degrés de complexité hiérarchique entre le niveau de la syllabe et le niveau du syntagme intonatif sont un sujet de controverse (cf. « Niveaux de constituance », p. 87).

Ainsi, Hirst & Di Cristo (1984) proposent une unité intermédiaire pour le français entre la syllabe et le groupe intonatif : l'Unité Tonale, qui, comme pour l'anglais, est l'unité de réalisation intonative de base, mais qui est constituée pour le français du gabarit tonal /BH/, alors que l'Unité Tonale en anglais possède un gabarit tonal /HB/. L'Unité Tonale délimite des groupes accentuels marqués par un accent initial ou un accent final. Elle est définie essentiellement sur une base lexicale.

Les auteurs étendent leur modèle en 1993 et proposent l'introduction d'un niveau de constituance supplémentaire entre l'Unité Tonale et l'Unité Intonative : il s'agit de l'Unité Rythmique. Les auteurs reconnaissent dès lors le rôle central du rythme, élément noyau de la langue et base de tout système prosodique. De plus, leur démarche est motivée par la conception selon laquelle le rythme est fondé sur une composante temporelle et mélodique. L'Unité Rythmique est délimitée par un accent primaire, et est immédiatement supérieure à l'Unité Tonale. Les unités rythmiques et tonales sont congruentes lorsqu'elles sont situées à la frontière d'un accent primaire. Si l'énoncé contient un accent initial, l'Unité Rythmique contient deux Unités Tonales.

Avec ce nouveau formalisme, Di Cristo & Hirst (1993) se rapprochent de la conception métrique du rythme. C'est d'ailleurs à cette date qu'ils adoptent la représentation en grille métrique parenthésée de Halle & Vergnaud (1987), qui rend compte à la fois des prééminences et de la notion de constituance. Rythme et intonation y sont présentés comme deux projections, temporelle et tonale, d'une même structure phonologique sous-jacente. L'Unité Rythmique, intermédiaire entre l'Unité Tonale et l'Unité Intonative, permet de rendre compte de faits qui relèvent à la fois de l'organisation tonale et de l'organisation temporelle. La mélodie relève en effet en partie du rythme, dans l'aspect dynamique des changements de hauteur, et pas seulement de l'intonation.

En 1997, les auteurs achèvent d'asseoir leur description de l'accentuation du français dans un cadre métrique. En proposant différents niveaux de constituance pour rendre compte des frontières de syntagme (frontières mineures) et d'unités intonatives (frontières majeures), les auteurs proposent l'indépendance de l'unité accentuelle par rapport à l'unité intonative, et se démarquent des descriptions du français 'langue sans accent'.

«Les études qui considèrent le Mot Prosodique et le groupe intonatif comme des unités coextensives, paraissent se rallier à la thèse du syncrétisme de

l'accent et de l'intonation (...) » (Di Cristo et Hirst, 1997 : 76).

En théorie, l'Unité Rythmique peut contenir un nombre illimité d'Unités Tonales, mais les auteurs pensent qu'on ne peut pas en trouver plus de trois dans la réalité. Ils en donnent l'exemple suivant (les syllabes accentuées sont indiquées en italique) :

le général de l'armée de l'air

où l'Unité Rythmique contiendrait les trois UT (le gé)(néral de l'ar)(mée de l'air).

Cet exemple nous semble arbitraire et par trop théorique. Si ce genre de schème accentuel existait réellement, il devrait être accompagné selon nous de 'traces' d'accent final intermédiaire, au moins sur la syllabe finale du premier mot lexical (par un Downstep et une trace d'allongement, comme le propose plus tard Di Cristo, sous presse), et ceci afin de permettre une prononciation acceptable.

Certains auteurs ont d'ailleurs intégré des règles dans leurs modèles ne permettant pas ce genre de réalisations accentuelles.

« Dans une succession de deux syllabes accentuées, au moins une sera de type accent final : l'accent initial étant un élément facultatif dans le groupe intonatif, on ne trouvera jamais deux AI de suite, sans qu'il y ait entre les deux un autre accent final. » (Mertens, 1992).

Nous proposons dans le paragraphe suivant une critique du niveau de l'Unité Tonale, que nous ne considérons pas pertinent dans une description du système accentuel du français.

1.6.6.2.1. L'unité tonale est-elle adaptée à la description prosodique du français ?

Sur le plan métrique, l'Unité Tonale est synonyme de 'Pied Métrique' ; elle caractérise sur le plan phonétique la substance d'une unité supposant une réalisation purement tonale de l'unité accentuable. Elle constitue une unité de traitement où se réalise la synchronisation des segments tonals et des segments phonématiques (Di Cristo & Hirst, 1993).

Une critique importante de l'Unité Tonale a été formulée par Jun & Fougeron (1995). Les auteurs relèvent une certaine contradiction entre la définition phonologique (mot, entité phonologique) et le domaine d'actualisation de l'Unité Tonale, qui ne tient pas compte des frontières du mot lexical. L'Unité Tonale ne respecterait pas le principe de la hiérarchie stricte proposé par Selkirk. Di Cristo & Hirst rétorquent à cela qu'il est clairement admis que le Pied Métrique ne respecte pas, intrinsèquement, les frontières de mots (Hirst & Di Cristo, 1996).

Comme Jun & Fougeron, nous ne souscrivons pas à l'existence de cette unité intermédiaire en français. En effet, dans la majorité des cas, cette unité est congruente avec l'Unité Rythmique et constitue donc le domaine de l'allongement final. Lorsqu'elle n'est pas congruente à l'Unité Rythmique, i.e. lorsqu'elle est délimitée à droite par un accent initial secondaire, elle est le domaine d'une autre forme d'accentuation. Ainsi, une catégorie phonologique unique rendrait compte de deux types d'accents (initiaux et finaux) pourtant distincts aussi bien sur le plan fonctionnel que formel. C'est d'ailleurs l'un des arguments qui avaient incité Di Cristo & Hirst (1993) à intégrer à leur modèle le niveau de

l'Unité Rythmique.

Nous verrons, notamment avec l'analyse des configurations tonales, que l'Unité Tonale ne peut pas recouvrir deux réalités physiques totalement différentes. Nos résultats vont d'ailleurs dans le sens des résultats de Jun & Fougeron (1995 ; 1997), qui proposent la représentation phonologique sous-jacente du syntagme accentuel comme étant marqué par un accent initial et final, réalisé ou non en surface, selon les contraintes de débit (Fougeron & Jun, 1996) notamment. Cette conception rejoint d'ailleurs la notion d'arc accentuel de Fónagy (1980), et la notion de marquage initial et final au niveau du mot (bipolarisation accentuelle) de Di Cristo (sous presse).

Selon Jun & Fougeron (1995), le syntagme accentuel doit être représenté dans sa représentation sous-jacente par une suite de tons /LHLH/, et non comme deux groupes /LH/ indépendants. Elles avancent trois arguments majeurs :

- les deux mouvements ascendants (l'un marquant l'accent initial et l'autre l'accent final de groupe) ont des timings similaires, quel que soit le nombre de syllabes dans chaque mouvement (de 200 à 300 ms).
- le temps de chute entre le H initial et le L suivant à l'intérieur d'un syntagme accentuel dépend du nombre de syllabes entre les deux points cibles, alors que le temps de chute entre le H et le L entre deux syntagmes accentuels est constant quel que soit le nombre de syllabes (entre 100 et 200 ms).
- l'association du H initial avec une syllabe est relativement lâche, et les auteurs en déduisent qu'il est associé de manière lâche à la deuxième syllabe du mot lexical.

Le deuxième argument constituerait la preuve d'une cohésion importante à l'intérieur du syntagme accentuel entre l'accent initial et l'accent final, invalidant par là même un niveau inférieur de constituance.

D'autres éléments ayant trait au domaine de l'accentuation sont en faveur de l'abandon de ce niveau d'analyse (Unité Tonale) dans la description du système accentuel du français. Levac et al. (1993) montrent pour le cas du français québécois que le domaine de la proéminence finale en français est plus important que le domaine de la proéminence initiale. Ainsi, les groupes rythmiques et intonatifs seraient associés à un allongement des deux dernières syllabes du constituant prosodique, ce qui constitue un marquage large des proéminences, alors que les proéminences secondaires ne se portent que sur une syllabe, ce qui constitue un marquage étroit de la proéminence. Nos résultats sur l'empan des configurations tonales vont dans ce sens également, mais distinguent le domaine de la proéminence plus en fonction des caractéristiques fonctionnelles des proéminences que de leurs caractéristiques distributionnelles (Astésano et al., 1997b ; 1998).

Par conséquent, si nous acceptons le niveau du Pied Métrique comme constituant à part entière de la hiérarchie prosodique, nous refusons l'Unité Tonale dans sa relation avec la représentation phonologique et phonétique de l'accentuation. Nous étayerons cette prise de position dans les chapitres expérimentaux sur l'organisation temporelle des proéminences et les configurations tonales.

Nous avons jusqu'à présent situé le modèle de Di Cristo & Hirst dans le cadre métrique dont il se réclame. Nous présentons maintenant les catégories accentuelles proposées par le

modèle, et leurs relations à une définition fonctionnelle de l'accentuation.

1.6.6.3. Catégorisation fonctionnelle des accents

Nous avons vu que le modèle de Di Cristo & Hirst est fondé sur une description métrique de l'accentuation du français, prenant en compte les règles d'assignation des accents au niveau phonologique profond. C'est au moment de l'actualisation du message, que les règles rythmiques et sémantico-pragmatiques vont opérer.

La catégorisation fonctionnelle des types d'accents n'est qu'une tentative pour rendre compte de ces règles rythmiques et sémantico-pragmatiques :

– d'un côté, les accents rythmiques (initiaux et finaux), dont la réalisation ne fait pas partie d'un processus conscient et volontaire de la part du locuteur, ont pour fonction de grouper le discours en unités de sens décodables par l'auditeur, en tenant compte des limites de la perception et de la production. Ces accents dépendent des contraintes propres au code de la langue.

– De l'autre, les accents de type sémantico-pragmatique, qui dépendent de facteurs d'assignation des accents nucléaires (focalisation large) et des accents emphatiques (focalisation étroite) : ces facteurs d'assignation sont motivés par le choix qu'effectue le locuteur en fonction de sa visée communicative, du contexte et du développement de l'interaction.

La prééminence, dans cette conception, n'est donc pas simplement la 'promotion phonique d'une syllabe', mais un phénomène complexe qui reflète à la fois l'organisation métrique d'un énoncé et les aspects pragmatiques de l'organisation discursive (Di Cristo, sous presse).

Ce type de modèle tente d'intégrer des considérations linguistiques dépassant le contexte purement syntaxique auquel l'accentuation est traditionnellement liée. De ce fait, il constitue un cadre particulièrement adapté à l'analyse de la parole en situation.

1.6.6.4. Le modèle actuel

Le modèle, dans sa version actuelle, est l'aboutissement de l'évolution du modèle depuis Hirst & Di Cristo (1984) et est présenté dans Di Cristo (sous presse). Pour Di Cristo, l'intégration des catégories accentuelles à l'analyse permet de concilier l'aspect relationnel et l'aspect paradigmatique et hiérarchique de l'accent. Les niveaux hiérarchiques d'accentuation correspondent à des catégories accentuelles qui constituent les têtes de domaines métriques. A chaque niveau de la structure, les unités concernées peuvent être définies par les traits de marquage des têtes qui font intervenir à la fois des aspects tonals, temporels et intensifs.

Le modèle tel qu'il est développé par Di Cristo (sous presse) suppose le *Principe de Bipolarisation Prosodique* qui sous-tend un principe de *Promotion des extrémités* (Fraisse, 1967 ; Hyman, 1975; Fónagy, 1980; Bolinger, 1989). Ce principe s'applique à l'ensemble des unités linguistiques, du mot au tour de parole. Selon ce principe, le mot (unité accentuelle minimale), possède dans sa représentation sous-jacente une prééminence initiale et une prééminence finale. Cependant, la prééminence finale est dominante, et la prééminence initiale, secondaire. Ainsi, l'accent final constitue la marque formelle

obligatoire de la frontière du *Mot Phonologique*. Le gabarit métrique du mot serait, selon ce modèle :

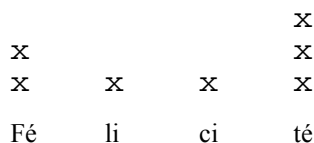


Figure 6 : Représentation métrique du mot « Félicité » (tiré de Di Cristo, sous presse).

Outre le mot (et le mot phonologique), l'Unité Intonative occupe un place importante dans le modèle de Di Cristo, et sa métricité serait motivée essentiellement par des contraintes sémantico-pragmatiques. Elle est le domaine du 'focus d'information', qui conditionne l'assignation de l'accent nucléaire.

Dans sa forme aboutie, le modèle métrique de l'accentuation de Di Cristo propose donc deux niveaux entre la syllabe et le syntagme intonatif (ou Unité Intonative) : le Pied et le Mot Phonologique (ou Mot Prosodique). La figure suivante illustre la constitution de la grille métrique parenthésée, rendant compte des relations hiérarchiques de proéminence pour la phrase : « il a rencontré les écrivains de la contestation » :

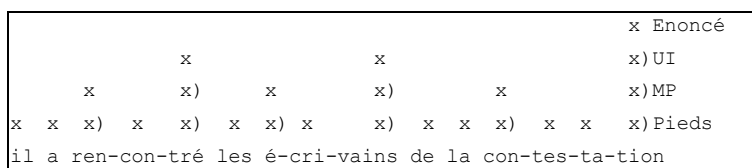


Figure 7 : Représentation métrique d'un énoncé et des constituants métriques de la phrase « il a rencontré les écrivains de la contestation » (Tiré de Di Cristo, sous presse).

Dans cette perspective, le modèle de Di Cristo privilégie l'Unité Intonative, au sein de laquelle les mots accentogènes sont représentés par un gabarit métrique à deux têtes. En ce sens, Di Cristo se démarque de ses recherches précédentes (notamment de Hirst & Di Cristo, 1984), pour lesquelles l'Unité Tonale (et par conséquent le Pied métrique) constituait l'unité métrique privilégiée ; dans sa version actuelle, le modèle met davantage l'accent sur le niveau du Mot Prosodique, comme unité d'actualisation du schéma bipolaire de l'accentuation du français (accentuation initiale, démarcative à gauche de l'unité lexicale, et accentuation finale démarcative à droite de l'unité lexicale). L'équilibrage rythmique ou visée eurythmique se traduit en surface par l'alternance des accents initiaux et finaux, et des syllabes inaccentuées et accentuées, dans le cadre plus large de l'Unité Intonative. Cette dernière est donc d'une part, le lieu de contraintes rythmico-lexicales (phonorythmiques ou phonotactiques), mais est également le lieu de contraintes pragmatiques, telles les accélérations de tempo qui influent notablement sur l'organisation rythmique des énoncés (réduction du nombre de groupes accentuels, inhibition des accents initiaux etc.) (Fougeron & Jun, 1996).

Par conséquent, le modèle de Di Cristo stipule une *dissymétrie fonctionnelle* entre l'accent initial et l'accent final, le premier étant générateur d'emphase, et le deuxième, générateur de frontière. A l'instar de Rossi (1981a) et selon les principes de la théorie

adaptative de Lindblom (1990), l'accent initial emphatique serait la forme de surface marquée (ou la version 'Hyper') de l'accent initial rythmique. Dans ce cas, l'accent initial rythmique contraint pragmatiquement dans le cadre d'une focalisation étroite, est supplanté par un accent emphatique. Pour ce qui est de l'accent final, des contraintes pragmatiques (visée du locuteur et contexte générale de l'interaction) peuvent également jouer, qui, dans le cadre d'une focalisation large, vont conférer à l'accent final le statut d'accent nucléaire.

Enfin, Di Cristo tente, dans une version ultérieure de son modèle, de formaliser le lien entre structure métrique du texte et analyse pragmatique du discours. Cette démarche permettrait de concilier enfin les différentes approches de la prosodie qui ont souvent été mises dos à dos (Di Cristo, sous presse).

« La visée métrique et la visée pragmatique ne sont pas mutuellement exclusives, et (...) l'accentuation constitue précisément un terrain privilégié pour tenter de les mettre en relation dans la perspective d'une approche unificatrice de la prosodie de notre langue ».

C'est en se concentrant plus particulièrement sur la description des liens entre structure informationnelle et procédés de focalisation, que Di Cristo tente cette mise en relation.

Nous avons présenté le cadre théorique dans lequel se situe notre recherche, et nous avons plus particulièrement insisté sur l'accentuation qui constitue selon nous le noyau dur du rythme. Nous nous proposons maintenant de présenter rapidement les notions de timbre et de débit qui constituent des paramètres du rythme mais que nous n'avons pas directement traités dans le cadre de cette recherche. Par contre, les phénomènes de pause seront présentés plus en détail car ils feront l'objet d'une analyse dans le chapitre consacré à l'organisation temporelle des prééminences (cf. p. 259 et suivantes).

2. LE TIMBRE

« [le timbre est la] qualité qui distingue des sons de même hauteur subjective, de même intensité et de même durée et qui est déterminée par l'intensité relative des harmoniques; le timbre peut notamment être sombre ou clair selon la répartition spectrale favorable aux fréquences graves ou aiguës. (...) » (Landercy & Renard, 1977).

Delattre (1966a) attribue au mode tendu, caractéristique du français, la stabilité du timbre des sons au cours de l'articulation. La tension musculaire constante réduit la mimique articulatoire. Ceci explique l'absence de diphtongaison et la neutralisation moins accusée des syllabes atones en français par rapport à l'anglais.

Selon Léon (1992), on peut résumer les grandes lois de la variation des timbres vocaliques du français à partir du fonctionnement accentuel et syllabique. En syllabes inaccentuées, les voyelles tendent à devenir moyennes, bien qu'elles ne soient pas centralisées comme les voyelles de l'anglais dans le même contexte accentuel. En syllabes accentuées, on rencontre deux types de syllabation: syllabation ouverte et fermée. Dans une syllabe accentuée fermée, la voyelle est ouverte et dans une syllabe accentuée ouverte, la voyelle est fermée. Ces deux types de syllabation sont en distribution complémentaire. Léon précise enfin que la perception du timbre s'effectue au niveau des deux premiers

formants alors que le troisième formant rend compte des caractéristiques individuelles du locuteur.

Le timbre n'a pas de rôle distinctif dans l'organisation accentuelle en français. Alors qu'en anglais par exemple, la quantité vocalique détermine le poids métrique d'une syllabe, le système phonologique du français ne connaît pas d'opposition de quantité vocalique pertinente.

Par contre, le renforcement des harmoniques supérieures du spectre des syllabes accentuées semble être une constante à travers les langues. Ce paramètre ('spectral tilt') serait notamment plus fiable que le paramètre de durée comme indicateur de proéminence en parole spontanée (Campbell, 1995b).

3. LE DÉBIT DE PAROLE

Le débit de la parole se mesure en nombre de syllabes par seconde. Il dépend énormément de facteurs individuels. La même phrase peut en effet être découpée en un nombre plus ou moins important d'unités de sens selon le débit du locuteur (Vaissière, 1990).

Le débit a une influence non négligeable sur l'organisation temporelle des proéminences (Vaissière, 1983). Un débit rapide a pour conséquence la réduction des degrés de frontières, voire la suppression pure et simple de certaines frontières, et par conséquent, en français, des accents situés à ces frontières. Cependant, tous les types d'accents ne sont pas forcément atteints par l'accélération du débit. Fougeron & Jun (1996) proposent ainsi que l'accent initial secondaire, qui peut disparaître en débit rapide, n'est pas soumis au même phénomène de replanification et de réorganisation prosodique que les accents situés à des frontières prosodiques. Selon elles, la disparition de l'accent initial en débit rapide serait la conséquence d'un recouvrement des gestes laryngés, i.e. une conséquence directe de contraintes physiologiques, et non une réorganisation phonologique. Leurs résultats confirment donc que la représentation tonale sous-jacente du syntagme accentuel, comportant un accent initial et un accent final, est /LHiLH*/ (cf. « L'unité tonale est-elle adaptée à la description prosodique du français ? », p. 93).

Par contre, leurs résultats les conduisent à conclure que l'accélération du débit affecte l'intonation à la fois dans sa réalisation phonétique et dans son organisation phonologique. Elles constatent une réduction de la dynamique de f0, et la réduction ou disparition de certaines frontières (disparition notamment des accents de syntagmes mineurs, dont elles déduisent qu'ils sont moins importants dans la structuration du discours). La réalisation de certains tons sous-jacents semble donc affectée. Cependant, la réorganisation du groupement prosodique d'une phrase en débit rapide est fortement dépendante des stratégies individuelles des locuteurs.

4. LES PAUSES

4.1. Nature des pauses

Goldman-Eisler a consacré de très nombreux travaux à l'étude des pauses. Dans son ouvrage de 1968, elle définit la pause comme étant une interruption du mouvement articulatoire. Elle distingue trois types de discontinuité de la phonation :

- lors de transitions articulatoires, par exemple quand deux occlusives se suivent. Une interruption est nécessaire pour permettre d'ajuster la position articulatoire.
- lors d'hésitations.
- lorsque la phase d'expiration se termine et que le locuteur a besoin de reprendre sa respiration.

Goldman-Eisler (1968) estime que la deuxième catégorie d'interruptions de la phonation (hésitations) représente des événements phonotactiques qui sont nécessaires au temps de programmation de l'énoncé. Elle les classe en deux catégories: les pauses vides (ou pauses non sonores) et les pauses remplies (ou pauses sonores), elles-mêmes divisées en deux sous-catégories: les hésitations vocales ("hum", [] isolés) et les allongements vocaliques. Mais elle inclut également dans ces événements phonotactiques les pauses de prise de souffle; elle a en effet constaté que ces dernières, tout en correspondant à un besoin physiologique, se situent dans l'énoncé à des endroits où le locuteur doit programmer la suite de son énoncé. En fait, le locuteur profite de la prise de souffle pour chercher ses mots, et inversement, on peut penser que le locuteur qui cherche ses mots en profite pour prendre son souffle.

Grosjean (1980) note que les hésitations qui consistent à prolonger anormalement les syllabes en fin de mot ou les mots monosyllabiques sont propres aux langues à syllabation ouverte, comme le français. Il ajoute que ces phénomènes de syllabes allongées ont en français presque la même ampleur que les hésitations vocales; elles représentent 43% des pauses sonores. Grosjean & Deschamps (1972) étudient l'occurrence des différents types de pauses. Ils constatent que 85% des pauses non sonores en français sont en partie ou en totalité des pauses de prise de souffle; elles ne représentent que 34% en anglais. Cette différence est liée au fait que les suites sonores sont plus longues en français. En ce qui concerne les pauses sonores, 60% des hésitations vocales sont précédées ou suivies d'un autre type de pause. En 1980, Grosjean compare les différents types de pauses du français et de l'anglais en parole spontanée. Il montre de nombreuses similitudes entre les deux langues, notamment en ce qui concerne le nombre total d'hésitations et le pourcentage de pauses remplies en fin et à l'intérieur des énoncés. Cependant, l'utilisation des deux types de pauses sonores (hésitations vocales et syllabes allongées) est différente d'une langue à l'autre. Grosjean émet l'hypothèse que l'anglais étant une langue à syllabation fermée, elle utilisera moins de syllabes allongées que le français mais compensera par l'utilisation plus systématique d'hésitations vocales.

Guaïtella (1991) distingue deux types de pauses : la pause (interruption *vocale*), qui est un fait segmental directement mesurable, et l'hésitation vocale (interruption *verbale*), qui

est un fait linguistique. Toutes deux peuvent avoir des fonctions similaires, notamment de programmation des segments de parole en parole spontanée. D'un point de vue formel, l'hésitation vocale a été décrite par Guaïtella comme une 'unité' rythmique de parole spontanée à part entière. Elle se caractérise par un allongement significatif de durée, qui la situe entre la durée syllabique moyenne et la durée des pauses silencieuses, ainsi que par un contour de f₀ plat.

4.2. Distribution et fonctions des pauses

Selon Fraise (1956), les pauses lient les unités rythmiques successives et participent à l'organisation unitaire des groupes rythmiques. Park (1986) définit d'ailleurs le groupe rythmique comme étant un groupe de phonation délimité par deux pauses externes. Les pauses ont donc un rôle démarcatif puisqu'elles sont exclues à l'intérieur d'un groupe rythmique.

Pour Duez (1987), la pause en français fait partie de l'allongement final, c'est à dire que la pause non sonore et l'allongement final relèvent selon elle d'un même phénomène phonotactique sous-jacent. C'est ainsi que les auditeurs tendent à percevoir une pause après une syllabe allongée même s'il n'y a pas trace physique d'un silence (Repp, 1992). Ceci est donc bien la preuve que les deux phénomènes sont équivalents sur le plan fonctionnel. Certains auteurs rejettent pourtant le rôle phonotactique de la pause silencieuse (Llorca, 1987). L'auteur montre que les silences sont utilisés pour renforcer les frontières entre les unités rythmiques; ils ont par conséquent un caractère redondant puisque ces frontières sont déjà marquées par des traits mélodiques, de durée ou de f₀. Elle en conclut donc que les pauses silencieuses sont certainement facultatives dans toutes les langues, à l'exception des pauses respiratoires, qu'elle inclut dans cette catégorie, et qui répondent à une nécessité physiologique.

Goldman-Eisler (1968) suggère que les pauses ont deux fonctions principales: tout d'abord, elles participent à la planification et à l'organisation des énoncés. Ensuite, elles permettent la prise de souffle. Les pauses de la première catégorie sont les plus nombreuses, ce qui laisse supposer que la fonction de planification et d'organisation est plus importante que la nécessité biologique de respiration.

Grosjean & Deschamps (1972) font une analyse statistique en parole spontanée de la distribution des différents types de pauses en français et constatent que les trois quarts des pauses interviennent à des coupures grammaticales. Cependant, 38% des pauses sonores sont réalisées en dehors des points de rupture grammaticaux, contre 25% des pauses non sonores. Goldman-Eisler (1968) a réalisé la même étude pour l'anglais, en parole spontanée également. Elle note qu'en parole spontanée, seulement 55% des pauses se trouvent à des jonctions grammaticales, ce qui laisse 45% de pauses réalisées à des points de rupture non grammaticaux. Elle en déduit que cela peut nuire au décodage de l'énoncé par l'auditeur, alors que l'on s'attendrait plutôt à ce que les pauses aient pour fonction de faciliter le décodage, en délimitant les groupes rythmiques. Au contraire, une grande proportion de pauses dans le discours spontané ne correspond pas à la structure linguistique de l'énoncé et n'aide pas à la communication. En fait, les pauses peuvent avoir une fonction rhétorique lorsqu'elles participent au groupement des unités linguistiques, et dans ce cas, elles facilitent la communication. Par contre, lorsque l'activité cognitive est intense, les pauses deviennent un aspect symptomatique du langage et ne servent aucune fonction dans le

décodage.

Grosjean & Deschamps (1975) montrent que pour l'anglais et le français, la répartition des pauses silencieuses (en nombre et en durée de pauses) se fera là où la structure syntaxique le permettra alors que la répartition des différents types de pauses sonores se fera par rapport à la phonotactique de la langue. Enfin, Grosjean (1980) met en évidence une différence essentielle entre l'anglais et le français en ce qui concerne la distribution des pauses. En anglais, l'éventail d'emplacement des pauses à l'intérieur de la phrase est plus large; on trouve notamment des pauses à l'intérieur des syntagmes verbaux, qu'il est impossible de rencontrer en français. Selon Grosjean, le nombre élevé de verbes prépositionnels favoriserait naturellement ces arrêts.

Les pauses ont bien évidemment un rôle rythmique majeur. Les travaux de Fant et al. (1987 ; 1990 ; 1991 a et b) indiquent que les pauses entretiennent un rapport entier (1, 2 ou 3, en termes de durée) avec les groupes rythmiques ou intervalles interaccentuels. Également, selon eux, le rythme de la parole serait plus prégnant dans les moments de pauses que dans les moments de verbalisation. A ce sujet, ils rejoignent les auteurs qui considèrent la pause comme un 'battement' rythmique à part entière (Abercrombie, 1964 ; Cutler, 1980 ; Repp, 1992 etc.). Nous verrons (cf. § « Pauses, unités rythmiques de référence », p. 272) que cette tendance existe quel que soit le style de parole, bien que dans des proportions spécifiques aux types de parole.

La pause remplit également une fonction de planification et certaines études permettent de mettre en évidence la fonction de démarcation des pauses, délimitant une unité de performance (le groupe de souffle). Ainsi, les parties de discours situées entre deux pauses silencieuses montrent une cohésion rythmique certaine (groupes rythmiques de tailles comparables, participant à une sensation d'équilibrage rythmique à l'intérieur de ces unités de performance), alors que les groupes rythmiques adjacents séparés par une pause sont significativement distingués sur la base de leur taille et leur durée (Wenk & Wioland, 1984).

4.3. Durée des pauses

Goldman-Eisler (1968) a fixé la durée minimale d'une pause silencieuse à 250 ms; en deçà, cette durée peut correspondre à la période de tenue d'une occlusive. Selon elle, la durée des pauses peut varier énormément. Elle est symptomatique des différences individuelles et dépend de la pression des interactions sociales. Le degré de familiarité avec le sujet réduit plus ou moins la durée et la fréquence des pauses. En général, les pauses ne dépassent jamais 3 s. et 99% des pauses sont même inférieures à 2 s. Les pauses respiratoires vont de 0,5 à 1 s. et occupent de 2,5 à 25 % du temps de locution. Les autres types de pauses (sonores et silencieuses) occupent de 40 à 50 % du temps total de locution.

Grosjean (1980) note que les suites sonores sont en moyenne plus longues en français qu'en anglais, mais cette différence est compensée par des pauses plus longues en français. Si l'on note beaucoup de similitudes dans la distribution et la fréquence des pauses dans les deux langues, le temps de pause est organisé différemment en anglais et en français. L'anglais fait plus de pauses et de durée moins importante que le français car il est plus concis.

4.4. Pauses et perception des unités linguistiques

Fraisse (1956) pense que les pauses influent au niveau de la perception du rythme: un intervalle trop court pourrait en effet entraîner des confusions en devenant un temps de la structure. De plus, la séparation temporelle plus grande des unités successives peut augmenter l'impression de proximité entre les éléments du groupe. Guaitella (1986) avance même que les pauses peuvent parfois jouer un rôle dans l'isochronie des mots prosodiques. Les 'battements silencieux' que constituent les pauses permettent de conserver une pulsation accentuelle entre deux séquences de parole temporellement non cohésives. Elles participent donc pleinement à l'équilibrage rythmiques des séquences de parole (Couper-Kuhlen, 1993).

Dans le discours spontané, lorsque l'activité cognitive est importante, les pauses peuvent intervenir de manière erratique et par conséquent nuire à la communication (Goldman-Eisler, 1968). L'auteur conclut que la parole spontanée étant le plus souvent "découpée" en séquences de 5 ou 6 mots, le flot de la parole et la facilité d'élocution ne sont qu'illusions perceptives.

La perception des pauses est facilitée si elles interviennent à des jonctions de syntagmes ou d'énoncé ; par contre, si elles interviennent à des endroits non attendus, elles sont moins souvent perçues. Ceci vaut aussi bien pour la parole que pour la musique (Repp, 1992). Ceci expliquerait pourquoi en parole spontanée, les hésitations vocales ou les pauses intervenant à des endroits non attendus (au milieu d'un groupe clitique par exemple) sont souvent non perçues ; par conséquent, nous estimons que le postulat de Goldman-Eisler (1968) selon laquelle les pauses erratiques en parole spontanée peuvent nuire à la communication, n'est pas entièrement recevable.

Wenk & Wioland (1982) proposent des principes rythmiques relatifs aux placement des pauses pour rendre compte de la spécificité du rythme du français, parmi lesquels le *Principe du placement des pauses*.

Dans les langues codachrones, une pause peut survenir entre 2 groupes rythmiques, et au début et/ou à la fin d'un groupe rythmique contenant nettement moins de syllabes que les groupes environnants. Il est frappant de constater le rôle rythmique que jouent les pauses à différents niveaux : au niveau syntaxique (marquage des frontières) ; au niveau du discours (marquage des tours de parole, et des unités du discours) ; au niveau stylistique (dénotation de la profondeur, du contrôle du discours...); aux niveaux biologique (respiration) et psycholinguistique, enfin. Sur le plan psycholinguistique, les pauses permettent la planification et l'exécution de la parole, et s'expriment en termes de cycles d'hésitation (conceptualisation) et de fluidité verbale ; c'est ce que Goldman-Eisler (1968) appelle le « rythme cognitif ».

Les pauses ont non seulement une fonction de démarcation, mais également une fonction de compensation, afin de maintenir l'impression d'isochronie lorsque des groupes rythmiques voisins ont un nombre de syllabes très différents.

Nous avons présenté les notions fondamentales liées au concept de *Rythme*. Dans cette perspective, nous avons porté une attention particulière au cadre théorique dans lequel se situe le modèle de l'accentuation de Di Cristo & Hirst. Ce modèle se réclame de l'approche métrique de l'accentuation, mais nous avons vu que Di Cristo tente de faire le lien, dans la

version actuelle du modèle, avec les propositions émanant des théories du discours fondées notamment sur une approche pragmatique. Une originalité de notre travail réside dans la prise en considération de divers types de discours dans l'analyse des faits accentuels en français. Avec une telle approche, nous visons à mettre en évidence la part de variabilité imputable aux diverses situations de communication, ainsi que les invariants rythmiques propres au système linguistique.

5. RESUME

Nous présentons les différents paramètres constitutifs du rythme linguistique, en insistant particulièrement sur l'accentuation qui représente l'essentiel de notre investigation. Tout en élargissant notre champ d'observation aux aspects rythmiques et accentuels de différentes langues, nous centrons notre réflexion sur l'accentuation du français. Nous montrons que les croyances selon lesquelles le français ne possède pas de système accentuel ne résistent pas aux arguments développés dans le cadre de la théorie métrique (Hirst & Di Cristo, 1984 ; Ladd, 1996 ; Di Cristo & Hirst, 1997). De plus, des analyses perceptives indiquent que la métrique du français (pieds iambiques) favorise la perception de la proéminence marquée essentiellement par la durée (Nord et al., 1990), paramètre accentuel important en français. Si, comme le montre Rossi (1972), le seuil de perception de la durée est universellement plus fin dans la condition où l'étalon (syllabe inaccentuée) précède la variable (syllabe accentuée), on peut considérer que le poids métrique de la syllabe est renforcé en position finale. La réalité perceptive étaye donc notre choix d'une analyse métrique de l'accentuation en français.

Le modèle de l'accentuation que nous avons choisi d'appliquer à nos analyses est présenté dans ce chapitre, dans le contexte théorique dont il est issu (théorie métrique). Ce modèle (Di Cristo & Hirst, 1997 ; Di Cristo, sous presse) propose de catégoriser les proéminences accentuelles sur le plan fonctionnel, afin de rendre compte des règles *rythmiques* participant du code de la langue et des règles *sémantico-pragmatiques* dépendant de la visée communicative du locuteur. La proéminence n'est donc pas simplement la 'promotion' phonique d'une syllabe mais un phénomène complexe qui reflète à la fois l'organisation métrique d'un énoncé et les aspects pragmatiques de l'organisation discursive.

Ce modèle, dans sa version actuelle, propose un « Principe de Bipolarisation accentuelle » qui suppose que le mot possède dans sa représentation sous-jacente une proéminence initiale et une proéminence finale. Dans cette perspective, le Mot Prosodique est l'unité d'actualisation du schéma bipolaire de l'accentuation du français, qui s'apparente clairement à la notion d'« Arc Accentuel » proposé par Fónagy (1980). En ce sens, le modèle de Di Cristo & Hirst prend en considération la réalité d'une accentuation initiale en français, qui coexiste avec l'accentuation finale traditionnelle. Le modèle permet donc de rendre compte de la fonction hiérarchique de l'accentuation (relations entre les proéminences) ainsi que de la fonction d'intégration rythmique des unités lexicales et sémantiques dans le cadre de l'arc accentuel (Fónagy, 1980 ; Di Cristo, sous presse).

- *CHAPITRE 3* -

Variabilité et Discours

1. INTRODUCTION

L'un des buts fondamentaux des recherches en phonétique est de mettre en évidence les structures invariantes dans chaque langue (Kohler, 1986). Cette approche s'inscrit notamment dans la tradition de la grammaire générative qui s'oppose à une démarche inductive visant à établir des généralisations sur la base de l'observation et de la comparaison d'un échantillon des faits de langue le plus large possible. Une démarche hypothético-déductive présuppose l'existence de catégories universelles caractérisant un système noyau.

C'est cette démarche que nous avons choisi d'adopter au travers d'une approche métrique de l'accentuation du français. Il s'agissait dans un premier temps de distinguer entre la variabilité, source de 'bruit' dans le système linguistique (variabilité aléatoire), et la variabilité structurelle (individuelle, contextuelle, stylistique etc.) (pour une discussion, voir Rossi, 1985a). Dans la perspective qui nous occupe, les phénomènes de variation de durée intrinsèque des phonèmes ou de caractérisation du locuteur (sur la base de leur f0 usuelle notamment) ne constituent pas une variabilité pertinente (cf. § « Pourquoi normaliser ? », p. 175). Notre but est de procéder à un examen de la variabilité accentuelle et rythmique pertinente dans la communication, mais également de mettre en évidence la part de variabilité participant de la caractérisation d'un type de discours. Le choix du modèle métrique de l'accentuation du français s'inscrit dans cette recherche de caractérisation de la variabilité stylistique, autour d'un système que l'on présuppose noyau.

Notre tâche est d'autant plus complexe que le caractère mobile et probabilitaire de l'accent français encourage une variabilité individuelle ainsi que des « *déviation*s considérables à partir d'une 'norme' floue » (Fónagy, 1980 : 162).

Nous proposons dans ce chapitre de faire le point sur les notions d'invariance et de variabilité, et sur les notions de normes et d'écarts à la norme qui en découlent partiellement. Nous tâcherons dans un deuxième temps de nous situer par rapport à une définition des styles de parole, en passant en revue un certain nombre d'études portant sur la caractérisation linguistique, phonétique et prosodique des types de discours. Nous insisterons plus particulièrement à cette occasion sur le style journalistique, qui constitue une partie de notre corpus, et qui présente l'intérêt, selon nous, d'ancrer l'évolution du système accentuel du français dans l'usage. De plus, ce style de parole suppose un encodage particulier du texte qui le situe dans une position intermédiaire, sur une échelle de spontanéité, entre la parole lue et la parole spontanée (cf. § « Choix des locuteurs et des types de discours », p. 135).

2. INVARIANCE ET VARIABILITE

Le lien existant entre les notions de variation et d'invariance est central en linguistique, comme il peut l'être dans d'autres disciplines scientifiques (la biologie par exemple). C'est en effet à travers l'analyse des phénomènes linguistiques que l'on suppose propre au système linguistique, et l'analyse des 'écarts' par rapport au système noyau, que l'on tente de comprendre les mécanismes linguistiques. C'est également à partir de ces observations que l'expérimentateur/théoricien peut envisager de formaliser un système linguistique.

« (...) à travers la façon dont sont appréhendés les variations et les invariants, c'est en effet la problématique générale des liens entre langage, pensée et cognition qui se trouve engagée. » (Fuchs, 1997 : 13).

A ce sujet, les recherches sur les pathologies du langage ou encore sur l'acquisition du langage sont centrales dans la compréhension des mécanismes linguistiques. L'observation de l'acquisition des formes linguistiques chez l'enfant, par le jeu incessant de l'essai/erreur permettant d'atteindre ce que l'enfant estime être les caractéristiques de son système linguistique, fournit des renseignements précieux sur les mécanismes cognitifs de construction du langage, et également sur les particularités du système linguistique cible. Ainsi, Konopczynski (1986) observe une variabilité des formes linguistiques plus importante dans le jais que dans le proto-langage, variabilité qui se réduit encore à l'âge adulte pour atteindre les formes canoniques du système linguistique.

Les notions d'invariants et de variabilité peuvent cependant être assimilées aux notions de 'norme' et d'écart à la norme, dont les travaux en stylistique rendent compte traditionnellement. La linguistique historique, lorsqu'elle a commencé à étudier dans le détail les transformations du langage, a montré que l'évolution de la langue a fréquemment son origine dans des façons de parler populaires, argotiques ou 'patoisantes' ; de sorte que la correction d'une époque ne fait souvent que consacrer les 'incorections' de l'époque précédente. Il y a donc une certaine réticence de la part des utilisateurs de la langue, qui induit un décalage dans la reconnaissance de l'évolution de la langue. Cette 'inertie' linguistique, aggravée par la lourde tradition normative en France, se retrouve dans les travaux de linguistique et phonétique générale (cf. § « L'organisation accentuelle en français », p. 78). C'est ainsi que le concept de norme a pu influencer les recherches des phonéticiens, lorsqu'ils considèrent par exemple la lecture comme une 'marque neutre' de discours. Toute autre forme de discours ne serait pas représentative du système prosodique d'une langue, les 'écarts' par rapport à la norme discursive étant jugés trop importants. Or, comme le font remarquer Ducrot & Schaeffer (1995), tout type de discours constitue une forme de marquage.

« Si la rupture introduit un marquage, la systématisme, au-delà d'une certaine limite, fait de même. » Ducrot & Schaeffer, 1995, p. 157.

Dès lors, le concept de parole 'neutre' s'évanouit. La variabilité discursive, loin de constituer une source de 'bruit', doit être considérée sous l'angle des variations dialectales, idiolectales, sociolectales, situationnelles etc., constitutives à part entière d'un système linguistique. De même, le marquage stylistique ne doit plus être défini par rapport à une norme (lecture, parole de laboratoire), qui impliquerait de le considérer comme un écart, une déviance par rapport à cette norme. Il doit être considéré avant tout comme une alternative résultant de la situation de communication.

*« Les notions de **choix** et de **variation** stylistiques, qui jouent un rôle important dans la stylistique générale, où elles sont utilisées notamment pour étudier les divers niveaux et registres discursifs présents dans une langue donnée, semblent plus prometteuses que la notion d'écart, ne serait-ce que parce qu'elles traitent les différenciations stylistiques comme des dimensions inhérentes à l'activité discursive plutôt que comme des éléments surajoutés à une base neutre. » Ducrot & Schaeffer, 1995, p. 158.*

Nous considérons pour notre part que les invariants rythmiques et accentuels constituent le système noyau de la langue. En tant que tels, ils vont être constitutifs de la langue parlée 'dans tous ses états'. La variabilité rencontrée représenterait les formes allophoniques, pertinentes sur le plan linguistique, de cette structure noyau. Mais elle constituerait également une trace d'appartenance à un groupe sociolectal, socioprofessionnel etc., ou serait tout simplement révélatrice d'une situation de communication (conversation, débat, dispute, entretien, conférence etc.). La variabilité observée à travers les différents styles de parole est cependant, dans une certaine mesure, assez systématisable, car elle est principalement le résultat d'une adaptation à un code linguistique, de manière inconsciente chez le locuteur.

« En définitive, malgré les aspects bien souvent très fluctuants (et parfois déroutants) de la variabilité prosodique observée, on peut dire qu'il y a, dans une large mesure, systématisation des variations. Celle-ci est probablement en partie inconsciente, mais permet, chez les locuteurs linguistiquement les plus favorisés, une adaptation « d'instinct » à un schéma prosodique lié à la nature situationnelle du message à transmettre. Ces schémas prosodiques sont à rattacher aux mécanismes de la verbalisation. » (Lucci, 1983 : 329).

Les mécanismes de la verbalisation auxquels Lucci fait référence désignent en fait les mécanismes cognitifs entrant en jeu dans le processus d'encodage de la langue. Ces processus seraient différents suivant la situation de communication, plus particulièrement dans les cas de lecture et de parole spontanée, que l'on présente souvent comme deux cas d'école radicalement opposés.

3. CARACTERISATION DES TYPES DE DISCOURS

L'analyse de corpus spontanés se démarque des analyses traditionnelles sur la parole contrôlée créée afin de valider les modèles théoriques. Les études sur la parole contrôlée ont apporté des connaissances considérables sur les mécanismes linguistiques. Cependant, nous pensons qu'il est temps de passer à l'analyse de la parole en situation (parole spontanée, parole réelle), afin d'étendre le champ d'investigation. Ce genre d'analyse peut de surcroît apporter un éclairage différent sur les notions de variabilité et d'invariance. Néanmoins, la parole spontanée introduit un grand nombre de degrés de liberté. Se pose alors le problème du contrôle des variables, qui nécessite le traitement d'un grand nombre de données. Cet obstacle est réduit en parole de laboratoire contrôlée, ce qui explique que l'on ait longtemps privilégié ce type de parole. Nous verrons cependant que certaines méthodes de normalisation des données, appliquées à la parole spontanée, permettent de s'affranchir, dans une certaine mesure, d'une part de variabilité (cf. § « Pourquoi normaliser ? », p. 175).

Comme le propose Lucci (1983), la vraie nature d'une langue (dans toute sa complexité et son dynamisme) ne peut être décrite que par le biais d'analyse de corpus non contrôlés. En ce sens, bien entendu, ce genre de description sera éloignée des descriptions normatives d'un français figé « *apparemment immuable et cohérent dans sa simplicité illusoire.* » (Lucci, 1983 : 12).

La description d'une langue en situation peut cependant rebuter le chercheur et induit

des comportements pusillanimes. Sans doute, cette frilosité trouve-t-elle sa source dans la crainte de découvrir, dans la réalité linguistique, des événements ne rentrant pas dans un cadre descriptif préétabli ; dans une grande mesure, l'étendue et la complexité de telles descriptions empiriques doit également entrer en ligne de compte. Certains auteurs se réfugient ainsi derrière le caractère apparemment 'phénoménologique' voire 'anecdotique' de telles entreprises. Cette attitude reflète l'ampleur de la mésestime que nourrissent certains chercheurs, appartenant le plus souvent à la tradition générativiste, à l'égard des recherches en sociolinguistique, pragmatique, ou tout simplement en phonétique et phonologie expérimentale.

Cette réticence endémique est à l'origine de la négation de l'évolution du système accentuel du français et de sa marginalisation. Elle a même été décrite par Fónagy (1980) en termes psychanalytiques, comme reflétant une certaine crainte vis à vis du changement du système linguistique « *susceptible de créer un sentiment d'insécurité, un certain malaise* » (Fónagy, 1980 : 176).

Si on reconnaît néanmoins aujourd'hui l'intérêt d'étudier la parole en situation, notamment pour les applications en synthèse et en reconnaissance de la parole (Sagisaka et al., 1997), un certain flou demeure autour de la notion de parole spontanée. On en donne très souvent des définitions par la négative, en la contrastant à la parole lue. De là découle une dichotomisation des types de discours, ignorant le plus souvent les multiples dimensions qui entrent en ligne de compte dans leur caractérisation.

3.1. Parole lue vs. parole spontanée

3.1.1. Parole lue et parole spontanée : deux processus d'encodage différents

L'opposition entre parole lue et parole spontanée peut être abordée dans un premier temps d'un point de vue cognitif. Les différences prosodiques entre la parole lue et la parole spontanée seraient directement imputables à la différence de processus d'encodage à l'origine de ces formes d'oralisation (Guaïtella, 1991). En parole spontanée, l'étape de conceptualisation est simultanée à celle d'oralisation, alors qu'en parole lue, la conceptualisation précède l'étape d'oralisation. L'intégration visuelle du texte en parole lue par groupes correspondant à l'empan de la mémoire visuelle, crée un rythme propre à la lecture. Selon Guaïtella par conséquent, le rythme de l'oralisation va être influencé par le rythme visuel de lecture, ainsi que par la ponctuation, alors que le rythme de la parole spontanée sera, lui, déterminé par l'activité de conceptualisation. Le rythme parolier serait donc un « *rythme de concepts* » (Borel-Maisonny, 1967 ; Goldman-Eisler, 1968). Cette activité de conceptualisation en parole spontanée est généralement identifiable dans le discours par des périodes d'hésitation, de répétition, de 'remplissage' verbal (notamment par des formes verbales canoniques) et par des pauses silencieuses.

Cette conception duale des types de discours qui sous-tend l'existence de deux systèmes organisateurs différents, se retrouve également dans les travaux de phonostylistique.

« Le jeu de la récupération sémantique de l'intonation est très important lorsque disparaissent les articulations syntaxiques à fonction logique, dans les énoncés du parler spontané. En fait, si l'on compare des phrases isolées du type « bien formé » avec des énoncés de parole spontanée, on arrive

même à la conclusion qu'on a affaire à des systèmes dont les traits prosodiques diffèrent pratiquement tous. » (Léon, 1993: 148).

Cependant, on peut se demander à juste titre si le rythme des différents types de discours est la trace de systèmes organisateurs différents (Guaitella, 1991), ou bien si le style de parole ne serait qu'un phénomène phonétique de surface (Barry, 1995), ce qui impliquerait l'existence d'un système noyau dans la représentation mentale sous-jacente (Di Cristo, sous presse).

3.1.2. Comment définir la parole spontanée ?

Souvent, les chercheurs donnent une définition par la négative de la parole spontanée : est spontané tout ce qui n'est pas lu.

« The 'spontaneous speech' which we want to model is speech that is not read to script » (Beckman, 1997 : 7).

L'inventaire que Beckman donne des techniques mises en place par les linguistes et phonéticiens pour tenter de faire produire des énoncés 'spontanés' aux locuteurs est d'ailleurs éloquent. Parmi ces procédés, on trouve les techniques de Map Tasks et du Wizard of Oz¹ notamment. Il est surprenant que l'auteur n'envisage pas les possibilités d'enregistrement en chambre anéchoïque de conversations ou de discussions, ainsi que la possibilité de 'piocher' sur les ondes radios ou télévisuelles une multitude de types de discours (interview, journaux, publicités, débats, etc.) produits dans des conditions d'enregistrement acceptables.

Il nous semble que les corpus 'spontanés' restreints proposés par Beckman sont le reflet d'une certaine crainte vis à vis de la diversité des formes que présente la parole spontanée dans de véritables situations de communication. L'immense variété des situations de communication et leurs implications pragmatiques, linguistiques et prosodiques complexifie en effet grandement la tâche des phonéticiens. Cependant, certains travaux en linguistique ont permis de cerner davantage les différentes facettes des types de discours.

Dès lors, la distinction traditionnelle entre parole lue et parole spontanée peut être abandonnée pour laisser la place à une nouvelle caractérisation des styles de parole sur des dimensions multiples, en référence aux participants du discours, leur nombre et leur relation, ainsi qu'aux besoins de la communication elle-même. Ces principales révisions sont dues en grande partie à des linguistes tels Eskenazi (1993) (cf. § « Notions de 'styles de parole' et de 'types de discours' », p. 130.).

Pour notre part, nous définissons la parole spontanée comme un type de discours pour lequel le locuteur construit son discours au fur et à mesure de l'émission, sans support écrit ou mémorisation préalable d'un texte. A partir de là, des considérations telles que la connaissance du sujet, la relation avec les interactants, l'état émotionnel du locuteur, son implication et sa visée communicative ainsi que la situation de communication, vont

¹ La technique de Map Task consiste à donner un plan (de ville par exemple) au locuteur, dont la tâche est ensuite de décrire un parcours à un auditeur (présent ou absent). Pour la technique du Wizard of Oz, l'expérimentateur demande au locuteur de dialoguer avec un ordinateur, dont l'expérimentateur simule les réponses.

participer à la catégorisation d'un type de discours sur une échelle de spontanéité.

3.1.3. *Intérêts de l'analyse de types de discours spontanés*

Un des arguments en faveur de l'étude de la parole spontanée réside dans la difficulté à reproduire de manière artificielle les paramètres prosodiques de focus dans leur diversité, ainsi que tout ce qui a trait aux émotions dans le discours. Également, l'intérêt d'étudier la parole en situation est de prendre en compte les variations dialectales, afin d'adapter les systèmes de synthèse et de reconnaissance.

Pour Vihanta (1993), la parole spontanée a l'avantage de permettre d'étudier les préférences du locuteur quant à l'utilisation de tel ou tel procédé prosodique. Pour lui en effet, le fait de générer le message plus ou moins simultanément avec sa réalisation, ouvre des perspectives beaucoup plus vastes pour la recherche de l'interaction des moyens syntaxiques et prosodiques dans l'organisation du message. Cependant, Vihanta insiste sur le fait que la parole spontanée recouvre des réalités diverses et variées : différents niveaux de registres, de variétés socio-situationnelles etc. Les situations de communication revêtent également une importance particulière : un discours public, même s'il est présenté sans texte, a un degré moindre de spontanéité qu'une conversation entre amis, car il aura été préparé, voire partiellement rédigé à l'avance, et certaines parties auront été probablement apprises par cœur. Enfin, les moyens prosodiques utilisés pour rendre compte des échanges linguistiques (prise, maintien ou cession de parole, et plus généralement les indices de niveau conversationnel) entre participants d'une conversation (ou d'un débat etc.), sont autant de procédés prosodiques qui viennent se surajouter aux procédés de hiérarchisation et de structuration du message.

Cette complexité définitionnelle amène Vihanta à mettre en doute l'efficacité des modèles théoriques dans la tâche de description des systèmes prosodiques de divers types de discours :

« Quand on a compris la diversité immense et étonnante de la langue parlée, on est obligé d'avouer que c'est une illusion que de pouvoir décrire la langue parlée à l'aide d'un seul modèle monolithique basé sur des réflexions théoriques. Il est clair que les moyens prosodiques et syntaxiques utilisés par le locuteur varient selon le genre, le registre et le style du discours, et ainsi la structuration et l'organisation du message se font suivant des règles différentes dans des situations différentes. Ces règles ne peuvent être décrites d'une façon fiable qu'à partir des analyses de corpus de discours authentiques les plus variés. Le travail s'annonce de longue haleine et la modestie s'impose. » (Vihanta, 1993 : 265).

Vihanta (1993) compare un type de discours 'interview' où le locuteur est interrogé en tant que spécialiste de la question, avec le style journalistique. Selon lui, la différence majeure entre les deux interventions réside dans le fait que le journaliste professionnel ne se trouve pas personnellement impliqué par ses propos et présente de manière objective et neutre son texte, alors que le premier (interview) s'engage dans les propos qui sont les siens. Ce degré d'implication aurait des conséquences sur la prosodie.

Nous pensons cependant que les différences prosodiques ne doivent pas exclusivement être le reflet du degré d'engagement et d'implication de la part du locuteur, mais également

de la visée informationnelle du discours. Il ne faut pas oublier que les locuteurs journalistes professionnels mettent un point d'honneur à adapter l'organisation sémantique, syntaxique et prosodique de leur discours à un type de discours oral, qui tendrait donc plutôt vers une imitation des caractéristiques de la parole spontanée.

Le style journalistique a fait l'objet d'un certain nombre de recherches fort intéressantes. Cependant, nous pensons que les descriptions de ce type de discours n'ont pas toujours pris en considération les aspects hybrides qui constituent la spécificité de ce style de parole.

3.1.4. Style journalistique : hétérogénéité linguistique

Le style journalistique emprunte à la fois les caractéristiques des styles de lecture et de parole spontanée. Selon les professionnels de la communication, les journalistes de radio ou télévision rédigent leurs discours au préalable, mais en s'appliquant à transcrire une façon de parler 'oralisée'. Ils s'imprègnent ensuite de leur texte et le reproduisent oralement, selon un principe équivalent à la récitation. Ils disposent bien évidemment de notes écrites, mais s'éloignent autant que possible de la lecture. Le style journalistique est donc dual : le texte est transcrit puis appris, mais se rapproche d'un point de vue syntaxique et prosodique (taille des groupes de mots etc.) de la parole spontanée (cf. § « Présentation du corpus », p. 135). Cette caractéristique duale du style journalistique n'apparaît pas dans la définition qu'en donne Vihanta (1991) :

« Dans ce type de discours, il s'agit presque uniquement de la lecture d'un texte rédigé à l'avance, dans lequel la structuration syntaxique a déjà été effectuée au préalable, et qui est formé principalement sur le modèle de la langue écrite. La seule liberté, toute relative, qui reste au locuteur, au moment de la production du message, est celle d'utiliser les moyens prosodiques, dans le cadre syntaxique donné, pour hiérarchiser l'information. » (Vihanta, 1991 : 422).

C'est à travers ce type de discours que l'évolution prosodique du français a été le plus souvent décrite. Il apparaît en effet ancrer l'évolution accentuelle barytonique du français dans l'usage. On peut se demander quelles en sont les raisons.

« L'accent d'insistance semble bien avoir de nos jours une fréquence élevée dans le discours. Quelles sont les causes de cette fréquence élevée ? On ne peut se hasarder qu'à quelques suppositions. Peut-être sont-elles sociolinguistiques. Le rythme accentué, saccadé, violent même de notre vie moderne pourrait-il avoir une influence sur le comportement des sujets parlants ? La grande place prise également par les moyens audiovisuels d'information - radio et télévision - de même qu'ils tendent à un certain 'nivellement' de la prononciation, peuvent aussi privilégier un certain style fortement impressif. On ne peut émettre que des hypothèses, mais il n'en reste pas moins que l'insistance est de plus en plus employée et qu'elle tend à entrer dans le système de la langue. » (Séguinot, 1976 : 45).

Léon (1993) et Fónagy & Fónagy (1976) remarquent également que l'accent 'didactique' prolifère dans ce style de parole, marqué non seulement par une accentuation forte mais aussi par les coups de glotte. Certains linguistes considèrent qu'il s'agit d'une marque

d'identité socioprofessionnelle. Nous pensons que le qualificatif d'accent 'didactique' est particulièrement adapté à ce type de discours, plus que le terme 'd'insistance', qui peut revêtir une multitude de fonctions (cf. § « L'accentuation initiale : évolution du système accentuel en français », p. 80).

Dans le style de parole journalistique, il apparaît que l'accentuation des enclitiques est caractéristique, et correspond à une contrainte métrique de ce style. Vihanta (1993) se demande néanmoins si le rythme, au lieu d'être la cause de ce type d'accentuation, n'en est pas au contraire le résultat. L'auteur explique cette tendance à l'accentuation des enclitiques comme une aide au décodage de la structure informationnelle du discours, qui, fondée sur des critères de l'écrit, serait trop difficile à décoder pour les auditeurs, à moins d'une aide particulière au décodage (accentuation des conjonctions, les subordinations etc.). Nous verrons dans le chapitre sur l'organisation temporelle des prééminences (cf. Chapitre 7 - L'organisation temporelle de la parole - , p. 259) qu'il s'agit sans doute plutôt d'un marquage particulier des lexèmes (marquage gauche, comme frontière lexicale). Dans sa fonction informationnelle et démarcative, il semblerait en effet que ce type de discours met en relief le mot plutôt que le syntagme (Léon, 1993).

Les auteurs qui ont décrit ce type de discours sont quasiment unanimes sur le caractère irritant, déplaisant, de la prosodie des annonceurs de radio ou de télévision.

La description décapante qu'en fait Bolinger (1989), est plutôt ironique.

« (...) The free-wheeling part of the brain makes its best guess, which is to accent a word if it looks big and important, to accent practically every word if that will make sure that nothing is missed, to shift the stress in compounds if they are not recognized as such (...), and to build almost every sentence toward a climax at the end. The result is an excess of fist-pounding, doubtless useful to keep drowsy motorists awake on the road but otherwise rather wearisome to listen to for long. » (Bolinger, 1989 : 393-394).

Le type d'encodage (lecture d'un texte sur un prompteur) serait à la base de cette surcharge accentuelle. A l'usage, et malgré l'évolution des techniques et les méthodes personnelles des annonceurs radio, ce type de prosodie se serait répandu à l'ensemble du discours journalistique, pour devenir un archétype d'une prosodie professionnelle (Bolinger, 1989 ; Fónagy & Fónagy, 1976).

« Maybe all this should be excused simply as the ingredients of a professional style. But for the constant listener it would be restful if occasionally the newscasters and their associates would just COOL IT [en majuscules dans le texte]. A daily exercise in taming the wild accents and toning down the ends of sentences might help. » (Bolinger, 1989: 397).

Le rôle majeur de la prosodie dans la caractérisation socioprofessionnelle est évident. Les auditeurs sont d'ailleurs à même de reconnaître, en parole filtrée, le style journalistique parmi d'autres styles de parole (Fónagy & Fónagy, 1976). Ces auteurs attribuent cela au caractère essentiellement *barytonique* de son système accentuel, opposé au système *oxytonique* traditionnel du français.

En comparant trois types de discours (conversation, lecture d'un conte, et nouvelles journalistiques), Fónagy & Fónagy (1976) constatent que, contrairement à l'impression

auditive, les trois styles n'ont pas de fréquence d'accents sensiblement différente. Par contre, la distribution des accents distingue les trois styles, le style journalistique permettant la réalisation d'un grand nombre de clachs accentuels. Si les cas d'arcs accentuels sont présents dans les trois styles, le style journalistique et la conversation se distinguent de la lecture de conte en ce sens que l'accentuation a une tendance 'centrifuge', i.e. qui tend à favoriser la formation de groupements cohésifs (arcs accentuels plus fréquents), alors que la lecture a une tendance accentuelle 'centripète', i.e. une certaine tendance à la dissimilation des événements. Également, la lecture se distingue des deux autres types de discours par une plus grande mélodicité de la voix, et une plus grande régularité d'apparition des accents. La lecture présente une distribution des schèmes rythmiques moins dispersée que les autres styles, car ce style exploite la redondance (syntaxico-sémantique et rythmique). Enfin, la conversation présente une dispersion inter-individuelle plus importante que le style journalistique et la lecture de conte.

Une autre caractéristique du style journalistique serait la présence d'un schéma accentuel très particulier vers la fin des énoncés, où la syllabe finale perd l'accent principal au profit de la première syllabe du syntagme (déplacement d'accent).

« Ce qui caractérise avant tout le rythme des actualités, c'est un changement de pattern rythmique qui intervient au cours de la phrase, ou plus précisément, vers la fin de la phrase ou plus rarement vers la fin d'une proposition. C'est à ce moment là que les accents se rapprochent tout à coup, que l'accent se déplace à l'intérieur des mots vers la première syllabe et qu'il frappe avec le plus de violence la première syllabe du syntagme, même si cette première syllabe est atone par sa nature, ou plus exactement, par l'absence de poids sémantique qui caractérise la classe de mot constituant la syllabe » (Fónagy & Fónagy, 1976 : 203).

Nous avons nous même observé ce phénomène, particulièrement évident à la fin des paragraphes (en fin de topique majeur). Il semblerait donc que le schème barytonique soit particulièrement marqué aux alentours des frontières prosodiques majeures, et ce proportionnellement à la force de la frontière (énoncé < topique). Cette caractéristique ne se retrouve pas en lecture ou en parole spontanée.

Cet 'effet' stylistique, proprement professionnel, est accompagné la plupart du temps d'un *ritardando* et d'un *staccato* final de clause.

« Tout contribue à accroître le suspense : la montée tonale [sur la syllabe précédant la clause], le ralentissement du débit, la voix étouffée puis suspendue. » (Fónagy & Fónagy, 1976 : 208).

Les actualités opposent donc un style dynamique, agité, nerveux, au style ondoyant et caressant des contes de fées. Sans aller jusqu'à proposer une analyse psychanalytique de l'inclination de ce type de discours pour une accentuation 'heurtée' (inclination que les auteurs comparent à la pulsion prégénitale sadique-anale (sic !)), nous partageons leur analyse en termes d'adéquation du matériau sonore au contenu sémantique, souvent teinté de 'sensationnalisme'.

« N'y a-t-il pas une certaine analogie entre la parole heurtée, saccadée par des accents précipités et violents, les formes mélodiques rigides, angulaires,

d'une part, et les tendances inhérentes au genre qui cherche à frapper d'étonnement, et qui présente plus d'affinités avec des changements dramatiques, violents, des accidents 'spectaculaires', qu'avec les aspects idylliques de la vie quotidienne. » (Fónagy & Fónagy, 1976 : 215).

Pour conclure, ce type de discours a l'avantage de présenter des caractéristiques accentuelles et rythmiques très particulières, fondées sur l'utilisation optimale de l'immense liberté rythmique que permet la langue française. De par le caractère hybride du type d'encodage ('récitation' d'un texte écrit en respectant la syntaxe et le rythme de l'oral), nous émettons l'hypothèse que ce style de parole se situe sur le plan prosodique entre les styles de parole lue et spontanée. Nos analyses expérimentales s'efforcent de mettre en évidence les caractéristiques rythmiques et accentuelles de trois styles de parole (*Lecture*, *Interview* et *Journal* radiophonique). Nous verrons que, sur tous les points de nos observations, notre hypothèse concernant le statut intermédiaire du style journalistique sur le plan prosodique est vérifiée.

3.2. Caractéristiques phonétiques et prosodiques des différents styles de parole

La nécessité d'intégrer dans les modèles de synthèses les caractéristiques phonétiques et prosodiques des styles de parole est à l'origine du développement des recherches sur ce sujet (cf. notamment Sagisaka et al., 1997). Si l'on peut constater qu'il existe plus de similitudes que de différences entre les caractéristiques phonétiques des logatomes et de la conversation (Klatt, 1976), il n'empêche que la parole spontanée présente une variabilité dans ses formes prosodiques et phonétiques nettement plus importante que la parole contrôlée. Pour le cas précis des durées segmentales, Campbell (sous presse) constate que les données de laboratoire ne sont pas suffisamment représentatives des distributions réelles des phonèmes mesurés en situation de parole réelle. L'auteur remarque qu'il existe plus de variabilité interstylistique pour un même locuteur que pour différents locuteurs d'un même style de parole.

Pour pallier cette variabilité, Campbell propose de modéliser les durées à différents niveaux de constituance. Il suppose en effet que la variabilité interstylistique et interlocuteurs est moindre aux niveaux de la syllabe ou du pied accentuel. En ce sens, Campbell se situe dans la lignée des travaux de Kiel (Kohler, 1986 notamment) où le modèle d'organisation temporelle est élaboré sur plusieurs niveaux de constituance métrique (segments, syllabes, pieds). L'aller-retour entre ces niveaux, ainsi que l'application descendante des règles temporelles permettent, selon ces auteurs, de minimiser la part de variabilité inter- et intra-locuteurs.

Nous verrons cependant que la variabilité temporelle notamment est présente à tous les niveaux, y compris aux niveaux de constituance supérieurs. Par contre, le degré de variabilité est moindre pour les styles de parole non spontanée (cf. Chapitre 7 - L'organisation temporelle de la parole, p. 259, notamment).

Dans la perspective d'une modélisation des caractéristiques de la parole spontanée, quels seraient les éléments pertinents à retenir ? Campbell (1995a) montre que, outre les paramètres globaux tels le débit ou les contours intonatifs, les caractéristiques segmentales doivent être prises en compte. Campbell estime que les systèmes automatiques, pour qu'ils soient réellement performants, doivent prendre en compte l'intention du locuteur, sa visée

communicative, qui détermine en grande partie la production de la parole. En ce sens, les actes de langage doivent être étiquetés, et c'est seulement en fonction de ces caractéristiques de niveaux supérieurs que les contraintes de bas niveaux (acoustiques) doivent être analysées, reproduites ou reconnues. Cette méthodologie a également l'avantage d'être économique, et de concentrer l'analyse sur les variations acoustiques pertinentes.

Dans le même esprit, Hirschberg (sous presse) a amplement cherché à caractériser les types de discours. Selon elle, la qualité de la voix, la réduction ou l'hyperarticulation segmentale, la présence ou l'absence d'aisance d'élocution, le rythme, les proéminences accentuelles, les contours intonatifs et l'étendue tonale, varient selon le type de discours et la situation de communication. Hirschberg déduit de ses résultats comparatifs que seuls les paramètres prosodiques globaux (étendue de la f_0 et organisation temporelle) seraient révélateurs des styles de parole, alors que les paramètres locaux (contours intonatifs, variations locales de f_0 etc.) refléteraient l'information structurale et sémantico-pragmatique indépendamment du style de parole.

Nous verrons néanmoins que certaines caractéristiques locales (allongement différentiel des composantes syllabiques sous l'effet de l'accent) peuvent également refléter la nature du style de parole (spontané vs. non spontané). Il reste cependant à vérifier si les variations locales sont pertinentes perceptivement pour la caractérisation des types de discours.

La caractérisation phonétique et prosodique des styles de parole se fait généralement dans une perspective contrastive. Les recherches s'effectuent sur la base d'une opposition binaire entre parole lue/parole spontanée, cette dernière se trouvant souvent dans un flou définitionnel. Nous nous proposons maintenant de présenter les principaux résultats émanant des études contrastives sur les styles de parole.

3.2.1. Principales caractéristiques phonétiques et prosodiques distinguant les styles de parole

Du point de vue de la production, il semblerait que l'organisation prosodique de la lecture obéisse à une règle de sélection à droite, i.e. qu'elle est déterminée par le dernier contour intonatif de la phrase (possibilité d'anticipation visuelle et donc conceptuelle), alors que l'organisation prosodique de la parole spontanée serait par contre déterminée par une règle de sélection à gauche, de par le caractère non programmé, non 'préfabriqué' de ce type de discours (Guaïtella, 1991). En parole spontanée, les réalisations prosodiques se structureraient donc plus en fonction de ce qui a été dit que de ce qui va être dit (Guaïtella, 1990). En ce sens, l'auteur se rapproche du modèle intonatif de Martin (1981a), selon lequel les 'mots intonatifs' (structures intonatives minimales) sont liés par une loi de dépendance qui détermine la formation de groupement.

« L'inversion de pente mélodique est le mécanisme essentiel qui marque, en français, la sélection à droite : tout mot ou syntagme intonatif qui se joint à un mot ou syntagme situé à sa droite pour former un groupe plus grand présente un contour de sens de variation mélodique opposé à celui du syntagme auquel il se joint. » (Martin, 1981a : 265).

Les contrastes de 'figures' contribueraient donc à la formation de groupements perceptifs. Cependant, Guaïtella propose que le principe de sélection à droite ne s'applique qu'à la parole pré-programmée, et propose un principe de sélection à gauche pour la parole

spontanée.

La parole spontanée serait également caractérisée par l'utilisation systématique de contours intonatifs de continuation (Hirschberg, 1995 ; sous presse ; Wichmann, 1993 ; Levin et al., 1982), alors que la lecture adapterait les contours intonatifs aux marques de ponctuation (contours montants et descendants). Cette différence d'organisation intonative serait la trace d'activités cognitives différentes (encodage différé ou simultané) et de la connaissance partagée entre les interlocuteurs (connaissance sur la situation de communication : conservation de la parole menacée ou non, par exemple) (Levin et al., 1982 ; Guaitella & Santi, 1992). Même en parole filtrée, les auditeurs sont capables de reconnaître le caractère spontané ou lu d'un discours, et ce essentiellement grâce à la distribution et la forme des pauses (hésitations vocales notamment). En parole spontanée, de fait, l'organisation rythmique apparaît plus variable et le tempo moins régulier qu'en lecture (Levin et al., 1982).

Une manifestation concrète de la distinction entre l'activité de lecture et l'activité de parole spontanée est la difficulté qu'éprouvent les auditeurs à ponctuer la parole spontanée, alors qu'ils sont beaucoup plus cohérents et précis dans cette tâche pour la lecture. Cette différence de performance pourrait être due au fait que les frontières prosodiques sont marquées par moins d'indices concomitants en parole spontanée qu'en lecture (Wichmann, 1993). Guaitella (1991) et Guaitella & Santi (1992) montrent notamment que la pause joue un rôle prépondérant dans la perception des frontières en parole lue, alors qu'il semblerait que l'amplitude du contour de f_0 soit plus pertinent pour la perception de la parole spontanée. La pause silencieuse est perçue en parole lue comme volontaire et donc pertinente, alors qu'elle est synonyme d'hésitation et non de structuration en parole spontanée. L'absence de l'indice pausal de structuration en parole spontanée est compensé par l'amplitude de variation de f_0 .

Il apparaît que l'auditeur

*« [...] s'adapterait donc aux contraintes pragmatiques liées au type d'oral, et interpréterait les phénomènes prosodiques en fonction de ces contraintes »
(Guaitella, 1991 : 134).*

La parole lue marque plus les contrastes vocaliques que la parole spontanée (Blaauw, 1992). Cette remarque va dans le sens des résultats de Campbell (1995a ; sous presse) selon lequel les phonèmes en lecture tendent à être clairement séparés en macro-classes phonémiques sur la base de leur durée, alors que la parole spontanée tend à réduire les différences moyennes entre les phonèmes. Les contrastes accentuels seraient plus importants en lecture qu'en parole spontanée (mouvements de f_0 et allongements plus importants sur les accentuées). Blaauw (1992) pense que l'interaction des paramètres (qualité vocalique, intelligibilité, variation d'intensité, taux d'articulation et placement des accents) est sans doute importante dans la distinction des types de discours, mais il n'y a pas de ligne de démarcation perceptive catégorielle dans la distinction entre les styles de parole. Aucun des paramètres classiques (f_0 , durée et intensité) n'est dominant dans la caractérisation d'un style de parole. Cependant, la f_0 joue un rôle facilitant pour la reconnaissance des types de parole, plus que la durée.

Au regard de ces études, on s'aperçoit que les principales distinctions entre la parole lue et la parole spontanée tiennent aux caractéristiques prosodiques globales : marquage des

frontières, contours intonatifs, qualité vocalique, débit etc. Sur le plan de l'organisation temporelle, les études sont beaucoup moins unanimes et présentent même parfois des résultats contradictoires, sur la tailles des groupes rythmiques notamment. Cela tient essentiellement, selon nous, aux méthodes d'investigation. La grande majorité des auteurs proposent en effet de comparer les différents styles de parole sur la base de textes identiques, en proposant une relecture d'un discours prononcé dans des circonstances généralement 'non contraintes'. Nous tâchons dans le paragraphe suivant de mettre en lumière les limites d'une telle démarche.

3.2.2. Critique des méthodes traditionnelles de comparaison des types de discours

La plupart des études procèdent à des comparaisons entre des situations de communications pseudo-spontanées (dialogues homme-machine ; Map tasks etc.) et la relecture par les mêmes locuteurs de ces dialogues ou monologues. L'avantage d'une telle méthode est le contrôle du contenu sémantique, syntaxique et segmental, et le contrôle de la variabilité inter-individuelle. A priori, la variabilité intra-individuelle observée ne devrait refléter que la différence de processus d'encodage.

L'étude de Bruce (1995) révèle ainsi que les principales différences entre les deux types de discours (spontané/lu) tiennent aux stratégies pausales, au morcellement en syntagme plus petits en parole spontanée, et à une plus grande dynamique de f0 locale sur les éléments en focus.

Lucci (1973, 1983) constate, dans les mêmes conditions méthodologiques, que la lecture se caractérise par un allongement des groupes rythmiques, un accroissement du débit, et une certaine stabilité de durée des pauses. En conférence, par contre, le temps de pause est fortement dispersé (variabilité accrue) et la proportion des pauses augmente. L'auteur déduit donc que le nombre et la durée des pauses sont inversement corrélés avec la taille des groupes rythmiques.

Nous verrons que nos propres résultats sur l'organisation temporelle de la parole indiquent exactement la tendance inverse, confirmant par ailleurs les résultats de Fónagy & Fónagy (1976) et Fónagy (1980) qui travaillent également sur de la parole en situation. Il apparaît notamment que les groupes rythmiques sont plus longs en parole spontanée qu'en parole lue. Les résultats de Bruce et Lucci sont sans doute à imputer au caractère très artificiel de l'expérience, qui implique la lecture d'un texte a priori conçu pour l'oral.

De telles méthodologies conduisent en fait à des résultats peu fiables, aux dires mêmes des auteurs. Ainsi, Hirschberg (1995) reconnaît que les types de discours spontanés qu'elle étudie (dialogue homme-machine, Map tasks) sont peu naturels. D'autre part, les présupposés sur ce qui constitue la spécificité de la parole spontanée sont nombreux : en l'occurrence, les faux départs et hésitations sont explicitement enlevés pour la tâche de relecture, ce qui constitue un biais majeur pour une éventuelle comparaison objective des différences entre les deux types de discours. L'étude de Blaauw (1992) comporte le même biais méthodologique.

Nous l'avons vu, les auteurs qui ont eu le mérite d'essayer de caractériser les styles de parole sur la base des composantes prosodiques et phonétiques, étudient la relecture par les mêmes locuteurs d'un discours plus ou moins 'spontané', le plus souvent produit dans des circonstances de parole très peu naturelle, où les locuteurs connaissent les visées

expérimentales de leurs productions. Or, comme le font remarquer Wenk & Wioland (1984), lorsqu'ils légitiment un travail de recherche sur la parole radiophonique, il y a un grand intérêt à étudier la parole en réelle situation de communication.

« (...) en parlant à la radio, les sujets ne modifient pas leur comportement verbal en fonction des buts supposés d'une étude à laquelle ils auraient prêté leur voix. Leur production reste, donc, relativement spontanée et soumise aux seules contraintes propres à ce type de communication. » (Wenk & Wioland, 1984 : 71).

A l'instar de Wenk & Wioland, et comme le fait remarquer Campbell (1995a), nous pensons en effet que les différences entre les types de discours se mesurent en termes de « fonction » du discours. Dans sa forme naturelle, non contrainte, la parole est interpersonnelle et orientée vers un but communicatif fonctionnel. Dès lors que le locuteur est en situation de discours contrainte (enregistrement en chambre sourde, à l'aide d'un micro, où la plupart du temps l'interlocuteur est absent et fictif), la parole a une visée de 'production' pure, sans tenir compte de l'auditeur (elle n'a plus sa visée naturelle de communication). Les diverses études, modèles et théories sont donc fondés sur ce type de parole non représentatif de la parole humaine, qui implique forcément une intention de communication. Rien d'étonnant alors que les formes prosodiques soient si différentes de ce que l'on trouve en réalité, dans la parole en situation.

La visée des études de Bruce, Hirschberg, Lucci etc. est louable en ce sens qu'ils essaient de sortir du carcan de la parole lue, qui a été pendant des années le seul matériau d'étude à partir duquel des modèles ont été créés et ont donné lieu à des applications (synthèse et reconnaissance de la parole). Cependant, ces auteurs tombent dans l'excès inverse. En réhabilitant la parole spontanée, ils dénaturent la parole lue. En effet, en proposant de considérer la relecture d'un discours spontané comme de la parole lue, ils s'écartent de l'essence même de la lecture. La lecture a forcément un lien avec le texte écrit. Le texte écrit constitue un système à part entière, avec ses propres règles d'élaboration (syntaxe, ponctuation etc.). Lorsqu'on transcrit un texte spontané, on transcrit un autre type de système linguistique : le système de l'oral. Et l'on sait que les systèmes de l'écrit et de l'oral ont chacun leur spécificité (Blanche-Benveniste & Jeanjean, 1987 ; Guaïtella & Santi, 1992 ; Guaïtella, 1991).

Ce genre de méthodologie comporte donc des biais considérables et s'inscrit en faux par rapport à une réalité linguistique indéniable. D'ailleurs, ces auteurs ne précisent que très rarement comment ils procèdent pour transcrire un texte spontané : comment choisissent-ils les signes de ponctuation, sachant que le système de ponctuation de l'écrit est inadéquat pour rendre compte de la prosodie de l'oral (Guaïtella, 1991 ; Guaïtella & Santi, 1992) ? ; Sur quoi se fondent-ils pour éliminer certaines caractéristiques de la parole spontanée (hésitations, faux départs etc.) et appliquent-ils des critères systématiques ?

Sur ces bases là, on peut se demander si la 'Lecture' telle qu'elle est définie dans ce genre d'études, correspond encore à un style de parole à part entière. Au risque de paraître sévère, nous pensons que ce type de recherches présente le double désavantage de travailler sur des corpus pseudo-spontanés et de la pseudo-lecture. On peut se demander dès lors dans quelle mesure les résultats qui en ressortent sont fiables et exploitables. Il est étonnant de voir que les auteurs, pourtant conscients des limitations de leurs études, n'aient pas cherché

à définir précisément la notion même de type de discours ou de style de parole, afin de promouvoir la comparaison de styles de parole ‘en situation’, où un maximum de biais seraient ainsi évités.

C’est dans cette dernière optique que nous situons notre recherche. Nous avons en effet opté, à l’instar de Campbell, Wenk & Wioland et Fónagy, pour une analyse prosodique de la parole en situation. Si le matériau linguistique est non contrôlé, l’intégrité des styles de discours est néanmoins préservée, et les observations qu’on peut faire sont par là même plus représentatives de la réalité linguistique et prosodique.

Le chapitre suivant est consacré à la présentation de notre corpus et de notre méthodologie. Nous tâchons de situer le cadre de notre recherche notamment par rapport aux théories linguistiques permettant de rendre compte précisément des différents types de discours. C’est sur ces bases que nous proposons ensuite de faire le lien avec une théorie de la prosodie.

4. RESUME

Dans la perspective où nous postulons l'existence d'un système linguistique noyau, il convient d'aborder les notions d'invariance et de variabilité, intimement liées. Nous montrons que ces notions sous-tendent très souvent les notions de 'normes' et 'd'écarts à la norme', qui portent tort à l'analyse de la parole en situation. Ainsi, nous pensons qu'aucun type de discours ne doit être considéré comme 'neutre' (typiquement, parole de laboratoire), car cela implique que les autres types de discours (parole continue en situation) ne sont définis qu'en relation à cette 'norme' linguistique. Au contraire, nous pensons que tout discours est marqué. Le marquage stylistique, quel qu'il soit, découle de la situation de communication. Il est cependant systématisable car il est le résultat d'une adaptation à un code linguistique. De fait, les réalisations rythmiques et accentuelles variables selon les types de discours ne constituent pas la preuve de systèmes organisateurs différents (Guaïtella, 1991), mais sont le reflet d'un phénomène phonétique de surface, impliquant l'existence d'un système commun dans la représentation mentale sous-jacente. Nous refusons par là même la dichotomisation entre parole lue/parole spontanée. Nous proposons plutôt de caractériser les types de discours sur une échelle de 'spontanéité', en prenant en considération les multiples facettes des situations de communication (Eskenazi, 1993).

Nous centrons notre analyse sur la parole en situation. Pour cela, nous avons choisi d'analyser trois types de discours produits en véritables situations de communication (*Lecture*, *Journal Radiophonique*, *Interview*). Nous nous démarquons par la même des méthodes traditionnelles de comparaisons inter-stylistiques (Lucci, 1983 ; Blaauw, 1992 ; Bruce, 1995 ; Beckman, 1997 ; Hirschberg, sous presse, etc.). Si ce genre de parole est moins contrôlé sur le plan du contenu segmental et linguistique, il est néanmoins plus représentatif de la réalité prosodique. Nous pensons en effet que les différences prosodiques entre les types de discours sont directement liées à la fonction, à la 'visée communicative' du discours (Wenk & Wioland, 1984 ; Campbell, 1995a).

L'analyse du style journalistique nous paraît intéressante pour deux raisons principales : d'une part, ce style de parole ancre dans un certain usage l'évolution accentuelle barytonique du français ; d'autre part, ce style présente une hétérogénéité linguistique, qui lui confère des caractéristiques accentuelles et rythmiques empruntant à la fois à la parole lue et à la parole spontanée. Dans la perspective de la mise en évidence d'un système noyau, la place intermédiaire qu'occupe le style journalistique sur une échelle de spontanéité est particulièrement intéressante pour juger de la réalité de structures rythmiques invariantes, et de la variabilité rythmique correspondant à un type de discours particulier.

- *CHAPITRE 4* -

Aspects Méthodologiques

1. OBJECTIFS DE L'ETUDE ET CADRE METHODOLOGIQUE

L'accentuation, le timbre, le débit, les pauses participent du rythme d'une langue (cf. Chapitre 2 - Les paramètres du rythme - , p. 51 et suivantes). Les formes et substances variées de ces phénomènes reflètent la complexité inhérente à la notion de rythme. Cette hétérogénéité se traduit également dans la difficulté à mettre en évidence des relations simples unissant les deux aspects de ce phénomène : la réalité acoustico-phonétique et la réalité subjective (perception).

Nous avons choisi de centrer notre étude sur l'observation acoustico-phonétique du rythme, et plus précisément sur l'organisation paradigmatique et temporelle de l'accentuation. Ce paramètre du rythme nous paraît en effet central, de par la relation étroite qu'il entretient avec la notion de constituance prosodique.

Notre objectif est donc la caractérisation du système accentuel du français en termes de ses corrélats acoustico-phonétiques. A cette fin, nous avons choisi d'appliquer à notre étude le modèle accentuel issu de Di Cristo & Hirst (1997), repris et développé dans Di Cristo (sous presse), qui proposent une typologie fonctionnelle, ainsi qu'une interprétation formelle du système accentuel français dans un cadre interprétatif métrique. Le choix de ce modèle est amplement discuté dans le paragraphe « Un modèle de l'accentuation du français », p. 90.

A l'instar de ces auteurs, nous émettons l'hypothèse que le rythme, en tant qu'ossature du système prosodique d'une langue, possède des structures invariantes. Ces structures invariantes constituent les points d'ancrage du rythme d'une langue donnée. Les variations observées suivant les types de discours seront à interpréter en termes « d'alloformes » de ces formes invariantes. Nous pensons, à la différence de certains auteurs (tels Guaïtella, 1996 par exemple), que les différents processus d'encodage (lecture, parole spontanée) n'affectent pas la structure profonde de l'organisation accentuelle. Néanmoins, nous sommes consciente qu'une certaine variabilité induite par la situation de communication va se manifester au niveau de la structuration rythmique des énoncés, aussi bien dans l'agencement des paramètres accentuels que dans l'organisation temporelle de la parole, et coexister avec les invariants rythmiques. Il nous semblait donc intéressant de rendre compte à la fois de cette variabilité et des invariants rythmiques du français. Pour cela, nous avons choisi d'étudier l'organisation rythmique de trois types de parole susceptibles d'être situés sur une échelle de spontanéité : *Interview*, *Journal Radiophonique*, *Lecture* (cf. la caractérisation de ces différents types de parole, § 3.1.2, p. 135).

Les théories métriques décrivant les systèmes accentuels des langues se fondent traditionnellement sur des phrases construites par l'expert à des fins descriptives (Selkirk, 1984 ; Halle & Vergnaud, 1987 ; Hayes, 1989 etc.). Plusieurs études portent ainsi sur la description de phrases lues (Delais, 1994 par exemple), voire de paragraphes ou de textes lus. La description de la parole 'spontanée' est restée essentiellement l'apanage des tenants de l'approche comportementaliste (Bolinger ; Fónagy ; Guaïtella, 1991 et suivantes etc.). Seule Couper-Kuhlen (1993) a envisagé, pour l'anglais, une analyse métrique de la conversation. Ces études varient cependant aussi bien par leurs visées théoriques que par les méthodologies et les matériaux linguistiques employés.

Il nous semblait donc intéressant de procéder à une étude de l'accentuation du français sur de la parole continue 'en situation', utilisant un même cadre descriptif et une même méthodologie. De fait, l'objectif central de cette étude est de pouvoir dissocier la variabilité stylistique de la structure invariante du système linguistique.

Un autre intérêt de notre étude réside dans la démarche même que nous employons. En choisissant un modèle proposant une typologie fonctionnelle de l'accentuation, nous nous situons dans une démarche hypothético-déductive, opposée à une démarche ascendante dont le but serait de déduire des règles phonologiques à partir des observations ; ce genre de démarche est coûteux, peu économique si l'on veut extraire des informations pertinentes sur de la parole en situation. L'intérêt d'une démarche descendante est la confrontation d'un modèle aux données expérimentales. L'aller-retour s'opérant entre les deux permet d'une part, une certaine efficacité de l'observation par le repérage au préalable d'unités pertinentes ; d'autre part, cette démarche permet un éventuel ajustement du modèle en fonction des limitations rencontrées.

2. PARAMETRES RETENUS POUR L'ANALYSE DE L'ACCENTUATION

Les paramètres acoustico-phonétiques de l'accentuation sont essentiellement la durée, la f_0 et l'intensité. Le rôle de la durée comme paramètre de l'accentuation en français n'est pas à démontrer. Le français a d'ailleurs souvent été décrit comme ayant une structure rythmique fondée essentiellement sur les contrastes de durée (alternance de syllabes courtes et longues).

Par contre, le caractère récessif de la f_0 dans les descriptions est sans doute à imputer au postulat selon lequel le français est une langue sans accent. Ainsi, certains auteurs invoquent le syncrétisme de l'accentuation avec l'intonation comme preuve de cette absence de structure accentuelle (Rossi, 1980). De fait, un grand nombre de descriptions en français concernent les contours de f_0 participant de l'intonation, et négligent le rôle de la f_0 comme paramètre de l'accentuation (paramètre phonétique d'une entité phonologique à part entière) (Delattre, 1966a, 1966b et 1966c ; Vaissière, 1974b ; 1975 ; Martin, 1981a ; Mertens, 1993, notamment). A ce titre également, la f_0 est souvent un corrélat de l'expressivité du message, de l'état émotionnel du locuteur, ou des fonctions modales de l'intonation (question, affirmation etc.). D'autre part, nous avons vu que l'influence de la tradition anglo-saxonne a longtemps pesé sur les descriptions des langues, et a participé à propager des croyances erronées sur le rythme et l'accentuation du français. L'accentuation en français n'est pas distinctive au niveau lexical. Or, dans de nombreuses langues à accentuation lexicale, le corrélat acoustique principal de l'accentuation est la f_0 . Par conséquent, nombre de descriptions ont négligé de prendre en compte ce paramètre pour le français et se sont focalisées sur le paramètre de durée.

Pour ce qui est de l'intensité enfin, son rôle dans la caractérisation de l'accentuation en français n'a pas été démontré. Si pour l'anglais, l'intensité joue un rôle important dans la réalisation de l'accent, les travaux semblent conclure que ce n'est pas un paramètre déterminant en français, tout au moins en ce qui concerne l'accentuation non emphatique. Delattre (1965) montre même que la voyelle d'une syllabe finale (accentuée) est en moyenne légèrement moins intense que la voyelle d'une syllabe non finale (inaccentuée). L'intensité peut jouer cependant un rôle dans la perception du poids syllabique, si l'on

prend en considération l'effet de la 'total amplitude' (intégration temporelle de l'amplitude du signal, ce que Rossi, 1967b, nomme la sonie).

Au vu des problématiques liées aux paramètres de l'accentuation en français, nous retenons dans le cadre de cette étude les paramètres de durée et de f_0 . Nous n'avons pas choisi l'intensité pour deux raisons essentielles :

- tout d'abord pour une raison technique : notre corpus est constitué pour les deux tiers d'enregistrements radiophoniques, pour lesquels nous n'avons aucune information sur les conditions et les techniques d'enregistrement (distance du locuteur par rapport au micro notamment, mais aussi problèmes liés à la technique de compression de dynamique, qui compresse le signal et rend par là même délicate l'étude de l'intensité relative).

- ensuite, nous avons vu que l'intensité ne semble pas jouer un rôle déterminant dans la caractérisation de l'accentuation en français. On peut penser que, dès lors qu'on se place dans une perspective de description pragmatique de l'accentuation, l'intensité entre en ligne de compte dans la distinction des accents emphatiques notamment. Cependant, ce paramètre peut être observé indirectement à travers l'observation des allongements de durée consonantique à l'initiale de mot ou de groupe. Plus précisément, l'accroissement de durée des éléments consonantiques initiaux sous l'accent emphatique est lié à une augmentation de l'intensité : les deux paramètres sont en effet corollaires de l'augmentation de l'effort articulatoire (Benguerel, 1973).

Notre étude va donc s'articuler autour de deux axes majeurs :

- l'organisation paradigmatique de l'accentuation à travers l'étude des durées différentielles des composantes syllabiques d'une part, et des aspects dynamiques des variations locales de f_0 (configurations tonales) d'autre part ;

- l'organisation temporelle de l'accentuation, à plusieurs niveaux d'actualisation du rythme : la syllabe, le Mot Prosodique et le groupe de souffle, en fonction du type de frontière prosodique.

Nous présentons ici les grandes lignes de notre méthodologie. Nous détaillerons davantage les différentes étapes de notre méthodologie lors des chapitres dédiés aux résultats de notre recherche (cf. Chapitres 5, 6 et 7).

2.1. Organisation paradigmatique de l'accentuation

Nous voulons dans un premier temps explorer les corrélats acoustico-phonétiques des différents types de catégories accentuelles définies dans le cadre du modèle de Di Cristo & Hirst (1997), et Di Cristo (sous presse). Nous procédons par conséquent à une description paradigmatique des catégories accentuelles. Nous postulons que les différences fonctionnelles et distributionnelles vont être matérialisées au niveau des paramètres phonétiques. La distinction des catégories accentuelles prédéfinies par les paramètres de durée et de f_0 validerait le modèle descriptif, et nous conforterait dans notre choix.

2.1.1. Durée : quelles unités d'observation ?

Il est délicat d'isoler des unités dans le flot continu de la chaîne segmentale (cf. § « Segmentation phonémique », p. 143). La parole se déroulant dans le temps, les événements s'enchaînant les uns après les autres, comment est-il possible de discrétiser la durée, et quelle(s) unité(s) de discrétisation choisir ? Notre démarche descendante nous permet d'avoir un cadre préalable d'analyse, facilitant la discrétisation des unités. Ainsi, nous postulons que l'unité de base, domaine de réalisation de règles phonologiques, est la syllabe (cf. § « La syllabe », p. 42). La syllabe est donc le lieu d'actualisation des phénomènes accentuels (Garde, 1968). Nous nous proposons d'observer les phénomènes d'allongement (ou de réduction) qui s'appliquent aux syllabes en fonction des catégories accentuelles. La syllabe ne constitue pas pour autant une unité phonologique irréductible. Ainsi, les constituants syllabiques (Attaque, Noyau, Coda) sont-ils décrits dans la phonologie non linéaire (Durand, 1990) comme des unités susceptibles de porter des informations prosodiques. Les règles d'ajustement de durée entre les composantes syllabiques sont ainsi la preuve d'une organisation à l'intérieur de la syllabe. De fait, il nous semblait intéressant à la suite de Campbell (1993, et suivantes), Fant et al. (1991a et b) et Beckman & Edwards (1992) notamment, de rendre compte également des phénomènes de durées différentielles dans le cadre de la syllabe. Nous verrons que les stratégies d'allongement différentiel varient suivant la catégorie accentuelle observée, et permettent de distinguer notamment les catégories accentuelles en fonction de leur distribution, et du degré de la frontière subséquente. Enfin, nous essaierons de dégager les caractéristiques accentuelles invariantes quel que soit le type de discours.

2.1.2. Hauteur : unités et paramètres d'observation

Nous proposons de poursuivre l'étude des corrélats phonétiques des catégories fonctionnelles prédéfinies, en termes de propriétés tonales. En effet, si le français est généralement défini sur une base principalement temporelle, des études récentes (Di Cristo & Hirst, 1997 ; Beckman, 1993) ont mis l'accent sur le rôle des indices tonals dans l'organisation rythmique du français, dans la mesure où certaines syllabes dans un mot peuvent être associées phonologiquement à un ton ou une séquence de tons. Terken (1993) insiste sur le rôle prépondérant de la hauteur dans la perception de la proéminence. Il est cependant partisan de l'étude combinée des facteurs prosodiques et notamment des facteurs liés à l'aspect dynamique des variations de f_0 . Dans le même ordre d'idée, t'Hart et al. (1990) fondent leur description en contours tonals (plutôt qu'en tons discrets) sur le caractère continu et dynamique de la f_0 .

A la suite de Terken (1993), et t'Hart et al. (1990), nous choisissons de prendre en compte les propriétés temporelles de la f_0 telles que les directions de pentes, l'empan syllabique et le timing des configurations, afin de rendre compte des spécificités de l'organisation tonale du français. Cette analyse est complémentaire aux études précédentes de l'organisation tonale du français en termes de tons discrets (Hirst et al., sous presse).

Pour cette étude, l'accent est mis sur les invariants tonals, i.e. sur les paramètres tonals permettant de caractériser les différentes catégories accentuelles, quels que soient les locuteurs ou les styles de parole.

2.2. Organisation temporelle de l'accentuation

Après une analyse détaillée des corrélats phonétiques des catégories accentuelles sur un axe paradigmatique, l'analyse de la relation qu'entretiennent les catégories les unes par rapport aux autres sur un axe syntagmatique nous paraît complémentaire.

Nous avons vu que le français, longtemps considéré comme une langue à isochronie syllabique stricte, possède une double identité accentuelle et rythmique qui déroutent les observateurs (cf. § « Introduction : le mythe de l'isochronie et de l'isosyllabité », p. 32 ; et § « L'accentuation initiale : évolution du système accentuel en français », p. 80). D'une part, de nombreuses études ont montré que les accents en français tendent à se produire à intervalles réguliers, indiquant par là que le rythme syllabique coexiste avec un rythme accentuel (Wenk & Wioland, 1982 ; Fant et al., 1991a et b, notamment). D'autre part, il existe une 'nouvelle' tendance (ou du moins, une tendance de plus en plus répandue) à marquer le début des mots ou des groupes de mots par un accent initial ayant une 'parenté' physique avec les accents de l'anglais par exemple, réputée langue à isochronie accentuelle. Cette tendance peut contribuer dans une certaine mesure à la perception globale d'une récurrence régulière subjective des accents.

Nous nous proposons de rendre compte de ces particularités accentuelles par l'analyse de l'organisation temporelle des Mots Prosodiques en français. Nous distinguons deux types de Mots Prosodiques :

- le Mot Prosodique simple que nous définissons comme une unité constituée d'un seul Pied Métrique borné à droite par un accent final de syntagme ou d'Unité Intonative ;
- le Mot Prosodique complexe, constitué de deux Pieds Métriques (PM), le premier PM étant borné à droite par un accent initial et le deuxième PM borné à droite par un accent final. Le modèle accentuel que nous utilisons (Di Cristo & Hirst, 1997 ; Di Cristo, sous presse) nous permet de mettre en évidence l'influence de la nature de l'accent sur l'organisation temporelle des Mots Prosodiques.

Nous testons également l'influence des différents processus d'encodage (lecture ou parole spontanée) sur l'organisation rythmique des Mots Prosodiques (Voir Guaïtella, 1991, pour une discussion détaillée). Nous observons ainsi des paramètres tels que la densité accentuelle (pourcentage de syllabes accentuées par rapport au nombre total de syllabe) et la variabilité des différents niveaux d'actualisation du rythme (syllabes, Mots Prosodiques simples et complexes, et groupes de souffle). Enfin, nous procédons également à une analyse de la distribution et de la nature des pauses, en fonction des groupes rythmiques.

3. PROTOCOLE EXPERIMENTAL

3.1. Description du corpus

3.1.1. Critères linguistiques de sélection du corpus

Comme nous l'avons déjà exposé, notre objectif central est l'étude des invariants rythmiques transversaux à différents types de discours. Dans la part inévitable de variabilité

qui ressortira de notre étude, nous désirons dégager la part de variabilité imputable seulement aux différents processus d'encodage. A cette fin, il est impératif de contrôler au maximum les sources éventuelles de variabilité qui pourraient parasiter, gêner l'interprétation de nos résultats. La tâche est énorme pour l'analyse de la parole, et nombres de zones d'ombre demeurent (Rossi, 1985a ; Lindblom, 1990).

Il est bien évident que l'analyse de corpus de parole continue, alliant un aspect de gradation sur une échelle de spontanéité, induit plus qu'elle n'inhibe une grande part de variabilité (cf. § « Invariance et variabilité », p. 107). Ce type de recherche est largement inspiré des travaux des sociolinguistes et des spécialistes de pragmatique. Ces derniers ont été en effet les précurseurs de l'analyse de la langue parlée dans différentes situations de communication, afin de rendre compte des dimensions de variabilité des faits de langue. Sous leur impulsion, les phonéticiens ont pris conscience de l'importance de l'étude de la langue 'en usage'. Parallèlement à la quête traditionnelle des invariants phonétiques et prosodiques, qui passe souvent par l'analyse de la parole contrôlée dite de 'laboratoire', est venue s'ajouter la nécessité pour les phonéticiens de décrire la langue dans toute sa variabilité. Dès lors, la diversité des situations de communication et les différentes formes d'interactions verbales sont prises en compte comme autant de sources de variabilité, participant de l'intégrité d'une langue.

Rares sont pourtant les études qui rendent compte de l'étendue de la variabilité prosodique à travers les différents types de discours en français (Guaitella, 1991 et suivantes ; Vihanta, 1991 et 1993 ; Fónagy & Fónagy, 1976 ; Lucci, 1973, 1983 ; Rossi, 1987).

3.1.1.1. Notions de 'styles de parole' et de 'types de discours'

Une première étape de contrôle de l'objet d'étude est le choix du corpus. Afin de mieux cerner notre objet d'étude, revenons sur les définitions des notions de 'styles' ou de 'types' de discours. Eskenazi (1993) dresse un bilan des études dont l'objet est la caractérisation des 'styles' de parole. Selon elle, les auteurs s'accordent généralement sur l'importance de la 'situation' du discours, i.e. le contexte dans lequel se trouvent les interlocuteurs. Ils n'atteignent cependant pas de consensus sur la terminologie attachée à la notion de style ; pour les uns, il s'agit du degré de familiarité du discours, pour les autres, du contexte social ; pour certains enfin, le style recouvre des notions aussi diverses que le mode de production de parole (spontanée/non spontanée), le type de discours produit (narration, conversation, lecture etc.) et le contexte dans lequel la parole a été produite et enregistrée (radio, télévision, chambre sourde, cadre familial etc.).

Il apparaît donc nécessaire de hiérarchiser les critères définitionnels afin de ne pas mélanger, voire oblitérer certaines notions essentielles à la caractérisation d'un type de discours. Cette étape nous paraît d'autant plus nécessaire qu'un de nos buts est de distinguer la variabilité prosodique induites par les différents types de discours. Un contrôle précis des variables discursives s'avère donc impératif.

Eskenazi énumère les facteurs de variabilité du style et propose par là même des éléments de définition. Le style est fonction de la perception qu'a le locuteur du statut de son auditeur et du type de situation dans lequel il se trouve. Il peut varier au cours du discours, en fonction de plusieurs facteurs externes et internes à la situation de

communication. Le style dans lequel le locuteur produit son discours est le mélange d'un effort conscient et inconscient de sa part, qui n'est d'ailleurs pas toujours perçu par l'auditeur dans le sens originel de l'intention du locuteur. La perception et la production d'un style particulier de parole sont en effet des paramètres très subjectifs, dépendant de stratégies individuelles liées aux contextes social, économique, psychologique, culturel etc. des interactants.

Ces réflexions nous amènent à percevoir les multiples dimensions de la notion de style de parole. La plupart des auteurs ne se positionnent d'ailleurs que par rapport à l'une ou l'autre de ces dimensions. Eskenazi (1993) tente d'embrasser toutes les acceptions de la notion de style et propose un modèle définitionnel sur un espace tridimensionnel permettant de situer les divers styles en fonction de leurs différentes manifestations. Elle propose trois axes graduels :

- l'axe d'*Intelligibilité* : selon la situation de communication, le locuteur doit fournir un effort de clarté plus ou moins important.
- l'axe du *degré de familiarité* : le style varie en fonction du degré d'intimité existant entre les interlocuteurs.
- l'axe de la *strate sociale* : selon le degré de culture sur lequel vont s'accorder les interlocuteurs pour s'entretenir.

A la croisée de ces différents axes, Eskenazi exprime les corrélats prosodiques et phonétiques des styles ainsi définis. Un locuteur dans une situation non familière combinée à une contrainte sociale élevée produira plus de pauses autour des mots pleins et liera son discours par l'insertion de schwas ; les segments présenteront des coefficients de variation réduits, et les différences articulatoires et acoustiques des macro-classes phonémiques se trouveront accentuées.

On pourrait penser qu'il existe une certaine corrélation entre ces différentes dimensions du style, notamment par exemple entre le degré de familiarité et l'intelligibilité. Eskenazi donne l'exemple d'un même degré de familiarité entre les interactants, mais où la tâche du locuteur va varier : la lecture d'une histoire à un enfant et la conversation de tous les jours relèvent a priori du même degré de familiarité, mais ne nécessitent pas le même effort d'intelligibilité. Dans le premier cas, le locuteur va avoir tendance à allonger les durées et à utiliser une étendue de f_0 large, alors que dans le deuxième cas, les segments seront plus coarticulés et les cibles formantiques moins précises.

Le modèle d'Eskenazi propose donc une distinction nette entre 'style de parole' et 'type de discours'. Ce dernier recouvre la notion de 'tâche (discursive)' que le locuteur va devoir accomplir : interview, lecture, conversation, informations radiophoniques ou télévisuelles, conférence, etc. Ces tâches nécessitent l'utilisation d'un ou plusieurs styles de parole particuliers, l'interdépendance entre les deux phénomènes étant sans doute à l'origine des nombreuses confusions trouvées dans la littérature.

Nous avons maintenant assez d'éléments pour situer notre corpus en fonction du style de parole, mais une description linguistique en termes d'unités d'analyse discursive nous semble incontournable et complémentaire.

3.1.1.2. A la recherche d'une unité d'analyse linguistique

Dans notre tâche de description prosodique transversale à plusieurs types de discours, nous ressentons le besoin de définir notre corpus sur la base d'une même unité linguistique. Afin d'éviter toute circularité en effet (autonomie des niveaux d'analyse), nous préférons avoir recours à des considérations pragmatiques plutôt que phonétiques. Cette étape de description préalable nous paraît essentielle pour l'analyse de parole spontanée et connectée.

Dans cette perspective, nous nous inspirons de certains concepts issus de la Linguistique Interactionnelle. Les linguistes issus de ce courant théorique considèrent que tout message est adressé, tout discours constitue une forme d'interaction, voire que « *tout comportement humain, quel qu'il soit, procède de l'interaction.* » (Vion, 1992 : 18). Ce concept de l'interaction comme centre de toute communication verbale (ou non verbale) nous séduit pour deux raisons principales :

- il est particulièrement bien adapté à l'analyse de la parole spontanée en situation ;
- il nous permet de disposer d'une unité d'analyse homogène, adaptée à des types de discours hétérogènes sur l'axe spontané/non spontané.

3.1.1.2.1. *Cadre typologique de l'analyse des interactions*

Nous choisissons de nous inspirer du cadre typologique défini par Vion (1992), qui propose une caractérisation des différents types d'interaction sur la base de critères définitionnels. Ce modèle discursif présente un intérêt triple : d'une part, il est l'un des plus récents en la matière en langue française ; il présente d'autre part un souci d'intégration des concepts développés dans les théories et les modèles antérieurs proposés dans ce domaine ; enfin, les critères définitionnels nous paraissent rendre compte de manière optimale de la variété des interactions possibles.

Vion définit tout d'abord les composantes générales constitutives de l'interaction :

- la *situation*, résultat des rapports sociaux antérieurs des interactants, ou produit de l'activité des sujets ;
- les *participants*, caractérisés par leur statut social, leur rôle et leur appartenance à un groupe ;
- le *cadre interactif*, correspondant à la nature du rapport social établi au début de l'interaction ; ce cadre peut évoluer au cours de l'interaction, modifiant par là même la nature de cette dernière.

Son modèle propose une structure hiérarchique de l'interaction, composée de 6 unités formelles :

L'unité maximale est *l'interaction* définie comme « l'ensemble de ce qui se produit entre deux ou plusieurs sujets, du début à la fin de leur rencontre ». Vient ensuite le *module* qui constitue une unité homogène sur le plan discursif, plusieurs modules pouvant constituer une même interaction. Dans ce modèle, le *module* est une unité dialogique, mais il peut

également constituer une unité monologale. Au dessous du module dans la hiérarchie se trouve la séquence, unité fonctionnelle ou thématique. Puis l'échange, défini comme l'unité minimale de dialogue. Aux niveaux inférieurs de la hiérarchie enfin, se situent l'intervention définie comme la plus petite unité monologale, et l'acte de langage comme l'unité fonctionnelle minimale (la réponse à une question constitue un acte de langage, l'ensemble constituant deux actes de langage successifs).

L'unité linguistique que nous choisissons de retenir pour la catégorisation de notre corpus est le Module. L'objet de notre recherche en effet n'est pas la caractérisation macro-prosodique des types de discours, dans leur complexité interactionnelle et pragmatique (stratégies des tours de parole dans un cadre dialogique par exemple). Dans la perspective d'une caractérisation rythmique d'énoncés produits dans diverses situations de communication, il nous paraissait essentiel de définir des extraits de parole présentant une certaine cohésion discursive, tant du point de vue de la thématique que des énonciateurs. A cet effet, le Module nous paraît être l'unité la plus pertinente, car elle est caractérisée par une continuité temporelle et une homogénéité discursive. Dans une certaine mesure, la définition du Module sur le plan discursif peut être renforcée par une caractérisation prosodique. En effet, ce type d'unité discursive est caractérisée sur le plan prosodique par une unité cohésive, d'ordre supérieur dans la hiérarchie prosodique : le paraton, tel qu'il a été défini par Brown & Yule (1983) et Nicolas (1995).

3.1.1.2.2. Traits définitionnels du Module

La caractérisation du Module se fait sur la base de deux critères principaux, correspondant aux conditions de production du discours et à la structure interne de l'interaction.

– Les conditions de production du discours font référence aux distinctions *spontané/non spontané*, et *parole publique/parole privée*. Nous considérons comme *non spontané* tout discours prenant appui sur un texte écrit, qu'il soit lu, récité ou interprété. Est considéré comme spontané un discours élaboré par le locuteur au moment de sa production (cf. § « Comment définir la parole spontanée ? », p. 111, pour une discussion détaillée). Pour ce qui est de la deuxième distinction, nous appelons *parole publique* tout discours destiné à être diffusé, pouvant dépasser le cadre de l'interaction définie 'ici et maintenant'. La *parole privée* fait référence à l'utilisation du discours à des fins personnelles, ne dépassant pas un cadre privé.

– La structure de l'interaction fait référence à un critère interne de caractérisation de l'interaction. On distingue ainsi deux types d'interaction : les interactions à *structure d'échange*, dans lesquelles les participants ont la possibilité de devenir énonciateurs, et les interactions *sans structure d'échange* pour lesquelles cette possibilité n'existe pas.

Une fois l'interaction définie sur la base de ces critères généraux, le modèle de Vion (1992) propose d'affiner la caractérisation des différents types de discours par l'utilisation de 4 critères binaires :

– *symétrie vs. complémentarité* : l'interaction est *symétrique* lorsque les participants positionnent leur discours sur un plan d'égalité (la conversation par exemple). Elle est *complémentaire* en revanche, lorsque le statut respectif des participants n'est pas le même (ex : interaction entre le médecin et son patient).

– *coopération vs. compétition* : il y a *coopération* entre les interactants lorsqu'ils recherchent un consensus (conversation, entretien amical), alors que la *compétition* se caractérise par sa nature conflictuelle (dispute, débat).

– *finalité interne vs. externe* : l'interaction a une finalité *interne* lorsqu'elle est centrée uniquement sur la qualité du rapport liant les interactants, qu'elle contribue à entretenir (conversation). La finalité est *externe* si l'interaction comporte un enjeu symbolique (débat) ou s'il y a recherche de connaissances et d'information (entretien, interview).

– *caractère formel vs. informel de l'interaction* : ce critère est fonction des conventions plus ou moins contraignantes entourant l'interaction (lieu, statuts et relations des participants etc.)

A la lumière de ces différentes caractérisations, il apparaît que ces 4 critères ont été définis à l'origine pour un type particulier d'interaction (la conversation), puis adaptés à la définition des autres types de discours. Ceci constitue une certaine limitation, mais permet également une souplesse d'analyse. Ainsi, l'utilisation de l'ensemble des 4 critères n'est pas forcément nécessaire à la caractérisation d'un type d'interaction. Il est également probable de rencontrer deux types différents d'interaction définies par les mêmes critères ; dans ce cas, la distinction doit pouvoir s'opérer sur la base des critères plus généraux de l'interaction, comme les conditions de production ou la structure des interactions. En tout cas, ce modèle présente l'avantage de réduire les risques de variabilité incontrôlée en permettant d'objectiver notre classification linguistique. La Figure 8 donne une représentation des critères généraux et spécifiques permettant la caractérisation des types d'interaction. Ces critères avaient été adaptés du modèle de Vion (1992) dans le cadre d'un projet de recherche (groupe de recherche GARP, projet PACOMUST) (Astésano et al., 1995a).

<i>Conditions de production</i>		
<input type="checkbox"/> spontané	non spontané <input type="checkbox"/>	
<input type="checkbox"/> parole publique	parole privée <input type="checkbox"/>	
<i>Structure de l'interaction</i>		<i>Types d'interaction</i>
<input type="checkbox"/> à structure d'échange / sans structure d'échange <input type="checkbox"/>		Théâtre/cinéma <input type="checkbox"/>
<div style="display: flex; justify-content: center; align-items: center; gap: 20px;"> <div style="text-align: center;"> <input type="checkbox"/> complémentaire / <input type="checkbox"/> coopératif / <input type="checkbox"/> finalité interne / <input type="checkbox"/> formel </div> <div style="text-align: center;"> <input type="checkbox"/> symétrique <input type="checkbox"/> <input type="checkbox"/> compétitif <input type="checkbox"/> <input type="checkbox"/> finalité externe <input type="checkbox"/> <input type="checkbox"/> informel <input type="checkbox"/> </div> </div>		Lecture <input type="checkbox"/>
		Récitation <input type="checkbox"/>
		Lecture interprétée <input type="checkbox"/>
		Conversation <input type="checkbox"/>
		Transaction <input type="checkbox"/>
		Consultation <input type="checkbox"/>
		Enquête <input type="checkbox"/>
		Entretien <input type="checkbox"/>
		Dispute <input type="checkbox"/>
		Débat <input type="checkbox"/>
		Discussion <input type="checkbox"/>
		Autre <input type="checkbox"/>

Figure 8 : Typologie de l'interaction adaptée du modèle de Vion (1992) pour les besoins du projet PACOMUST. Ce tableau est issu de Astésano et al. (1995), p.37.

A partir des critères définitionnels des interactions, nous sommes en mesure de

caractériser les extraits de parole qui constituent notre corpus, et de justifier notre choix.

3.1.2. *Choix des locuteurs et des types de discours*

Notre corpus se compose d'environ 10 minutes de parole connectée, collecté à partir de la base de données constituée pour les besoins du projet PACOMUST. Nous avons sélectionné 3 types de discours qui nous semblaient représentatifs d'une gradation sur une échelle de spontanéité : *Lecture*, *Lecture interprétée*, *Interview*. Le caractère 'spontané/non spontané' d'un type de discours est la base définitionnelle d'une grande majorité de descriptions prosodiques. Cette notion ne recouvre cependant pas toujours les mêmes dimensions suivant les auteurs (cf. « Comment définir la parole spontanée ? », p. 111) et constitue une caractérisation assez sommaire de la réalité discursive. Nous pensons avoir montré ci dessus (cf. § 3.1.1, p. 129 et suivantes) la nécessité de définir plus précisément les notions de 'styles' et 'types' de parole et d'opérer une sélection du matériau linguistique sur la base d'un modèle de l'interaction. C'est sur ces bases que nous présentons maintenant notre corpus. Précisons que chaque type de parole est représenté par deux locuteurs de sexe opposé. Cette distinction n'est pas centrale pour notre étude, bien que certains auteurs aient cherché (et parfois trouvé) des différences significatives liées au sexe du locuteur (Bolinger, 1989; Léon, 1993). Nous verrons néanmoins que certains de nos résultats indiquent des tendances rythmiques liées à l'appartenance sexuelle des locuteurs (cf. Chapitre 7 - l'organisation temporelle de la parole -, p. 259).

Il est bien entendu que le choix de deux locuteurs par style de parole constitue une limite de notre étude ; nos résultats sont d'ailleurs interprétés avec précaution eu égard à la taille de notre corpus. Cependant, l'analyse de 10 minutes de corpus représente déjà une masse de travail considérable (analyses instrumentales, linguistiques, statistiques etc.). Le besoin d'une automatisation des relevés et d'une vaste base de données sur le français se fait cruellement sentir.

3.1.2.1. Présentation du corpus

- *Lecture* : il s'agit de la lecture d'un paragraphe d'un texte extrait de « "*L'univers inconnu*" dans *le monde de Walt Disney* », déjà utilisé par Nicolas (1995) pour une description de l'organisation intonative du texte lu (Référence : « L'univers inconnu » dans *le monde de Walt Disney*, extrait du texte "Notre ami l'atome" par Heing Haber, p. 6-93 : "Un univers d'atomes", p. 12.). Ce paragraphe constitue une seule unité discursive. Les sujets, natifs de la langue française, avaient pour consigne de lire le texte avec une intonation neutre, sans effet littéraire, correspondant à un style de lecture courant destiné à un auditoire potentiel. Après avoir pris connaissance du texte, afin de permettre le décodage ainsi que la compréhension du texte pour parvenir à un stade de locuteur-lecteur-énonciateur le meilleur possible, les locuteurs effectuent deux lectures dont une seule sera sélectionnée répondant le mieux aux critères de qualité de la lecture (sans hésitation ni reprise, et la plus naturelle possible). Nous avons choisi un locuteur et une locutrice, dorénavant désignés par *Lecture M* et *Lecture F* respectivement. Chaque locuteur produit ainsi environ 1 minute de parole connectée lue (47 sec. pour *Lecture M* et 54 sec. pour *Lecture F*).

- *Journal radiophonique* : le style journalistique a été longuement décrit par Bolinger (1989), Vihanta (1991), Fónagy & Fónagy (1976) etc. (cf. § « Style

journalistique : hétérogénéité linguistique », p. 113). Il nous semblait intéressant d'étudier un type de discours qui, nous semble-t-il, stigmatise et ancre dans le langage courant la tendance du français à marquer le début des mots ou groupes de mots par des accents initiaux rythmiques et emphatiques, plus systématiquement qu'un autre style de parole. Les professionnels de la communication qualifient ce type de discours de '*lecture interprétée*', indiquant par là l'oralisation stylisée d'un texte écrit. Cette pratique s'apparente selon eux à l'interprétation d'un texte de théâtre ou de cinéma, où l'acteur « s'approprie » le texte de l'auteur en s'efforçant de paraître « improviser » le message au moment de sa production. Dans cet ordre d'idée, le travail du journaliste est aussi de lire un texte au préalable rédigé dans un style 'oralisé', c'est à dire dont la construction grammaticale et syntaxique est plus proche de celle de l'oral (moins de phrases subordonnées, de conjonction de subordination etc., phrases en règle générale plus courtes qu'à l'écrit) (Brown & Yule, 1983). Ce faisant, les journalistes prennent en compte les différences de traitement cognitif liées aux divers processus de décodage de la langue ; ainsi, la mémoire visuelle impliquée dans l'activité de lecture se distingue de la mémoire auditive par le type de support textuel lui-même (Guaïtella, 1991). Dans l'activité de lecture, le sujet a la possibilité de revenir en arrière dans le texte pour rafraîchir sa mémoire à court terme, alors que cette démarche n'est pas possible si le support est un texte oral (non enregistré). Au niveau de l'encodage, le locuteur désireux de transmettre de manière optimale son message tendra à produire des groupements syntaxiques et prosodiques plus courts et plus simples afin de tenir compte de la taille de la mémoire à court terme (Frasse, 1956). Le journaliste professionnel a donc ce souci d'optimiser son message, en adaptant son discours aux règles de construction de l'oral, afin de faciliter le décodage du texte par l'auditeur.

Il semblerait toutefois que les commentaires journalistiques radiophoniques, ou télévisuels en voix off, aient une prosodie comparable (densité accentuelle plus importante, marquage plus systématique des prééminences initiales démarcatives à gauche de mot lexical ou de groupe), alors que le discours télévisuel où le présentateur à l'écran s'adresse directement aux auditeurs est moins marqué prosodiquement ; la technique d'encodage est pourtant la même (lecture interprétée). Cela est sans doute dû au feedback visuel (parole audiovisuelle), facilitant le décodage de la parole. Cette remarque relativise un peu l'argument selon lequel la prosodie de la parole journalistique est à imputer uniquement au caractère écrit du texte ; la réalité est sans doute plus complexe, englobant plus généralement le contexte discursif (présence ou non d'un interlocuteur, même si celui-ci est représenté par une caméra, ou un autre journaliste ; visée du discours etc. ; ainsi que contact visuel ou non des interlocuteurs etc.).

Nous avons choisi deux extraits de ce type de parole, issus de la station de radio 'France Info'. Les deux extraits ont une structure textuelle légèrement différente, l'un étant une Revue de Presse par un journaliste professionnel (*Journal M*) faisant référence à l'actualité rapportée par la presse écrite, alors que le deuxième est un compte-rendu de vulgarisation d'une découverte scientifique par une journaliste professionnelle (*Journal F*). Les deux locuteurs produisent environ 2 minutes de parole chacun (2 mn 10 pour *Journal M*, 1 mn 54 pour *Journal F*), et chaque extrait est une entité discursive ininterrompue.

- *Interview* : ce type de discours est caractérisé par son caractère spontané. Il est

bien entendu que ce terme recouvre de nombreuses situations différentes, allant de la conversation entre amis autour d'un repas à des situations beaucoup plus contraintes comme la discussion formelle ou l'interview. Il ne faut pas prendre ce qualificatif de 'spontané' comme référant à une seule situation de communication, mais comme une des caractéristiques définissant un type de discours dans lequel le locuteur n'a aucun support écrit et où il a le loisir de construire son discours au fur et à mesure. La situation que nous avons choisie n'est par conséquent qu'un échantillon de la réalité que recouvre le terme 'spontané'. En effet, la situation de communication est particulière : le locuteur en situation d'interview est un locuteur privilégié en ce sens où il se trouve au centre de la communication. Tout l'échange est construit de façon à mettre en valeur le discours de l'interviewé : son tour de parole n'est pas menacé, le contenu de son discours n'est a priori pas remis en question, et son interlocuteur n'intervient que pour initier le discours, l'enrichir ou le réinitialiser. De plus, l'interview radiophonique (ou télévisuelle) est orientée autour du locuteur invité, qui s'exprime en règle générale sur un sujet qui le touche de près, l'impliquant personnellement ou professionnellement. Malgré cet aspect 'idéal' de la communication, le caractère spontané du discours n'est pas remis en cause ; quand bien même le locuteur maîtrise parfaitement le sujet de son discours (promotion personnelle ou commerciale), il n'en construit pas moins son discours au fur et à mesure de la production. Il doit mettre en valeur ce qu'il a à défendre ou promouvoir. Cette stratégie n'est pas forcément éloignée des stratégies de communication de la conversation par exemple, archétype de la situation de communication spontanée idéale, si l'on considère que le locuteur se construit ou construit son identité sociale, par le discours. Rassembler un corpus de conversation pour une analyse prosodique fine n'est cependant pas une mince affaire : pour ne pas nuire à la spontanéité (l'enregistrement en chambre sourde étant idéal pour la qualité de l'enregistrement mais nuisible à la situation de communication), l'enregistrement devrait pouvoir se faire à l'insu des locuteurs, mais cette pratique a des limites déontologiques. De plus, les caractéristiques discursives de la conversation (chevauchements nombreux, variations erratiques du niveau d'intensité sonore, etc.) ne facilitent en rien une analyse prosodique fine, mais correspondent davantage à des descriptions linguistiques larges. C'est pourquoi notre choix s'est plutôt porté sur l'interview radiophonique, où la qualité sonore est bonne et le discours plus aisément analysable.

L'interview radiophonique a été décrite par Charaudeau (1984). Celui-ci relève 3 contraintes propres à cette situation de communication :

– *Le canal de l'oralité* : la communication va s'établir en fonction d'une 'mise en scène radiophonique' qui s'adresse à des auditeurs. La spécificité du canal radio induit une stratégie particulière d'évitement des chevauchements de parole, ce qui contribue à distinguer la situation d'interview de la conversation ou du débat par exemple. La distribution de la parole sera donc plus ou moins organisée à l'avance, afin d'éviter ces phénomènes nuisibles à la compréhension. De même, les temps morts et les silences devront être évités, afin d'assurer le continuum sonore propre à la radio.

– *La présence physique des deux interlocuteurs* : les gestes et mimiques inhérents à toute communication ne seront accessibles qu'aux interlocuteurs (interviewer et interviewé). L'auditeur devra probablement reconstruire l'information visuelle

manquante (marquage rythmique et expressif de la parole) par le biais d'une écoute attentive de l'intonation, du débit et de l'articulation phonétique.

– *Une situation à trois protagonistes* : la situation dialogique unissant l'interviewer et l'interviewé implique l'alternance de la parole. L'auditeur est un tiers à la fois présent et absent.

Les contraintes énumérées ci-dessus vont faire du style *Interview*, un style de parole spontanée très particulier. Dès lors, les résultats et observations que l'on pourra inférer de ce type de discours ne seront pas forcément généralisables à l'ensemble des situations de communication dites 'spontanées'.

Notre corpus est constitué de deux extraits d'interview enregistrés sur 'France culture'. Le premier locuteur (*Interview M*) présente en tant que directeur de la maison d'édition d'un dictionnaire connu, la version électronique de ce dictionnaire. La locutrice (*Interview F*) quant à elle, est interviewée sur son livre dédié à la carrière d'un homme politique. Sous couvert d'entretien culturel, le but de ces interviews est donc promotionnel. Dans les deux cas, les journalistes remplissent leurs rôles d'animateurs de l'entretien et contribuent à la mise en valeur du 'produit' présenté par leurs invités respectifs. Par conséquent, les journalistes ne coupent pratiquement pas la parole à leurs invités, et produisent de très rares régulateurs audibles. Le critère de sélection est simple : il s'agit d'avoir un extrait de parole non interrompu (compact), sorte de monologue constituant une entité thématique discursive. Il s'agit dans les deux cas de la réponse de l'interviewé à une question du journaliste, pendant laquelle les locuteurs développent entièrement librement leur pensée et leur discours. L'allocutaire privilégié de l'invité est le journaliste, mais le locuteur est conscient que son discours ne s'adresse pas qu'à lui, mais à de très nombreux auditeurs anonymes, absents de la situation de communication. Chaque locuteur a fait l'objet d'une interview dont nous avons extrait environ 2 minutes de parole monologique ininterrompue (2 mn 7 pour le *Interview M*, et 1 mn 38 pour *Interview F*).

3.1.2.2. Caractéristiques linguistiques communes aux différents types de discours

La sélection du corpus s'est opérée de sorte à éliminer un maximum de variabilité incontrôlable. Si l'on se réfère aux critères linguistiques définis dans le § 3.1.1, p. 129 et suivantes, notre corpus comporte une homogénéité sur des paramètres généralement négligés dans ce type d'études. Ainsi, l'unité de base est le *module*, que nous avons défini comme une unité discursive homogène sur le plan du discours. Il s'applique également à des corpus de parole non spontanée, si l'on considère que tout discours est interaction. Chaque extrait de notre corpus de parole non spontanée est un module à part entière, mais ils se caractérisent par leur caractère monologique (intrinsèque à la tâche linguistique). Dans la hiérarchie de la typologie de Vion (1992), les modules des styles *Journal* et *Lecture* sont congruents à l'unité immédiatement supérieure (l'interaction), ainsi qu'aux unités inférieures. Pour le style *Interview* en revanche, l'unité de base serait plutôt *l'acte de langage*, puisqu'il s'agit d'une réponse à une question. Cependant, chaque couple de question/réponse est un module à part entière, car le thème ainsi que le rapport entre les interactants évoluent constamment. Ces sous-modules constituent également des unités discursives homogènes monologiques. Nous considérons par conséquent avoir optimisé l'homogénéité des unités linguistiques de description.

Les extraits de parole peuvent tous être qualifiés de *parole publique*. Ils sont tous destinés à une audience, potentielle pour les types de discours non spontanés, et à la fois réelle et potentielle pour le style *Interview*. Les discours produits par l'ensemble des locuteurs ont des *finalités externes*, que ce soit la transmission d'une information (style *Journal*), ou d'un savoir (styles *Interview* et *Lecture*). Le caractère des diverses interactions est plutôt *formel*. Si l'on prend en considération les critères purement stylistiques définis par Eskenazi (1993), on constate que l'on pourrait situer les 3 types de discours à peu près au même endroit de son 'plan stylistique' tridimensionnel : les locuteurs sont plutôt situés vers la *strate* supérieure (à la fois de par leur statut d'énonciateurs privilégiés et leur statut social) ; les situations de communication induisent un rapport de *non familiarité* avec les auditeurs (réels ou potentiels), et les visées discursives poussent les locuteurs à produire une parole plutôt très *intelligible*.

3.1.2.3. Caractéristiques linguistiques distinguant les différents types de discours

Le processus d'encodage (spontané/non spontané) constitue la caractéristique distinctive essentielle. Le corrélat pratiquement indissociable de cette différence d'encodage est la structure interne de l'interaction : à *structure d'échange* pour le style *Interview*, *sans structure d'échange* pour les styles *Journal* et *Lecture*.

Les autres caractéristiques linguistiques étant communes et maîtrisées, nous pensons pouvoir par la suite interpréter la variabilité trouvée dans nos résultats par cette différence essentielle de processus d'encodage et de type d'interaction.

D'aucuns pourraient nous opposer que la variabilité stylistique observée dans notre étude ne peut être distinguée de la variabilité individuelle. En effet, chaque style de parole contient des locuteurs différents, interdisant une comparaison des stratégies discursives intra-individuelles. De nombreuses études consacrées à l'analyse des spécificités prosodiques des styles de parole proposent des comparaisons intra-individuelles (Lucci, 1983 ; Blaauw, 1992 ; Hirschberg, 1995 et sous presse ; Levin et al., 1982, entre autres). Ces types d'analyses supposent le plus souvent que les locuteurs relisent leurs productions de parole 'spontanée', afin de permettre une comparaison entre ces deux types de processus d'encodage en minimisant la variabilité syntaxico-sémantique. Or, aux dires des auteurs eux-mêmes, cette méthodologie induit une grande part d'artifice, et pervertit le but même de l'analyse de la parole 'en situation'.

Nous avons donc pris le parti de favoriser le 'naturel' de chaque situation de communication. Pour cela, aucun montage particulier n'a été réalisé, afin de saisir au plus près la réalité de la parole 'en situation'. Nous verrons que, si ce parti-pris nous empêche de nous dédouaner de la variabilité inter-individuelle, une certaine cohérence dans les stratégies prosodiques entre les locuteurs d'un même style de parole nous invite à minimiser la portée de ce choix méthodologique.

3.1.2.4. Intérêt méthodologique de la gradation des types de discours sur un axe spontané/non spontané

La particularité du processus d'encodage du style *Journal* radiophonique lui confère des caractéristiques pragmatiques empruntées à la fois à la lecture et à la parole spontanée. En essayant d'adapter son discours à un discours de type spontané, quelles sont les spécificités

prosodiques qu'un journaliste professionnel tente de reproduire ? Nous pensons que cette particularité du style journalistique, en se situant à l'intersection des deux autres types de discours sur l'axe spontané/non spontané, nous permettra de repérer plus précisément les caractéristiques propres à l'un ou l'autre des styles.

3.2. Niveaux de transcription et d'étiquetage du corpus

3.2.1. Considérations techniques

L'enregistrement des locuteurs du style *Lecture* s'est effectué en chambre anéchoïque sur un magnétophone analogique. L'enregistrement des styles *Journal* radiophonique et *Interview* s'est fait également sur bande magnétique à partir d'une base de données de plusieurs heures d'enregistrement radiophonique, constituée pour les besoins du groupe de recherche *GARP* (Astésano et al., 1995a). Les extraits ont ensuite été numérisés sur station Sun à une fréquence d'échantillonnage de 16 kHz avec une résolution de 16 bits. L'analyse des données a été effectuée à l'aide de l'éditeur de signal MES développé au Laboratoire Parole et Langage (Espesser, 1996). La détection de la fréquence fondamentale est réalisée à l'aide de trois méthodes de détection de la f_0 : les fonctions Peigne (Martin, 1981b ; Espesser, 1982), Amdf et Autocorrélation (Fohr & Laprie, 1989). A la suite de ces trois détections, un algorithme de décision effectue un vote afin d'obtenir la meilleure détection. Il faut noter que cette méthode de vote permet de pallier certaines faiblesses des méthodes de détection prises individuellement, comme les sauts d'harmoniques notamment sur les accents emphatiques (caractérisés par une augmentation brutale et rapide de la f_0), et la détection du voisé/non voisé.

3.2.2. Quels types de transcriptions pour quels types d'études ?

Toute analyse linguistique et prosodique suppose une étape préalable de transcription du corpus. Autessere et al. (1989) proposent, pour la transcription et l'étiquetage d'un corpus de parole, quatre niveaux de transcription :

- un niveau de transcription orthographique pour annoter les éventuelles erreurs de langage et événements inattendus ;
- un niveau de transcription phonologique qui détermine les frontières de morphèmes, de mots et de syntagmes, où les phones sont transcrits phonologiquement selon les symboles de l'API, et où des symboles diacritiques indiquent les enchaînements consonantiques et les liaisons obligatoires.
- un niveau de transcription phonotypique, interface entre la transcription phonologique et la transcription phonétique, qui prend en compte les variables contextuelles indépendantes du locuteur (règles d'élision de phonèmes par exemple) et qui s'affranchit des symboles de frontières. Des règles de réalisation archiphonémique des voyelles sont également prévues dans les cas phonologiques ambigus.
- un niveau de transcription phonétique enfin, seul niveau qui rend compte des réalisations effectives des locuteurs.

Ces étapes de transcriptions fines d'un corpus sont-elles nécessaires pour une étude des événements prosodiques ?

3.2.3. Niveaux de transcription adaptés à notre étude

Les étapes de transcriptions phonologique et phonotypique conviennent particulièrement aux études phonétiques de type acoustiques ou articulatoires. Une analyse des événements suprasegmentaux nécessite davantage la prise en compte des événements segmentaux susceptibles d'avoir une influence sur la réalisation prosodique d'un énoncé, tels les phénomènes d'interruption du flot de parole (coups de glotte, pauses, faux départs etc.), les phénomènes de présence ou d'absence de liaisons inattendus, les phénomènes de débit remarquables (sous-articulation d'une suite de phonèmes ou d'un passage), ainsi que les phénomènes segmentaux susceptibles de nuire à une étude fine des durées (assimilations de phonèmes, dévoisement ou voisement contextuel total d'un phonème etc.).

C'est pourquoi nous avons choisi dans notre étude de procéder à deux étapes essentielles de transcription : une étape de transcription orthographique fine, précédant une étape d'étiquetage phonétique large (étiquetage simultané au repérage des frontières phonémiques lors de l'étape de segmentation : cf. § « Segmentation phonémique », p. 143).

- *Niveau de transcription orthographique* : nous nous sommes inspirée des conventions de transcription orthographique du Groupe Aixois de Recherche en Syntaxe, qui propose une transcription fine des corpus de parole (Blanche-Benveniste & Jeanjean, 1987). Le texte doit être transcrit sans marque textuelle (ponctuation, majuscules, paragraphes etc.). Seuls les changements de locuteurs sont indiqués par un passage à la ligne. Nous ne conservons de leurs conventions que celles qui nous ont semblé pertinentes pour notre étude. Nous ne distinguons pas par exemple les différents types de pauses (courtes, moyennes, longues etc.) : cette étape de distinction se fera d'elle même après extraction des données issues de la segmentation. Nous distinguons par contre, dans la mesure du possible, les pauses respiratoires des pauses silencieuses vides : la plupart du temps, la différence est en effet perceptible et également visible sur le signal spectrographique (zone de bruit peu intense en moyennes fréquences) ; nous indiquons également les hésitations vocales (transcrites 'euh'). Enfin, nous indiquons les phénomènes de chevauchements d'énoncés, d'allongements vocaliques, de prononciation de consonne finale, de présence ou d'absence de liaisons remarquables, ainsi que tout autre fait de prononciation, de prosodie ou de débit, de rires, de bruits ou événements de la situation dignes d'intérêt pour notre étude.

Cette transcription orthographique est une étape préalable à l'étiquetage phonétique large et constitue en outre une aide à la décision pour la syllabation. (cf. § « La syllabe, unité rythmique », p. 42).

- *niveau d'étiquetage phonémique* : Nous nous situons à un niveau d'étiquetage phonétique large (i.e. sans symbole diacritique). Après écoute et transcription orthographique du corpus, notre étape d'étiquetage phonémique consiste à marquer ce que le locuteur a réellement dit : phénomènes de liaison, coups de glotte, timbre réalisé, pauses parmi lesquelles nous distinguons les pauses respiratoires, les pauses vides et les hésitations ; certains phénomènes de sous-articulation sont pris en compte, et lorsqu'une partie du discours, que ce soit un ou plusieurs phonèmes, une syllabe ou un mot, sont sous-articulés, nous prenons le parti de les mettre sous une même étiquette précédés d'un signe diacritique (!) afin de les marginaliser et de ne pas les intégrer par la suite dans les statistiques, notamment pour le traitement des durées des phonèmes ; outre les phénomènes

de sous-articulation qui sont à exclure des durées phonémiques, nous excluons également les phénomènes de coarticulation trop importante dont nous suspectons qu'ils ont des conséquences sur les durées intrinsèques des phonèmes : ainsi, les phénomènes de voisement ou dévoisement total des consonnes sourdes et sonores, ou encore le dévoisement total des voyelles hautes (en particulier /i/ et /y/) au contact de consonnes sourdes dentales (/t/, /s/) ou palatales (/k/, /j/). Pour ce qui est du timbre de certaines voyelles, nous sommes parfois amenée à prendre des décisions phonologiques lorsque leur timbre n'est pas tranché ; nous préférons cette solution à la transcription archiphonémique car, étant seule juge de la segmentation, nous pensons reproduire une certaine régularité dans ce genre de décision. Nous ne préjugeons pas de la réalisation phonétique du locuteur lors de l'étape d'étiquetage et de segmentation, même si la connaissance partagée des règles phonologiques du français peut dans une certaine mesure influencer notre perception. Ainsi, nous ne distinguons pas a priori les cas d'élision obligatoire, facultative ou interdite de la voyelle latente / /. Nous notons le / / lorsque le locuteur le réalise : nous gardons tout de même cette catégorie phonologique comme telle, i.e. nous ne choisissons pas de transcrire la réalisation de timbre de ce phonème (/ø/ /œ/ suivant le contexte phonétique et/ou prosodique) car son statut phonologique particulier de voyelle latente lui confère une durée intrinsèque propre. Nous ne présumons donc pas des réalisations du locuteur, et par conséquent, les problématiques purement phonologiques de type 'élision de consonnes liquides, élision du schwa, les phénomènes d'épenthèse (de type [uRs bla]), liaisons mal ou fort - à - propos, dévoisement ou voisement totaux, etc.' sont pris en compte et transcrits phonétiquement. Il est difficile de faire la différence entre une variabilité contextuelle allophonique d'un phonème et une variabilité contextuelle non prévisible et reproductible (comme parfois des phénomènes de dévoisement complet d'une occlusive sourde ou le voisement complet d'une occlusive sonore). Les phénomènes d'influence du contexte phonétique sont éliminés de la base de données phonémique par le signe diacritique (!) lorsqu'ils sont trop importants et non motivés phonologiquement. Les phénomènes de variantes contextuelles allophoniques sont par contre notés selon leur macro-classe phonémique : ex : si un locuteur prononce [l a le vo kl''] 'elle a les cheveux clairs' au lieu de [l a le vœ kl''], le [œ] de [vœ] ayant été postériorisé pour des raisons indéterminées, non phonologiques, sera transcrit non pas comme une réalisation allophonique mais selon la réalisation macro-phonémique à laquelle il appartient, i.e. [œ]. C'est ainsi que la transcription du /R/, phonème particulièrement instable en français, sera phonologique plutôt que phonotypique (/r/, /x/ ou /z/).

On ne doit pas cependant considérer que ce type de transcription phonétique est étroite ; nous considérons plutôt qu'il s'agit d'une transcription phonétique large, tout en n'étant cependant pas phonémique. On se situe quand même plus près du niveau acoustique (même si tous les événements acoustiques ne sont pas pris en compte s'ils ne sont pas phonologiquement pertinents, comme par exemple les événements articulatoires de type 'labialisation' ou 'palatalisation' etc.). Nous nous situons donc entre deux niveaux : le niveau acoustique (et les traces articulatoires très fines) et le niveau phonémique : si une réalisation phonétique contextuelle non allophonique doit nuire à la macro-classe phonémique dont elle est issue, elle sera précédée d'un signe diacritique pour la marginaliser afin de ne pas l'inclure dans les statistiques de durées.

Nous utilisons les conventions de transcription phonémique de SAMPA pour le français, développées dans le cadre du projet ESPRIT 2589 SAM (Wells et al., 1992).

La caractérisation phonémique est simultanée à la tâche de délimitation de frontière de l'unité phonémique.

3.2.4. Segmentation phonémique

3.2.4.1. Quel type de segmentation pour quel type d'étude ?

Une fois les conventions de transcription phonémique adoptées, se pose le problème du type de segmentation à appliquer au corpus afin de relever les aspects pertinents pour l'étude prosodique. Selon Meunier (1994), le but de la segmentation est double : il s'agit de maîtriser l'aspect continu de la parole, par conséquent de faire de la parole un objet d'analyse exploitable ; il s'agit ensuite de rechercher dans les événements du signal le caractère discret, donc linguistique, des unités de la langue. Un certain nombre de difficultés théoriques et pratiques sont associées à la détermination des frontières entre segments. Une des solutions adoptée parfois consiste à ne pas proposer de frontières et à se contenter de repérer le centre du segment acoustique correspondant à l'unité linguistique abstraite. Nous avons préféré indiquer explicitement les frontières. Il existe actuellement des systèmes semi ou complètement automatique de segmentation de la parole. Par rapport à un étiquetage manuel, ces dispositifs possèdent l'avantage d'un découpage déterministe, non dépendant de la subjectivité du segmentateur humain, et permettent surtout un gain de temps considérable. Cependant, ils comportent encore des lacunes et demeurent à l'état de prototypes. Nous avons donc opté pour un étiquetage manuel.

A l'instar de Meunier (1994), nous sommes consciente que les frontières ou discontinuités, indiquant les limites d'une unité phonétique sont uniquement issues d'un choix arbitraire opéré en fonction d'indices pertinents par le segmentateur. La notion de segment est une décision et non une vérité incontestable. Le but est de dresser une liste de critères permettant d'assurer une certaine cohérence. Pour la tâche de segmentation, la faculté auditive a des limites : d'une part, elle n'est d'aucune aide pour la délimitation des frontières phonétiques, car elle est inadaptée à l'écoute de segments très courts ; d'autre part, elle est terriblement subjective et soumise à de nombreux facteurs externes. Elle est cependant utile pour la transcription et le repérage global des événements. La recherche des discontinuités pertinentes (zones de passage d'une unité phonétique à une autre) se fera donc par repérage visuel sur l'oscillogramme et le spectrogramme, selon un certain nombre de critères issus des représentations canoniques qu'a le segmentateur des unités phonétiques.

3.2.4.2. Critères et choix arbitraires de segmentation phonémique

Nous présentons ici les problèmes typiques rencontrés lors de l'exercice de segmentation pour lesquels nous avons pris des décisions permettant une certaine cohérence et un maximum de reproductibilité. Nous rappelons que l'objectif majeur de notre segmentation phonémique est d'analyser la gestion par les locuteurs des durées segmentales sous l'effet de contraintes prosodiques et discursives. Une segmentation phonémique rigoureuse est donc nécessaire afin de réduire une source de variabilité supplémentaire non phonologique (bruit). Nos 10 minutes de parole ont donc été segmentées et étiquetées

manuellement à l'aide de l'oscillogramme et du sonagramme. Nous rappelons que l'utilisation de l'ouïe ne se faisait que pour le repérage global des suites sonores. Le repérage visuel des unités phonétiques s'est fait sur les bases suivantes :

Segmentation des voyelles :

- Fréquemment, des coups de glotte se produisent dans certains cas d'enchaînement vocalique, ou pour les voyelles à l'initiale de mot ou suivant une pause. Deux possibilités de segmentation se présentent alors, en fonction du contexte prosodique : soit il s'agit d'un 'accident' phonatoire et le coup de glotte fait partie intégrante de la voyelle ; l'étiquette de la voyelle est alors posée au début du phénomène de coup de glotte ; soit l'apparition du coup de glotte est conditionnée par le contexte prosodique (accentuation initiale rythmique ou emphatique ; enchaînement vocalique entre deux unités prosodiques différentes) et dans ce cas là, le coup de glotte est dissocié de la voyelle et porte une étiquette propre, au même titre qu'un événement prosodique de type 'pause'. Les étiquettes des voyelles sont posées sur les parties stables des voyelles, les parties instables constituant l'étiquette du coup de glotte.

- Voyelle suivie d'une occlusive ou d'une fricative sourde : l'amortissement de la voyelle, contraint par les phénomènes de coarticulation, appartient à la consonne suivante. La segmentation de la consonne se fait par conséquent à la fin de F2 et F3.

- Consonne vocalique + voyelle : les transitions entre la voyelle et la consonne sont lentes. Le début de la voyelle correspond à la stabilisation de la structure spectrale.

- Occlusive ou fricative sourde + voyelle : les voyelles hautes, /i/ /y/ notamment, peuvent être partiellement ou totalement dévoisées. La frontière des voyelles correspond à la structure formantique et non au début du voisement.

- Voyelle en position finale avant pause : la phase d'amortissement de la voyelle est comprise dans l'étiquette de la voyelle, contrairement aux cas de coarticulation avec des consonnes. Le relâchement de la voyelle ne peut en effet pas se produire pendant la prise de souffle, ou un silence.

Segmentation des consonnes :

- Occlusives sourdes à l'initiale d'un énoncé ou après pause silencieuse : il n'y a aucun moyen de savoir où commence la tenue de l'occlusive. Ces phonèmes sont donc marqués au début de l'explosion par le signe diacritique (!) de sorte à les exclure des statistiques de durées phonémiques.

- Consonne en position finale avant pause : on observe très souvent l'apparition d'une voyelle d'appui, notée [], que cette voyelle ait une valeur phonologique ('basse', 'rate' etc.) ou non ('Irak', 'rapt' etc.).

- Groupe d'occlusives sourdes : il n'est pas rare de trouver une seule phase de tenue + explosion pour l'articulation conjointe de deux occlusives sourdes (ex : 'apte'). Dans ces cas là, une seule étiquette est posée pour les deux occlusives, avec un signe diacritique (!) pour les exclure des statistiques de durées.

- Voyelles + fricatives sourdes : le début de la fricative est posée à la fin de la

structure formantique de la voyelle ; la voyelle est en effet coarticulée avec le début de la fricative (amortissement de la voyelle).

- Consonnes liquides : L'utilisation de la courbe d'intensité est parfois nécessaire à la segmentation des liquides intervocaliques notamment. Le cas du /R/ est particulièrement épineux car il s'agit d'un phonème polymorphe, très sensible à son entourage phonétique: il est segmenté immédiatement après la barre d'explosion d'une occlusive sourde ; lorsqu'il suit une voyelle, il est segmenté dans la partie formantique perturbée de la voyelle ; lorsqu'il suit une fricative, il débute lorsque le pôle de bruit de la fricative est perturbée et moins intense.

3.2.5. Base de données phonémique de notre corpus

Se lancer dans l'étude des phénomènes accentuels et plus précisément des durées segmentales, suppose de prendre toutes sortes de précautions méthodologiques. Notre étude ne porte pas expressément sur la description des durées acoustiques des voyelles et des consonnes du français, mais plutôt sur leur comportement phonologique dans le cadre de la syllabe, sous l'effet des contextes prosodiques et discursifs.

L'organisation prosodique prend appui sur le contenu segmental et les deux niveaux s'influencent mutuellement. Afin de retirer l'information prosodique pertinente, il est nécessaire de contrôler les événements de niveau inférieur afin de limiter les sources de bruit aux niveaux d'interprétation supérieurs. Une étape préliminaire de catégorisation et de traitement de données est nécessaire avant les étapes d'étiquetage phonologique et prosodique, notamment pour l'étape délicate de la normalisation.

Notre travail se fonde, par essence, sur l'étude de discours non contrôlés sur le plan lexical, syntaxique ou sémantique. Nous n'avons par conséquent aucun contrôle du contenu segmental. Ainsi, aucun corpus n'est phonétiquement équilibré, i.e. que tous les contextes phonétiques ne sont pas représentés. Cependant, le corpus est statistiquement équilibré (cf. Figure 9, p. 146).

Avant d'aborder les notions de durées intrinsèques des phonèmes, de contexte phonétique et de distribution des phonèmes, nous nous sommes demandée si notre corpus est représentatif des phonèmes du français, i.e. si tous les phonèmes sont équitablement représentés en fonction de leur occurrence moyenne, et ceci afin de ne pas biaiser les statistiques de durée, dans la perspective d'une normalisation. Pour cela, nous avons comparé l'occurrence des phonèmes de notre corpus avec les pourcentages d'occurrence des phonèmes du français de Wioland (1985). Cette étude constitue en effet une référence de par la taille du corpus utilisé : 20 échantillons de français parlé non méridional composés chacun d'une suite de 10 000 phonèmes, soit un corpus de 200 000 unités phonémiques. Un intérêt supplémentaire de cette base de données réside dans le fait que les types de discours étudiés sont très diversifiés, allant de la parole conversationnelle à la parole déclamée théâtrale.

Les résultats de Wioland montrent que les voyelles et les consonnes représentent respectivement 43.45 % et 56.55 % des phonèmes du français. Les occurrences des phonèmes de notre corpus sont sensiblement comparables à celles de la base de référence de Wioland, et ce quel que soit le type de parole ou le locuteur (Figure 9, p. 146. Cf également Astésano (1999) pour une comparaison détaillée des pourcentages d'occurrence

phonémique par locuteur et par style de parole).

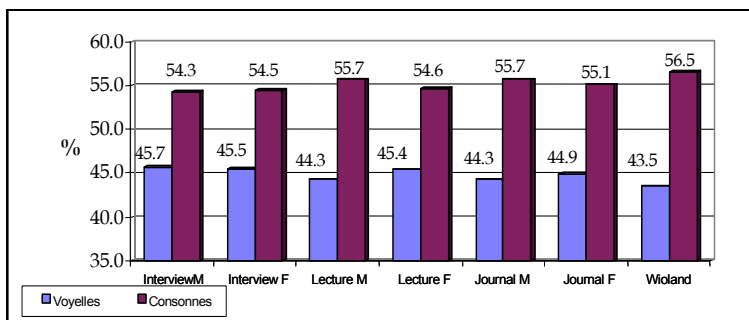


Figure 9 : Pourcentages d'occurrences des voyelles et des consonnes par locuteur, comparés à ceux de la base de données de Wioland (1985).

Notre corpus, bien que non contrôlé, est donc représentatif des occurrences des phonèmes en français, preuve que les échantillons de parole que nous avons sélectionnés pour chaque locuteur sont suffisamment importants.

Ci-après (cf. Figure 10, p. 146) sont représentées les durées moyennes en millisecondes des consonnes et des voyelles par locuteur. On peut observer d'ores et déjà une légère tendance d'augmentation de la durée des voyelles entre la parole spontanée (*Interview*) et la *Lecture*, l'allongement étant plus conséquent pour le style journalistique professionnel. Ces tendances sont sans doute à mettre en relation avec le débit de parole (cf. Tableau 4, p. 150).

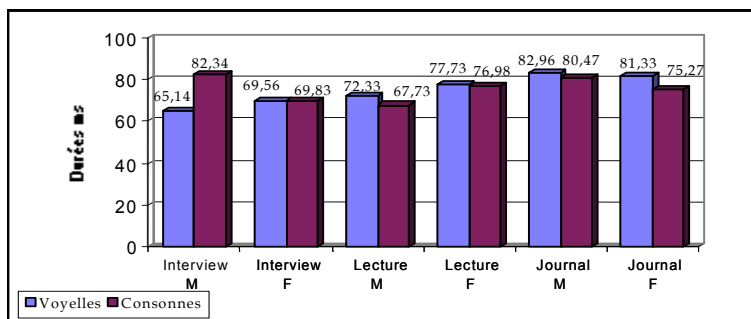


Figure 10 : Durées en millisecondes des voyelles et des consonnes par locuteur

3.2.6. Classification des phonèmes et affranchissement par rapport au contexte

Il est pertinent, pour l'étude de la durée, de regrouper les phonèmes en macro-classes en fonction de critères articulatoires et acoustiques, afin d'éviter toute redondance et de multiplier inutilement les variables à analyser. Plus précisément, on a coutume de regrouper les consonnes selon leur mode de production (voisé/non voisé) et d'articulation (occlusives, constrictives etc.). Les voyelles sont quant à elles regroupées selon leur apertures et le critère d'oralité/nasalité. De nombreuses études ont en effet mis en évidence le lien entre les critères articulatoires et les durées des phonèmes (cf. § « Variabilité induite par les

événements articulatoires (durées intrinsèques) », p. 170). A l'instar de Di Cristo (1985), nous choisissons de classer les voyelles orales en voyelles hautes, moyennes et basses, selon leur place dans le triangle vocalique. Les **voyelles hautes** correspondent aux **voyelles fermées** ([i], [y], [u]); les **voyelles moyennes** correspondent aux **voyelles mi-fermées**, ([e], [ø], [o], [ɘ]) et les **voyelles Basses** correspondent aux **voyelles mi-ouvertes** et **ouvertes** ([ɛ], [œ], [ɔ], [a]). Les **Voyelles Nasales** sont ([ʎ], [œ̃], [ã], [õ]). Le Tableau 2 p. 147 illustre les phonèmes du français, classés en fonction de critères articulatoires (Ghio, 1997). Nous avons préféré classer le [ɘ] parmi les voyelles moyennes car nous considérons que son articulation est plutôt mi-fermée, alors que le tableau ci dessous le qualifie plutôt de voyelle ouverte.

Tableau 2: Les phonèmes du français (Ghio, 1997 : 62)

CONSONNES			
<i>occlusives non voisées</i>	[p] paie	[t] taie	[k] quai
<i>occlusives voisées</i>	[b] baie	[d] dé	[g] gai
<i>occlusives nasales</i>	[m] mais	[n] nez	[ŋ] gagner
<i>constrictives non voisées</i>	[f] fait	[s] sait	[ʃ] chez
<i>constrictives voisées</i>	[v] vais	[z] zéro	[ʒ] geai
<i>semi-consonnes</i>	[w] ouais	[H] huer	[j] yé-yé
<i>liquides</i>	[l] lait	[R] raie	
VOYELLES			
<i>fermées</i>	[i] lit	[y] lu	[u] loup
<i>mi-fermées</i>	[e] les	[ø] leu	[o] lot
<i>mi-ouvertes</i>	[ɛ] lait	[œ] leur	[ɔ] lotte
<i>ouvertes</i>	[a] la	[ɘ] le	
<i>ouvertes (nasales)</i>	[ʎ] lin	[œ̃] l'un	
<i>ouvertes (nasales)</i>	[ã] lent	[õ] long	

Une fois l'étiquetage des phonèmes en macro-classes effectué pour chaque locuteur, nous effectuons une analyse de variance (dorénavant Anova) afin de mettre en évidence la gestion des durées brutes de chaque locuteur en fonction de la macro-classe phonémique. Cette étape préalable nous permet de nous assurer que notre corpus est représentatif des durées des français, et cohérent par rapport aux résultats trouvés dans la littérature.

L'Anova a été réalisée avec le logiciel Statview pour PC. Les locuteurs et les macro-classes ont été pris comme variables indépendantes, et la durée phonémique brute comme variable dépendante.

Avant de décrire plus avant les résultats de l'Anova, précisons qu'une pré-analyse graphique des données a été effectuée à l'aide du logiciel S+ sur station Sun², afin de nous assurer de la fiabilité de nos résultats. L'examen des graphiques indique que la distribution

² Nous tenons à remercier Michel Piterman pour son aide précieuse dans la mise en œuvre de la méthodologie statistique concernant l'analyse des données segmentales, et dans l'interprétation des résultats à la sortie du logiciel S+.

pour le facteur 'locuteurs' n'est pas gaussienne, car elle s'étale vers la droite, signe de l'influence des durées longues sur la moyenne. L'intervalle interquartile, indice de la dispersion des données autour de la médiane, est relativement constant quel que soit le locuteur ou la macro-classe. On en déduit donc que la condition d'homosédasticité est remplie. Ce dernier critère est essentiel pour la fiabilité des analyses Anova et résultats à venir. Il est en effet signe d'une certaine robustesse de notre corpus, à travers les types de discours et les locuteurs.

Les résultats de l'Anova montrent que les macro-classes et les locuteurs sont distingués sur la base des durées phonémiques brutes (respectivement F-value (105,026), P-value (<,0001), et F-value (21,772), P-value (<,0001)). La Figure 11, p. 149 et le Tableau 3, p.148 montrent la durée moyenne des macro-classes phonémiques tous locuteurs confondus.

Tableau 3 : Durées moyennes brutes (ms) par macro-classe, tous locuteurs confondus

Tableau des Moy. pour durées				
Effets : Macroclasses				
	Nombre	Moy.	Dév. Std	Err. Std
ConsNas	385	70.727	23.007	1.173
FricNV	426	105.215	33.189	1.608
FricV	266	73.439	24.134	1.480
Liquides	867	58.272	27.798	.944
OccNV	775	91.890	29.074	1.044
OccV	340	66.979	25.726	1.395
SemiCons	133	55.802	25.067	2.174
VoyBas	738	75.151	36.087	1.328
VoyMoy	873	67.285	32.321	1.094
VoyHaute	564	67.899	34.557	1.455
VoyNas	446	100.242	41.471	1.964

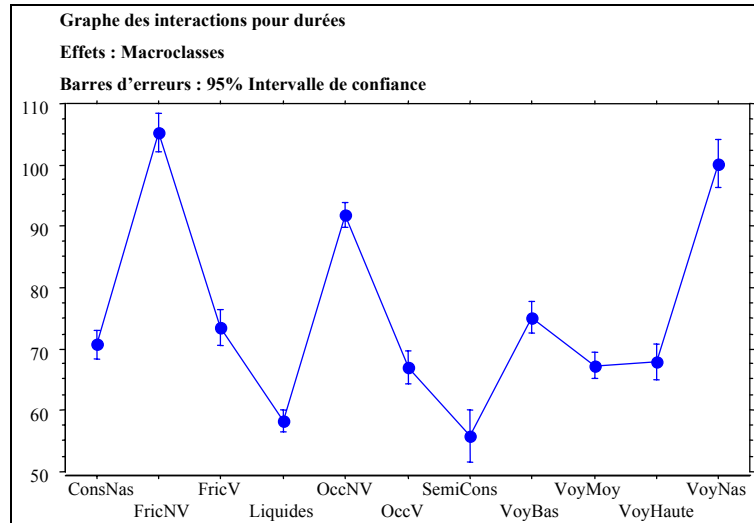


Figure 11 : Durées moyennes brutes (ms) des macro-classes, tous locuteurs confondus

Pour les consonnes, le mode d'articulation a une influence sur la durée des segments ; ainsi, les fricatives ont-elles une durée plus importante que les occlusives. A l'intérieur de ces macro-classes, le mode de production induit une supériorité des macro-classes non voisées sur les macro-classes voisées. Les test PLSD de Fisher (niveau de signif. 5%) comparant les macro-classes deux à deux, indiquent que les consonnes nasales ont sensiblement la même durée intrinsèque que les fricatives et occlusives voisées. De même, les semi-consonnes et les liquides ne se distinguent pas sur la base de leur durée intrinsèque. Nos résultats montrent également un comportement attendu des voyelles ; ainsi, les voyelles nasales sont elles nettement plus longues que les autres types de voyelles. Les voyelles basses présentent une durée significativement plus importante que les voyelles hautes et les voyelles moyennes, ces deux dernières macro-classes ne se distinguant pas significativement entre elles sur la base de la durée brute (cf. § « Caractéristiques des durées segmentales en français », p. 170).

L'interaction entre les locuteurs et les macro-classes est également significative, indiquant que les locuteurs ont des stratégies de durées phonémiques intrinsèques propres, et assez variables (F-value (4,189), P-value (<,0001)). On voit sur la Figure 12 p. 150 que les locuteurs du style *Journal* ont des durées intrinsèques plus cohérentes que les locuteurs des styles *Lecture* et *Interview*. Ces différences inter-locuteurs de durées intrinsèques sont sans doute à relier au débit, et le caractère contrôlé du discours des locuteurs (lecteurs-interprètes professionnels). Les tests PLSD de Fisher indiquent également dans la comparaison des locuteurs deux à deux, que les paires de locuteurs *Interview M/Lecture F*, *Interview F/Lecture M* et *Lecture F/Journal F* ne se distinguent pas sur la base des durées phonémiques. On peut relier cela directement au débit hors pauses (vitesse d'articulation), calculé comme le nombre de syllabes par seconde : les débits respectifs de chaque paire de locuteurs sont en effet très proches (Tableau 4).

Tableau 4 : Débit hors pause par locuteur (nombre de syllabes par seconde)

Locuteurs	Débit ou Vitesse d'articulation (syllabes/seconde, hors pauses)
Interview M	5.9
Interview F	6.1
Lecture M	6.3
Lecture F	5.6
Journal M	5.2
Journal F	5.5

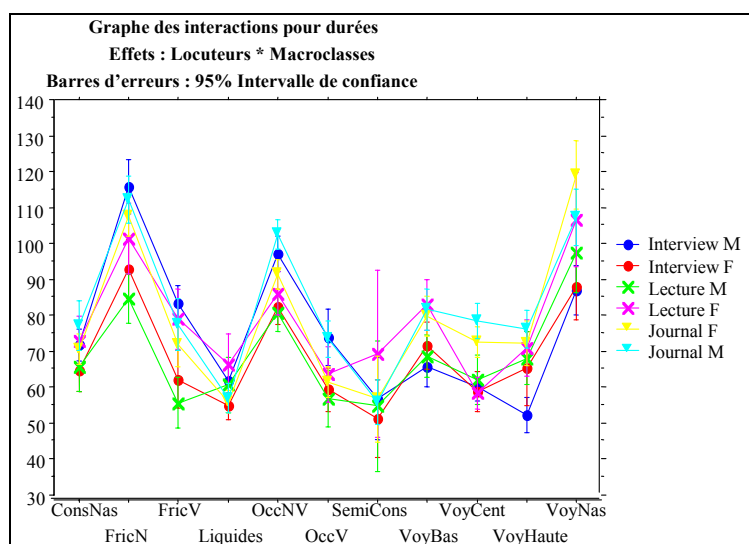


Figure 12 : Durées moyennes brutes (ms) des macro-classes par locuteur.

Le calcul de durée des phonèmes regroupés en macro-classes s'est effectué sans prise en compte du contexte phonémique, distributionnel ou prosodique (voir les descriptions de l'influence de ces divers facteurs sur les durées intrinsèques des phonèmes dans le § 2.2, p. 170). Nous rappelons que notre corpus n'est pas contrôlé, ni phonétiquement équilibré. Par conséquent, l'influence de ces divers facteurs sur les durées intrinsèques des phonèmes de notre corpus n'est pas contrôlée. Nous disposons cependant d'environ 1000 occurrences de phonèmes par locuteur, et nous avons vu par ailleurs que leurs pourcentages d'apparition respectent les statistiques d'occurrence des phonèmes du français issues de la base de donnée de Wioland (1985) (cf. § « Base de données phonémique de notre corpus », p. 145). Nous considérons donc que les variations contextuelles sont estompées par le nombre élevé d'occurrences. Les résultats de durée phonémique que nous obtenons vont d'ailleurs dans le sens des descriptions trouvées dans la littérature, descriptions élaborées la plupart du temps à partir de corpus contrôlés, phonétiquement équilibrés et contraints prosodiquement

(fricatives plus longues qu'occlusives ; segments non voisés plus longs que segments voisés ; voyelles basses plus longues que voyelles hautes etc.). Comme le soulignent Crystal & House (1982) à la lumière de leurs données issues de parole continue, les différences inter-locuteurs sont plus importantes que les différences imputables aux seules contraintes contextuelles.

3.2.7. Syllabation et étiquetage des phonèmes en Attaque/Noyau/Coda

L'étude de l'accentuation suppose de définir au préalable une unité d'analyse qui constitue le cadre interprétatif des événements observés. Toute décision de ce type revêt un caractère arbitraire et limitatif, mais est cependant incontournable. Nous avons choisi la syllabe comme unité de réalisation optimale des paramètres accentuels, et ceci au vu des arguments développés dans le paragraphe sur la syllabe, p. 42.

Nous avons utilisé les fichiers phonémiques de notre corpus comme base pour la concaténation en syllabes. Le passage d'unités phonémiques en unités de constituant supérieur suppose la prise en compte de règles phonétiques et phonologiques, permettant de définir les frontières syllabiques. Comme pour les décisions de segmentation en phonèmes (cf. 3.2.4.2, p. 143), la délimitation de frontières syllabiques doit se fonder sur des critères stricts, généralisables et reproductibles.

3.2.7.1. Critères définitionnels pour la tâche de division syllabique

A la suite de De Saussure notamment, les auteurs se sont attachés à définir des critères universels de division syllabique. Un large consensus s'établit autour du critère de pic d'intensité sonore comme noyau syllabique, condition de l'existence même de la syllabe. Ce critère confère à la voyelle une place centrale dans la syllabe, d'autant plus qu'elle peut constituer une syllabe à elle seule, alors qu'il n'en est rien pour les consonnes. Le critère de sonorité concerne également certaines consonnes (nasales, liquides, semi-consonnes) susceptibles de porter des marques prosodiques, et est souvent invoqué pour les intégrer dans le noyau syllabique. Cependant, ce critère ne répond pas à la question des limites de la syllabe.

Dans cette perspective, des auteurs comme Delattre (1966b) ont appliqué les principes de syllabation fondés sur les séquences d'aperture inspirés de De Saussure. Il s'agit d'une démarche descendante induite par les connaissances implicites du locuteur natif, qui s'oppose à la démarche ascendante de certains auteurs, visant à déduire de la substance des règles phonologiques. Les 5 degrés d'aperture qu'il définit correspondent, sur une échelle d'aperture croissante, aux modes d'articulation distinguant les occlusives, les fricatives, les nasales, les liquides et les semi-consonnes. Par conséquent, la syllabe abstraite [pSljawRst] comprenant 8 consonnes et 1 voyelle est prononçable, car elle est constituée d'ouvertures croissantes et décroissantes ininterrompues.

La règle est que la division syllabique s'opère avant la consonne la plus fermée dans une série intervocalique (ex : 'pa-trie' ; 'par-tie'). Lorsque deux consonnes consécutives ont la même aperture, la division s'effectue entre les deux consonnes (ex : 'ac-tif').

Delattre (1966b) montre que la division syllabique s'effectue en français et en espagnol sans tenir compte des frontières de mots dans le groupe de sens, ce qui favorise la syllabation ouverte. En anglais et en allemand en revanche, la forte anticipation

consonantique favorise l'occurrence de syllabes fermées. La structure syllabique est donc fortement liée à la structure morphophonologique et lexicale de chaque langue.

3.2.7.2. Critères de division syllabique appliqués à l'objet d'étude

Nous avons choisi d'appliquer à notre corpus les critères de délimitation des frontières syllabiques définis par Delattre (1966b), car ils respectent les règles phonotactiques du français et correspondent aux divisions syllabiques intuitives. Ce type de critères favorise de surcroît la syllabation ouverte, propre au français (Wioland, 1985). Il s'agissait pour nous de respecter autant que possible les structures internes des syllabes, car en tant que constituants métriques, elles participent pleinement au rythme du français. De fait, les syllabes constituent les unités de base de l'analyse de nos données temporelles et tonales.

• Syllabation :

Comme pour notre méthodologie de segmentation des phonèmes, nous nous fondons sur les réalisations des locuteurs dans notre tâche de délimitation des frontières syllabiques. Certains cas de figure récurrents nous ont amenée à appliquer des critères de syllabation ad hoc et à adapter ainsi les critères de Delattre (1966b) à la réalisation phonétique de l'énoncé :

– le français étant caractérisé par une forte anticipation vocalique et par le phénomène de liaisons, notre syllabation ne respecte pas les frontières de mots morphologiques. Nous avons pris le parti cependant de favoriser la syllabation lexicale dans les cas d'ambisyllabité, correspondant généralement à des frontières de constituants prosodiques. Ainsi, lorsque le locuteur *Journal M* par exemple, produit la partie d'énoncé « (confirme)(le figaro) » (les parenthèses indiquent les frontières de constituants mineurs), nous avons choisi de poser les frontières syllabiques comme suit : « con-firme-le-fi-ga-ro », favorisant ainsi la dépendance des phonèmes par rapport aux constituants prosodiques.

– les phénomènes d'épenthèse ou de réalisation d'un / / muet en fin de mot sont traités de la même façon ; ils ne peuvent constituer une syllabe à part entière et sont donc rattachés à la syllabe précédente. Un mot comme « syllabe » sera divisé en « syllabe » et non pas « sy-lla-be ». C'est la notion de macro-syllabe, constituant métrique à part entière selon Selkirk (1984) (correspondant dans sa hiérarchie prosodique, au niveau du 'Pied').

• Etiquetage des constituants syllabiques :

Une fois les frontières syllabiques délimitées, nous avons procédé à un étiquetage phonologique des constituants syllabiques en termes d'attaque (A), noyau (N) et coda (C) (cf. § « Structure syllabique », p. 43 pour une argumentation sur ce point). Ce sont les constituants immédiats de la syllabe. Une syllabe CV est représentée phonologiquement comme ayant une coda vide ($A N C_{(\emptyset)}$) ; une syllabe VC possède une attaque vide ($A_{(\emptyset)} N C$) ; une syllabe comportant un groupe de consonnes initial et/ou final possède une attaque complexe et/ou une coda complexe, et non pas plusieurs attaques et plusieurs codas, fonction du nombre de phonèmes les constituant. L'attaque et la coda sont donc des constituants phonologiques à l'intérieur desquels s'opèrent des règles et des contraintes phonétiques et articulatoires (compensation de durée, assimilation etc.) (cf. Figure 13, p.

153).

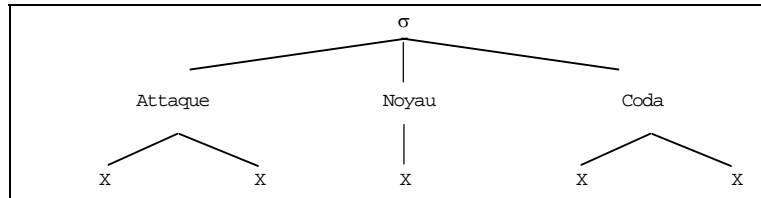


Figure 13 : Constituants immédiats de la syllabe : Attaque (A), Noyau (N), Coda (C). (σ = syllabe ; X = phonèmes)

Nous prenons le parti d'intégrer les consonnes sonorantes dans l'attaque syllabique (partie consonantique) Ainsi, les semi-consonnes vont appartenir à l'Attaque, et non pas au Noyau comme le suggèrent les tenants de la syllabation fondée sur l'échelle de sonorité. Nous sommes consciente cependant des arguments avancés par certains phonologues, qui considèrent que l'appartenance des semi-consonnes aux éléments d'Attaque ou de Noyau dépend de règles phonologiques précises. Ainsi, le mot 'oie' posséderait une Attaque vide, permettant le rattachement d'une consonne dans le cas de liaison (/leswa/), alors que le mot 'watt' ne permettrait pas de phénomène de liaison (/lewat/), son Attaque étant déjà occupée par une semi-consonne. Selon Durand (1990), ces phénomènes expliqueraient l'ambivalence de positionnement des semi-consonnes dans la syllabe. Nous ne partageons pas ce point de vue, car nous postulons non seulement l'existence d'Attaques vides, mais également le cas d'Attaques complexes, ce dont Durand (1990) ne semble pas tenir compte dans ce cas précis. Selon nous, la liaison ne s'opère pas dans /lewat/ tout simplement parce qu'il s'agit d'un mot d'emprunt.

Par conséquent, la suite de phonèmes dans le mot 'droit' sera ainsi étiquetée (cf. Figure 14) :

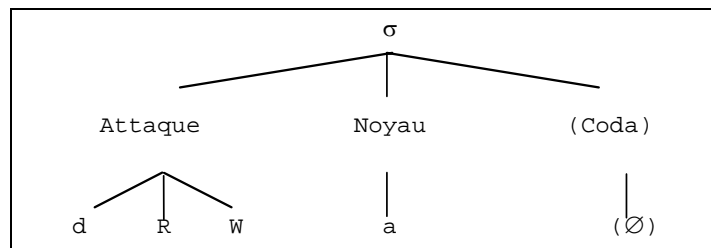


Figure 14 : Représentation du mot « droit » en constituants syllabiques

3.2.8. Étiquetage prosodique

3.2.8.1. Pauses

Conjointement à l'étape de segmentation phonémique, nous avons opéré une catégorisation des phénomènes d'interruptions de la chaîne parlée. Nous avons ainsi distingué les pauses silencieuses, les coups de glotte, les hésitations vocales, et tous les phénomènes de faux départs et de reprises. Ces derniers ont été marqués d'un signe (!) et ne

sont pas pris en compte dans notre analyse.

Pour la segmentation des pauses, un seuil de 240 ms est proposé par Goldman-Eisler (1968) afin d'éviter les confusions possibles entre les pauses silencieuses et les tenues d'occlusives. Or, l'observation empirique de notre corpus a révélé que ce seuil était beaucoup trop élevé. En effet, nos données indiquent que la durée moyenne des occlusives de nos locuteurs varie de 63 ms à 121 ms. Nous avons donc fixé le seuil inférieur des phénomènes d'interruption à 150 ms. Nos critères de segmentation incluant dans la durée des occlusives la période de VOT, il est clair qu'un seuil de 150 ms était amplement suffisant pour distinguer les pauses silencieuses des tenues d'occlusives. Enfin, le croisement entre les données acoustiques et des critères linguistiques tels que la fonction des pauses silencieuses nous ont permis d'écarter tout danger de confusion dans notre tâche de segmentation manuelle. La segmentation des pauses a été réalisée objectivement à l'aide du signal oscillographique et du spectrogramme à bandes larges. Nous avons distingué plusieurs types de pauses, en nous inspirant des critères de Goldman-Eisler (1968), de Grosjean & Deschamps (1972, 1973, 1975) et de Guaïtella (1991) :

- les pauses silencieuses : à la différence de Grosjean (1980), nous ne spéculons pas sur les fonctions des pauses silencieuses (fonction d'hésitation, de marquage syntaxique etc.). Par contre, nous classons les pauses silencieuses selon leur forme, en distinguant les pauses vides des pauses respiratoires. Cette distinction ne concerne que les locuteurs des styles *Interview* et *Journal*, car les pauses respiratoires sont parfaitement audibles et perceptibles visuellement sur le spectrogramme (bande de fréquences moyennes peu intenses). Par contre, nous avons créé une macro-classe 'Pauses' comprenant les pauses respiratoires et les pauses vides pour le style *Lecture*, car malgré l'enregistrement de très bonne qualité réalisé en chambre sourde, les pauses respiratoires sont difficilement repérables auditivement et visuellement. Cette différence de réalisation des pauses silencieuses du style *Lecture* est sans doute due au processus de planification de la respiration propre à ce type d'encodage.

- Les pauses sonores (à valeur d'hésitation) qui ne se retrouvent que dans le style *Interview*. Les deux autres types de parole étant fondés sur du texte écrit, les locuteurs n'ont pas recours aux pauses sonores à valeur d'hésitation qui se différencient des pauses silencieuses à valeur d'hésitation par la fonction qu'elles remplissent : le locuteur indique à son auditoire qu'il est en phase de planification de son énoncé, mais qu'il désire garder la parole, situation qui n'a pas raison d'être en parole lue ou interprétée. Si les locuteurs des styles lus devaient hésiter, ce ne serait pas pour une raison de planification de leur énoncé (au sens appropriation du message), mais de planification de leur lecture. Cette différence de processus d'encodage se reflète sur le plan de la gestion des pauses par l'utilisation exclusive en lecture ou lecture interprétée de pauses silencieuses.

Grosjean et Deschamps distinguent deux types de pauses sonores :

- les pauses remplies, c'est à dire les [œm] et les [] isolés ;
- et les allongements syllabiques comprenant, outre les allongements vocaliques, les [], [œ], [] ou [œm] en fin de mots.

Nous n'avons pas retenu dans notre distinction les allongements vocaliques définis par

ces auteurs pour deux raisons principales : il nous semblait difficile de différencier nettement l'allongement syllabique à valeur d'hésitation, de l'allongement syllabique caractéristique de l'effet de l'accentuation sur la durée syllabique. Pour établir une différence entre les deux, il aurait été nécessaire de calculer la durée moyenne des syllabes accentuées pour chaque locuteur et de ne conserver que les allongements syllabiques supérieurs à la moyenne + deux écarts-types, par exemple. Cette démarche aurait été coûteuse pour un résultat peut-être peu précis et peu pertinent. Nous avons d'autre part remarqué dans les extraits du style *Interview* que la quasi totalité des phénomènes d'allongements se manifestent par des hésitations vocales en fin de mots, et non pas par des allongements syllabiques "purs". C'est ce qui a motivé notre choix. Nous avons donc adapté notre typologie des pauses aux corpus observés et à la stratégie des locuteurs dans le domaine des pauses. Les allongements syllabiques à valeur de pauses étant le plus souvent des hésitations vocales, même si elles sont 'liées acoustiquement' à certains allongements finaux (Guañtella, 1991), nous avons choisi de ne conserver dans la catégorie des pauses sonores que les hésitations vocales.

Notre corpus comporte donc des pauses vides, des pauses respiratoires et des hésitations vocales. A cela, il faut rajouter ce que nous avons appelé 'interruptions', i.e. les coups de glotte réalisés entre des syntagmes prosodiques, ou bien avant un accent initial rythmique ou emphatique. Si les hésitations vocales sont uniquement l'apanage du style *Interview* pour des raisons évidentes, les autres types d'interruption de la chaîne segmentale se retrouvent dans tous les styles.

3.2.8.1.1. Critère de grammaticalité des pauses :

On trouve souvent dans la littérature l'analyse des pauses en fonction de leur distribution. Plus précisément, les auteurs cherchent à déterminer dans quelle mesure les pauses se trouvent à des jonctions grammaticales ou non de la chaîne parlée. La grammaticalité serait un facteur facilitant le décodage et une aide à une transmission optimale de l'information sémantique.

"Ideally the act of speaking serves the communication of meaning. This is achieved best if the gaps introduced into the stream of speech are coincident with semantic groupings. Ideally then pauses should be found between phrases, at the joint points between clauses or at the end of sentences; in short they should be semantically determined and occur at grammatical junctures." (Goldman-Eisler, 1968 : 13).

Selon cet auteur, les points de rupture non grammaticaux sont les suivants :

- au milieu d'un syntagme
- entre des mots ou des syntagmes qui se répètent
- au milieu d'un verbe composé

Nous avons choisi de ne pas avoir recours à une analyse distributionnelle des pauses pour les raisons suivantes :

- L'étude de la distribution des pauses n'est réellement pertinente que pour la parole spontanée. Or, notre corpus comporte un seul style relevant de ce type

d'encodage. Dans la perspective d'une étude contrastive entre les types de discours, ce critère n'est donc d'aucune utilité.

– Nous avons montré lors d'une étude précédente sur des extraits de parole spontanée (Astésano, 1993) d'une part que le pourcentage de pauses non grammaticales est relativement faible (de l'ordre de 25%, alors que Goldman-Eisler, 1968, proposait le chiffre de 45 %), mais également que la distinction entre pauses grammaticales et non grammaticales n'est pas pertinente sur la base de leurs durées intrinsèques (tests statistiques non significatifs).

Nous procédons néanmoins à une étude de la fréquence des différents types de pauses, ainsi qu'à l'étude de leurs durées intrinsèques (cf. § « Nature et durée des pauses », p. 277).

3.2.8.2. Repérage auditif des catégories accentuelles

Le but de notre étude est de distinguer les catégories accentuelles définies dans un cadre métrique sur la base des paramètres acoustico-phonétiques tels que la durée et la f0. C'est donc dans une démarche descendante, hypothético-déductive que nous nous situons. Le caractère proprement hypothético-déductif de notre démarche nous amène à rechercher les propriétés acoustico-phonétiques d'unités au préalable repérées dans le signal. La démarche inverse, qui consisterait à définir des catégories accentuelles sur la base d'observations sur le signal, nous paraît circulaire et dangereuse en ce sens que les différents niveaux d'analyse (acoustique, phonétique, phonologique et fonctionnel) se trouveraient mêlés ; l'aller-retour entre les différents niveaux, directement lié au concept de l'autonomie des niveaux d'analyse, ne serait alors pas possible.

Nous proposons de procéder à un repérage auditif des proéminences perçues dans le message. En effet, l'essence même du rythme est selon nous subjective. C'est le lien entre "l'expérience" que l'on a du rythme et sa réalité acoustico-phonétique qu'il nous intéressait de mettre en évidence.

Le nombre d'écoutes n'est pas limité. Nous procédons tout d'abord à un repérage global des syllabes proéminentes, que nous entourons sur la transcription orthographique des différents types de corpus. Lors de cette même étape, nous notons les frontières majeures que nous percevons, i.e. les frontières d'unités intonatives non terminales (notées /) et terminales (notées //). Une deuxième étape nous permet de repérer les accents à caractère emphatique (accent emphatique initial de mot ou non). Sur la base de la distribution des proéminences (initiales ou finales de mots ou de groupes), et de leur association avec une frontière majeure ou non, nous procédons dans une dernière étape à la caractérisation et à l'étiquetage des catégories accentuelles. Cette transcription prosodique s'effectue tout d'abord sur la transcription orthographique du corpus, puis dans un deuxième temps, cette information est reportée manuellement sur le signal de parole (concaténation syllabique des phonèmes) afin de pouvoir mettre en relation les phonèmes, leurs temps et durée, et leurs caractéristiques prosodiques.

A la suite de Di Cristo & Hirst (1997), et Di Cristo (sous presse), nous classons les syllabes proéminentes en deux macro-classes accentuelles majeures et six catégories:

– 1. la classe des accents rythmiques, comprenant l'accent final de syntagme (AFS) et l'accent rythmique initial de mot ou de groupe (AI) ;

- 2. la classe sémantico-pragmatique, comprenant l'accent emphatique initial (EMP), et l'accent final d'unité intonative terminale (UI-T) ou non terminale (UI-NT).
- Les syllabes non proéminentes sont classées 'inaccentuées' (INA).

Cette typologie donne lieu à une distinction des accents à la fois distributionnelle (initiaux vs. finaux) et fonctionnelle (rythmiques vs. sémantico-pragmatiques), un croisement s'opérant entre les deux sub-divisions (voir catégories accentuelles ci-dessus). La Figure 15 (p. 157) présente les catégories accentuelles définies par le modèle retenu, selon leur distribution et leur fonction.

Précisons toutefois que cette catégorisation ne prend pas en compte le degré de hiérarchisation prosodique. Par exemple, si nous prenons le cas de l'accent final nucléaire, nous n'avons pas distingué les accents nucléaires marquant les fins de paragraphes ou de topiques, des accents nucléaires situés à un degré moindre de frontière (énoncés non final de paragraphe, par exemple). Cette distinction n'aurait pu s'opérer que sur la base d'une analyse conjointe de la structure discursive des textes analysés. Ceci constitue par conséquent une limite de notre méthodologie, et pourrait en partie expliquer une part de la variabilité prosodique observée. Cependant, comme nous l'avons précisé p. 133, l'unité d'analyse linguistique que nous avons choisie (le Module) constitue une entité linguistique assimilable au paragraphe textuel et au paraton prosodique. Par conséquent, on peut considérer que les degrés de frontière prosodique sont assez homogènes à l'intérieur d'un même module; notre limite méthodologique s'en trouve par là même réduite.

		Fonction	
		<i>Rythmique</i>	<i>Sémantico-Pragmatique</i>
Distribution	<i>Initiale</i>	AI	EMP
	<i>Finale</i>	AFS	UI-NT UI-T

Figure 15: Distribution et fonctions des catégories accentuelles du modèle (Di Cristo & Hirst, 1997 ; Di Cristo, sous presse)

3.2.8.2.1. Légimité de l'approche perceptive par experts

Le repérage des frontières prosodiques, associées ou non à des pauses subjectives et/ou objectives, ainsi que le repérage des accents, est une étape délicate de la transcription. En effet, cet exercice est dans une certaine mesure dépendant du transcripateur. Nous avons au préalable effectué un repérage purement auditif des événements prosodiques auprès de trois auditeurs 'experts' (phonéticiens) sur 3 des 6 locuteurs de notre corpus, dans le cadre d'une

étude préliminaire (Astésano et al., 1995b). Après avoir effectué un vote par la médiane, nous nous sommes aperçue que le jugement des trois experts était plutôt cohérent (le plus souvent unanime) pour le repérage des accents (emphatique/non emphatique) et notamment pour le repérage des frontières prosodiques majeures (frontières d'unités intonatives non terminales, et frontières d'unités intonatives terminales). Il est intéressant de remarquer que le repérage des frontières prosodiques majeures s'est montré très cohérent à travers les différents transcripteurs, alors que la perception de ces phénomènes prosodiques est très subjective, et repose sur des critères variés comme l'allongement syllabique, les discontinuités mélodiques et parfois également les pauses silencieuses (respiratoires ou vides). De plus, ce critère de frontière, qui sous-tend le concept d'«unités» et de constituance, est de plus en plus remis en cause par les détracteurs de l'analyse métrique et les tenants de l'analyse rythmique (Guaitella, 1996; Wichmann & Knowles, 1995). Cependant, les trois experts ont obtenu les plus gros scores d'accord sur ces événements prosodiques. Ce constat nous amène à une réflexion concernant la perception des accents : le français a été décrit comme une langue de frontières par opposition aux langues accentuelles (Beckman, 1992 ; Vaissière, 1990). Ceci peut expliquer la plus grande cohérence de perception des accents de frontière entre les auditeurs, spécialement de frontières majeures, les accents de frontières mineures et les accents initiaux rythmiques notamment étant moins prégnants perceptivement. En effet, cette analyse perceptive a révélé une certaine variabilité entre les auditeurs, au niveau notamment de la distinction entre certains accents initiaux rythmiques et emphatiques, ainsi que de l'étendue du domaine accentuel emphatique. On peut émettre l'hypothèse que la distinction perceptive entre accents finaux et initiaux est catégorielle, alors que la perception des accents de distribution identique (initiale ou finale), est plutôt continue, et ce malgré leurs caractéristiques fonctionnelles différentes (accent initial rythmique vs. initial emphatique sémantico-pragmatique par exemple, ou accent final rythmique vs. accents finaux d'unités intonatives sémantico-pragmatiques). En règle générale cependant, il a été plus difficile de trouver un consensus dans la perception des accents rythmiques (AFS et AI), que dans la perception des accents à la frontière d'unités intonatives. On peut expliquer cette tendance par le marquage relativement faible des accents rythmiques au niveau acoustique. Il est frappant en revanche de trouver un fort consensus dans la perception des accents associés à une frontière d'unité intonative, eu égard à la variabilité trouvée dans la réalisation de ces accents (Vaissière, 1983 ; Astésano et al., 1997b et 1998) et aux difficultés de définition formelle de ces constituants prosodiques (cf. discussion § « L'accent nucléaire », p. 69). Ceci explique sans doute pourquoi des auteurs comme Fónagy (1980) ont montré une grande dispersion dans la perception de l'accent en français par des groupes d'auditeurs.

Le but de notre étude n'étant pas de déterminer la réalité perceptive des catégories accentuelles du français d'un point de vue psychoacoustique, nous n'avons pas fait appel à un groupe d'auditeurs pour réaliser un test de perception. Sur la base du test perceptif préliminaire, où nous avons observé une grande cohérence entre les 3 auditeurs experts, nous avons pris le parti de procéder au repérage auditif des proéminences pour le reste du corpus par un seul expert (l'auteur). Les locuteurs utilisent l'organisation prosodique afin de transmettre un message : les différentes catégories accentuelles qu'ils utilisent pour cela ont un usage linguistique particulier. C'est ce lien entre fonctions linguistiques et perception par l'auditeur que nous désirons mettre en valeur. Sur la base des catégories accentuelles fonctionnelles prédéfinies, nous voulons montrer que ces fonctions

linguistiques diverses sont reflétées au niveau acoustico-phonétique par des stratégies particulières en termes de gestion de la durée et de la f_0 .

Ce n'est qu'après le repérage auditif que les extraits ont été analysés au niveau de la réalisation des paramètres prosodiques. Aucune correction de l'analyse perceptive n'a été effectuée a posteriori sur la base de l'analyse acoustique.

4. RESUME

Nous proposons de caractériser le système accentuel du français sur le plan acoustico-phonétique. Nous choisissons pour cela le modèle de Di Cristo & Hirst qui propose une typologie fonctionnelle et une interprétation formelle du système accentuel du français dans le cadre de l'approche métrique. Une originalité de notre travail consiste à proposer un cadre d'analyse identique à différents types de discours, comprenant notamment un type de parole spontanée qui reste généralement l'apanage des descriptions issues du courant comportementaliste.

Notre corpus comprend 3 types de discours situés sur une échelle de spontanéité : *Lecture*, *Journal Radiophonique* et *Interview*. Un repérage auditif des proéminences a été effectué. Les proéminences ont ensuite été catégorisées en deux macro-classes et 6 catégories accentuelles :

- Accents rythmiques, comprenant l'accent rythmique final de syntagme (AFS) et l'accent rythmique (secondaire) initial (AI) ;
- Accents sémantico-pragmatiques, comprenant l'accent emphatique initial (EMP) et les accents d'unité intonative terminale et non terminale (UI-T et UI-NT).

Cette typologie donne lieu à une distinction à la fois *distributionnelle* (initial vs. final) et *fonctionnelle* (rythmique vs. sémantico-pragmatique) des proéminences accentuelles.

Notre corpus comprend 10 minutes de parole continue réparties sur les 3 types de discours (2 locuteurs par type de discours). L'étiquetage phonémique et les différentes étapes d'extraction des données ont été effectués manuellement. Nous totalisons 790 proéminences, 2613 syllabes et environ 6000 occurrences de phonèmes. Si notre corpus n'est pas contrôlé et par conséquent non phonétiquement équilibré, il est en revanche statistiquement équilibré. De même, nos valeurs de durées phonémiques vont dans le sens des études réalisées sur des corpus de parole contrôlée, indiquant par là même que notre corpus est suffisamment conséquent et robuste pour nous permettre de dégager des tendances à l'issue de nos analyses expérimentales.

Nous proposons deux axes de recherche :

- une analyse de l'organisation paradigmatique de l'accentuation (durée différentielle des constituants syllabiques, et aspects dynamiques des variations locales de f0) ;
- une analyse de l'organisation temporelle de la parole à plusieurs niveaux d'actualisation du rythme.

- 2^{IEME} PARTIE -

Analyses Expérimentales

<i>Chapitre 5 - Durées segmentales -</i>	<u>165</u>
1. Problématique des durées segmentales : Introduction	<u>167</u>
2. État des connaissances sur la durée segmentale en français	<u>169</u>
3. Tendances d'allongement différentiel des segments	<u>172</u>
4. Modélisation des durées segmentales dans un cadre syllabique : le modèle de Campbell	<u>173</u>
5. Analyses Empiriques	<u>181</u>
6. Discussion et conclusion générale sur les tendances d'allongement différentiel en fonction des contraintes prosodiques et des styles de parole	<u>225</u>
7. Résumé	<u>233</u>
<i>Chapitre 6 - Configurations Tonales -</i>	<u>235</u>
1. Introduction	<u>237</u>
2. Caractérisation des configurations tonales des proéminences en français	<u>240</u>
3. Conclusion	<u>255</u>
4. Résumé	<u>257</u>
<i>Chapitre 7 - L'organisation temporelle de la parole -</i>	<u>259</u>
1. Introduction	<u>261</u>
2. Plan expérimental	<u>263</u>
3. Résultats	<u>268</u>
4. Discussion et Conclusion	<u>283</u>
5. Résumé	<u>287</u>

- CHAPITRE 5 -

Durées Segmentales

1. PROBLEMATIQUE DES DUREES SEGMENTALES : INTRODUCTION

L'un des problèmes majeurs concernant le développement des systèmes de synthèse de la parole à partir du texte est l'implémentation de valeurs de durées segmentales adéquates (pour un état de la question sur ce sujet, voir Campbell, sous presse). Un nombre considérable de facteurs influent sur les durées segmentales (variabilité des durées internes aux segments, variabilité induite par les contextes phonémique, prosodique, syntaxique, sémantique, pragmatique etc.). Les combinaisons de facteurs entrant en jeu dans la détermination des durées segmentales sont multiples, surtout lorsque l'objet d'étude porte sur la parole spontanée non contrôlée. Le nombre de combinaisons possibles est tel que, pour la reconnaissance automatique de la parole, même un corpus d'entraînement très large ne contiendrait qu'une infime proportion des combinaisons possibles (van Santen, 1997).

Afin d'éviter le problème de la rareté des données segmentales, Campbell (1992) propose de considérer une unité de niveau supérieur - la syllabe - comme unité de prédiction des durées. La durée des segments individuels est ensuite prédite dans le cadre de la syllabe, selon un principe « d'accommodation » (cf. § « Les principes d'élasticité segmentale et d'accommodation segmentale dans le cadre de la syllabe », p. 178), qui prévoit, dans sa version 'forte', que chaque segment dans une syllabe donnée est allongé ou réduit de la même manière (le même coefficient k étant appliqué à tous les segments constitutifs d'une même syllabe). Une procédure d'aller-retour est ensuite appliquée, qui permet d'adapter la durée de la syllabe à l'influence de certains segments individuels (comme la durée différentielle des voyelles devant des occlusives voisées et non voisées).

Campbell révisé par la suite cette hypothèse forte d'accommodation et propose d'assouplir l'implémentation du facteur k , afin de tenir compte notamment des phénomènes d'allongement plus importants des éléments de Noyau et de Coda devant les frontières prosodiques majeures, et des allongements plus marqués des éléments d'Attaque sur les syllabes initiales de mots lexicaux ou d'accents emphatiques.

La procédure de prédiction des durées proposée par Campbell se fait donc en deux étapes :

- calcul de la durée des syllabes à partir du contexte prosodique, en ignorant dans un premier temps le contenu segmental de la syllabe ; cela suppose un principe d'indépendance segmentale, i.e. que la syllabe est indépendante des segments qu'elle contient ;
- calcul des durées des segments en fonction de leur élasticité, et suivant un principe d'accommodation segmentale, qui tient compte dans la deuxième version du modèle de l'allongement différentiel suivant le contexte prosodique ; cela implique un principe de médiation syllabique, i.e. que la durée segmentale est soumise à (ou en tout cas dépend de) la durée syllabique.

Van Santen (1997) n'adhère pas à ces principes et montre pour l'anglais américain et le chinois mandarin que la durée des syllabes dépend fortement des durées intrinsèques des segments qui la composent. Le débat sur l'unité de prédiction de durée la plus appropriée reste donc ouvert. Des tests psycholinguistiques seraient sans doute indispensables à la

détermination d'une unité.

Des résultats d'allongement différentiels semblables à ceux de Campbell pour l'anglais ont été trouvés pour le français (Fant et al., 1991a et b ; Astésano et al., 1995b). Ainsi, les syllabes accentuées situées à des frontières prosodiques majeures (unités intonatives) présentent un allongement plus important des éléments de Noyau et de Coda, la Coda étant particulièrement allongée à la frontière d'unités intonatives terminales.

Cependant, le choix d'une unité prosodique de prédiction des durées ne s'impose pas de lui-même. Ainsi, Campbell (1992) note que les segments de Coda ont une stabilité relative plus importante que les segments d'Attaque, et émet l'hypothèse qu'une unité intermédiaire (la Rime) pourrait être plus appropriée pour rendre compte des phénomènes d'allongement segmental. D'autres auteurs proposent une unité supérieure à la syllabe, le Groupe Inter P-Center (GIPC), où le centre perceptuel correspond à l'attaque syllabique après une pause, et à l'attaque vocalique dans les autres positions (Barbosa & Bailly, 1994). Cette unité de programmation rythmique permettrait de rendre compte de manière optimale de la variabilité des données analysées (phrases lues isolées).

Se pose alors le problème de l'objet d'étude. En effet, la plupart des études réalisées sur la durée segmentale ont pour objet la parole contrôlée de laboratoire. Or, les recherches sur la parole spontanée en français révèlent une organisation accentuelle plus complexe que ce que les descriptions traditionnelles du français laissent paraître. Ainsi, certains auteurs ont mis en évidence l'existence en français de deux tendances accentuelles complémentaires : une accentuation (mélodique) initiale de mot, comparable à celle de l'anglais, et l'accentuation finale traditionnelle, marquée par l'allongement (Fónagy & Fónagy, 1976 ; Di Cristo & Hirst, 1984 ; Rossi, 1985b ; Vaissière, 1991). Comme nous l'avons argumenté dans le paragraphe « L'organisation accentuelle en français », sur les modèles accentuels (p. 78), il nous semble impératif de tenir compte de ces caractéristiques accentuelles afin d'englober l'ensemble de la réalité rythmique spécifique au français. Il est nécessaire pour cela d'appuyer les recherches sur des modèles adaptés aux différents types de parole en situation. Le modèle de Di Cristo & Hirst (1997), et plus particulièrement sa version actuelle développée par Di Cristo (sous presse), nous semble correspondre à cette attente.

Nous proposons donc de rendre compte des phénomènes de durée en français en fonction du contexte prosodique et des styles de parole.

Avant d'exposer les résultats de nos analyses empiriques, nous proposons un exposé des recherches sur la durée, avec une attention particulière portée aux recherches sur le français.

Nous présentons ensuite les principales tendances concernant l'allongement des segments (allongement différentiel, fonction du contexte prosodique), trouvées dans les langues comme l'anglais, le suédois, le hollandais et le français (Fant et al., 1991a et b ; Campbell, 1992 et suivantes ; Hofhuis, 1993 ; Astésano et al., 1995b).

Enfin, nous présentons le modèle de Campbell, dont notre étude est largement inspirée. Ce modèle est fondé sur l'analyse de l'allongement différentiel des segments dans un cadre syllabique. Nous présentons donc la méthode de normalisation des durées segmentales proposée par Campbell, qui permet de rendre compte de l'allongement différentiel des segments en fonction de leur élasticité propre.

2. ÉTAT DES CONNAISSANCES SUR LA DURÉE SEGMENTALE EN FRANÇAIS

2.1. Diversité des études

De nombreux phénomènes affectent la durée de la parole, comme l'accentuation, les frontières de mots, les unités syntaxiques, le contexte phonétique, le nombre de syllabes dans un groupe accentuel etc. A cela, il faut ajouter l'importante variabilité intra-locuteur : en effet, la durée de chaque phonème peut varier considérablement d'une répétition à l'autre d'une même phrase par un même locuteur. (O'Shaughnessy, 1984).

La tâche de modélisation de la durée segmentale n'est donc pas aisée, compte tenu de la variabilité rencontrée à différents niveaux. Cette difficulté se reflète dans la diversité des études entreprises dans le domaine de la durée segmentale en français, tant du point de vue des corpus (en taille et nombre de locuteurs) que du point de vue des objets d'étude.

Les motivations des chercheurs sont à l'origine variées : descriptions phonétiques et phonologiques comparatives (Delattre, 1966b, 1981 ; Boudreault, 1968) ou non (Benguerel, 1971 ; O'Shaughnessy, 1981, 1984 ; Di Cristo, 1985) ; applications à la reconnaissance (Bartkova & Sorin, 1987 ; Bartkova, 1988) ou à la synthèse vocale (Aubergé, 1991 ; Barbosa & Bailly, 1994).

De plus, les corpus vont de la *parole 'naturelle'* (Delattre, 1966b), en passant par la lecture de *phrases porteuses* ou de *phrases* (Boudreault, 1968 ; Benguerel, 1971 ; O'Shaughnessy, 1981 ; Di Cristo, 1985), ou bien de *paragraphes* (O'Shaughnessy, 1984), ou encore par la *lecture de nombres* (Bartkova, 1988), jusqu'aux *logatomes* (Bartkova & Sorin, 1987 ; Barbosa & Bailly, 1994).

L'objet des études est également différent :

- il s'agit des *phonèmes* pour Boudreault (1968), O'Shaughnessy (1984), Bartkova & Sorin (1987), Bartkova (1988), Aubergé (1991) et Barbosa & Bailly (1994) ;
- des *syllabes* pour Delattre (1966b) et Benguerel (1971) ;
- des *voyelles*, ou du moins une partie du système vocalique, pour Delattre (1981), O'Shaughnessy (1981) et Di Cristo (1985) ;
- et enfin des *consonnes* pour O'Shaughnessy (1981).

Certaines de ces études ont validé pour le français la distinction des phonèmes en macro-classes articulatoires et acoustiques sur la base des durées segmentales. Di Cristo (1985) a ainsi observé une distinction significative entre les voyelles hautes, moyennes et basses, alors que Bartkova (1988) a comparé les durées des macro-classes consonantiques traditionnelles (classées selon le mode d'articulation et de phonation).

Les unités segmentales ainsi définies sont analysées en fonction de divers contextes :

- contextes *prosodiques* et *distributionnels* (Delattre, 1966b ; Boudreault, 1968 ; Benguerel, 1971 ; O'Shaughnessy, 1981 ; Di Cristo, 1985 ; Bartkova & Sorin, 1987 ; Bartkova, 1988, Aubergé, 1991 et Barbosa & Bailly, 1994) ;

– contextes *phonétiques* (O’Shaughnessy, 1984), et le plus souvent une combinaison des deux.

Enfin, les types de mesures effectuées finissent de complexifier davantage le paysage de nos connaissances sur la durée segmentale en français :

– nous disposons ainsi de *moyennes* seules (Delattre, 1966b, 1981 ; Boudreault, 1968 ; Bartkova & Sorin, 1987, et Aubergé, 1991) ;

– de *ratio* et de *pourcentage* d’allongement ou de raccourcissement (Benguerel, 1971 ; O’Shaughnessy, 1981) ;

– et enfin des *moyennes* et *écarts types*, autrement dit des ‘caractéristiques d’élasticité’ des unités segmentales (O’Shaughnessy, 1984 ; Di Cristo, 1985 ; Bartkova, 1988 ; Barbosa & Bailly, 1994).

Malgré la grande richesse que représentent ces études, la diversité des objets d’études et des critères de mesures conduit à une trop grande variété et rareté des résultats. Le besoin d’une base de données de référence des durées segmentales en français, comme il en existe pour l’anglais, se fait toujours sentir.

2.2. Caractéristiques des durées segmentales en français

Malgré leur hétérogénéité, les diverses études ont permis de cerner davantage les caractéristiques des durées segmentales en français.

La durée se caractérise par son élasticité, sa variabilité, ce qui rend difficile son utilisation dans un algorithme de reconnaissance de la parole, a fortiori un algorithme multi-locuteurs (Bartkova, 1988).

On distingue essentiellement trois types de variabilité des durées segmentales :

– la variabilité induite par les phénomènes articulatoires (durées intrinsèques des phonèmes)

– la variabilité induite par l’influence du contexte phonétique (durées co-intrinsèques)

– la variabilité induite par le contexte prosodique.

Il convient de rajouter à cette liste la variabilité d’origine paralinguistique.

2.2.1. Variabilité induite par les événements articulatoires (durées intrinsèques)

Une des particularités du français est que la durée n’a pas de rôle phonologique distinctif. Les différences de durée vocalique notamment (ex : maître/mettre) n’ont plus, de nos jours, de valeur distinctive en français. Les rares oppositions de paires minimales qui demeurent se réalisent par une différence de timbre, celle-ci étant généralement accompagnée d’une différence de durée vocalique (ex : cote/côte ; pomme/paume). Ces variations distinctives sont très marginales (dialectales ou désuètes) et instables. En revanche, la variabilité induite par le contexte phonétique, de caractère involontaire, est très stable (Delattre, 1966a).

La variation intrinsèque la plus frappante en français est la différence de durée entre les voyelles nasales et leurs voyelles orales correspondantes, et ce quel que soit le contexte prosodique. Cette différence contribue à l'identification des voyelles nasales. Plus la voyelle orale est intrinsèquement longue, moins la différence de durée avec la voyelle nasale correspondante sera marquée (Delattre, 1962 ;1981).

La durée des voyelles en français varie donc en fonction du trait de nasalité, mais également en fonction de la hauteur (O'Shaughnessy, 1981). Di Cristo (1985) constate ainsi que les voyelles hautes sont intrinsèquement plus courtes que les voyelles basses.

Pour les consonnes, le mode d'articulation a une influence sur la durée des segments : les constrictives ont une durée intrinsèque plus importante que les occlusives. De même, le mode de phonation induit une variation de durée, les segments non voisés étant sensiblement plus longs que les segments voisés.

2.2.2. Variabilité induite par le contexte phonétique (durées co-intrinsèques)

Lors de la phase de phonation, les phonèmes s'influencent les uns les autres (coarticulation). Plus précisément, la durée vocalique est influencée par le mode d'articulation des consonnes adjacentes et le voisement de la consonne subséquente (O'Shaughnessy, 1981). Le mode d'articulation exerce une influence, les voyelles étant généralement plus longues devant les constrictives que les occlusives. Le phonème [R], souvent considéré comme une fricative sur le plan phonétique, a le même effet allongeant que les constrictives. Il est d'ailleurs probable que la durée importante de la voyelle qui précède le [R] (voisé par assimilation) soit un indice de présence du [R] en finale de syllabe, et qu'elle permette le contraste avec une syllabe ouverte, car les indices acoustiques du [R] sont souvent trop subtils (O'Shaughnessy, 1981).

La durée de la voyelle est également proportionnelle au recul du lieu d'articulation de la consonne subséquente, et ceci aussi bien pour les voyelles longues que pour les voyelles brèves (Di Cristo, 1985). Mais le facteur qui influence le plus la durée vocalique est le mode phonatoire de la consonne subséquente : les consonnes voisées ont une influence allongeante sur les voyelles. Enfin, les variations co-intrinsèques de la durée vocalique ne dépendent pas uniquement des consonnes subséquentes : elles sont aussi conditionnées par les consonnes qui précèdent, le caractère voisé de la consonne étant un facteur significatif (Di Cristo, 1985).

A ces variations co-intrinsèques de durée, fondées sur les caractéristiques articulatoires des segments, Vaissière (1977) ajoute un facteur lié à la place du phonème dans la syllabe (cf. § « Tendances d'allongement différentiel des segments », p. 172) et le nombre de phonèmes dans la syllabe. Il y aurait en effet un phénomène de compensation de durée à l'intérieur de la syllabe : la voyelle a tendance à être plus courte lorsqu'elle est suivie d'un groupe consonantique, et cela peut être expliqué par la tendance à l'isochronisme syllabique qui veut que des syllabes de structures différentes aient une durée approximativement identique (Bartkova, 1988). Di Cristo (1985) remarque les mêmes tendances de compensation de durée à l'intérieur de la syllabe (abrègement ou allongement de la durée vocalique suivant le contexte consonantique) mais il s'interroge sur la validité de l'explication (tendance à l'isochronie syllabique); en effet, les voyelles devant les constrictives sourdes sont allongées, alors que ces dernières sont déjà nettement plus

longues que les occlusives. Les processus de compensation des durées segmentales à l'intérieur de la syllabe constituent cependant une preuve supplémentaire de l'existence de la syllabe en tant qu'unité de pré-programmation (Massaro, 1975).

Pour ce qui est de l'élasticité des segments enfin, Bartkova (1988) observe une durée moyenne plus longue et un écart type plus grand pour les voyelles nasales que pour les voyelles orales, ces dernières étant par conséquent plus homogènes. La durée vocalique est plus longue et l'écart type plus important lorsque le contexte droit est une consonne allongante. L'auteur en déduit donc que la valeur de l'écart type est corrélée avec la valeur de la moyenne.

2.2.3. Variabilité de durée induite par le contexte prosodique

La variation de durée segmentale est inversement corrélée avec la croissance du débit de parole, la longueur des mots et la fréquence d'occurrence du mot dans le discours. Bartkova (1988) remarque cependant que les consonnes sont plus robustes aux contextes prosodiques : elles ont des durées plus homogènes que les voyelles et sont moins sensibles notamment à la longueur du mot.

Le caractère accentué ou inaccentué influence également la durée des phonèmes : ainsi, les voyelles seront nettement plus longues et auront un écart type plus important en position accentuée qu'en position non accentuée. La durée des consonnes est systématiquement plus longue dans une position prépausale et cela d'une façon plus évidente pour les constrictives non voisées que pour les autres types de consonnes (Bartkova, 1988).

Benguerel (1971) remarque cependant que les variations co-intrinsèques (notamment le caractère voisé/non voisé de la consonne subséquente) sont un facteur moins important que les variations induites par l'intonation et la structure syllabique, dans l'accentuation non emphatique en français.

3. TENDANCES D'ALLONGEMENT DIFFERENTIEL DES SEGMENTS

L'information de durée peut aider à la reconnaissance automatique des groupements prosodiques. A cette fin, Campbell (1993) utilise les informations de durées différentielles dans le cadre syllabique. Il remarque que les segments dans les mots et les syntagmes proéminents ou emphatiques sont allongés de manière significative comme le sont ceux qui précèdent une frontière prosodique de syntagme. Cependant, les différences d'allongement des segments d'Attaque et de Coda permettent de distinguer les deux contextes d'allongement, et par conséquent de différencier les effets de proéminence (allongement sur le début de la syllabe), des effets de frontière (allongement sur la fin de la syllabe). Le degré d'allongement des composantes syllabiques observé à l'intérieur de la syllabe permettrait donc de révéler la structure d'un énoncé pour l'analyse automatique de la prosodie. Pour Campbell (sous presse), les variations de durée sont donc des indices majeurs, qui aident à l'interprétation du sens de l'énoncé par un marquage spécifique des frontières et des proéminences.

Klatt (1976) remarque que l'augmentation des unités phoniques (voyelles et consonnes) situées en fin de propositions ou de phrases en anglais et en français n'est pas un phénomène universel, donc non physiologique; par conséquent, il s'agirait d'un

apprentissage linguistique (cf. § « Cas particulier : l'allongement final, un phénomène universel », p. 58, pour une discussion). Cette variation de durée aurait un rôle prééminent dans l'identification des structures syntaxiques. Klatt se demande si le locuteur apprend à réaliser cet allongement des segments terminaux afin d'aider l'auditeur à décoder le message ou si c'est une tendance naturelle à opérer un ralentissement au terme des séquences motrices et des unités de programmation. Cette tendance pourrait servir à préparer la planification de l'unité syntaxique suivante (Di Cristo, 1985).

Plus spécifiquement, Campbell (1992, 1993, sous presse.) et Campbell & Isard (1991) observent que les segments en position finale d'énoncé subissent un allongement plus important du Noyau et de la Coda, alors que l'allongement des accents lexicaux en anglais ('stress') affecte plutôt les segments d'Attaque. Les segments accentués à l'intérieur des énoncés, pour des raisons rythmiques et accentuelles, sont allongés de manière plus uniforme à travers la syllabe. Les auteurs concluent que l'allongement différentiel permet de distinguer clairement entre les allongements syllabiques prépauxaux, qui, confinés à la Rime, sont considérés comme des phénomènes de frontières, et les allongements observés dans les syllabes internes d'énoncés, qui semblent être une propriété de la syllabe entière.

Il semblerait que ces caractéristiques de distribution différentielle de l'allongement sur les segments, en fonction de la place qu'ils occupent dans l'énoncé, soient un phénomène répandu dans diverses langues. Nombreuses sont en effet les données sur l'anglais (Beckman, 1992 ; Campbell, 1991 et suivantes), sur le hollandais (Hofhuis, 1993 ; Hofhuis et al., 1995 ; Cambier-Langeveld, 1997), sur le français (Astésano et al., 1995b ; Hirst et al., 1998 ; Astésano, 1998b), qui confirment ces tendances.

4. MODELISATION DES DUREES SEGMENTALES DANS UN CADRE SYLLABIQUE : LE MODELE DE CAMPBELL

4.1. Quelle unité phonologique pour rendre compte des phénomènes temporels ?

Un des débats en phonologie concerne la nature des éléments accentuables : s'agit-il des voyelles (Noyaux), des Rimes ou des syllabes ? Hayes (1989) et Selkirk (1984) proposent la syllabe, tandis que Halle & Vergnaud (1987) privilégient les noyaux syllabiques. Or, comme nous venons de le voir, de nombreuses études portant sur des langues différentes montrent que les consonnes initiales et finales sont également affectées par l'accent, ce qui tend à prouver que l'unité accentuable est plus grande que le Noyau ou la Rime. Il s'agirait fort probablement de la syllabe, ou bien, comme le suggèrent certaines expériences sur l'étendue du domaine accentuel en anglais, du pied métrique (Turk et al., 1997).

La question de l'unité la plus à même de rendre compte des divers phénomènes temporels se pose également en synthèse vocale. Quelle unité doit-on modéliser : le segment, le couple Attaque/Rime, la syllabe, le mot ... ? Il semblerait intuitivement plus plausible que le contrôle temporel et articulatoire se situe au niveau suprasegmental. En effet, des études sur l'intelligibilité et le naturel de la synthèse vocale ont montré la nécessité de prendre en compte des informations d'ordre rythmique (comme la régularité perceptive des pieds métriques) pour améliorer la qualité des systèmes. Pourtant, la plupart des systèmes de synthèse effectuent une programmation au niveau segmental et ne prennent

en compte que les phénomènes phonétiques de durées intrinsèques et co-intrinsèques, ainsi que la place du phonème dans le syntagme (Klatt, 1976 ; Bartkova & Sorin, 1987 pour le français).

4.1.1. La syllabe, interface rythmique

Campbell (1992 et suivantes) propose quant à lui de prendre en compte le niveau de la syllabe dans son modèle, en tant qu'interface entre le niveau phonétique et le rythme. Les structures de niveau supérieur détermineraient donc la durée des segments individuels. Un principe fondamental de cette hypothèse serait que les consonnes et les voyelles ont des caractéristiques d'allongement similaires si on les considère en termes de variance. Campbell tente de rendre compte par là de la manière dont les segments dans une syllabe partagent la place disponible en fonction de leur propre élasticité. Des règles de combinaison et de compensation des segments composant l'unité ainsi déterminée s'appliqueraient pour que l'addition de durée des différents segments atteignent la taille de l'unité de niveau supérieur. Les processus d'organisation temporelle semblent opérer à trois niveaux différents :

- au niveau de l'énoncé (phénomènes de frontières ; effets de l'accentuation et du rythme dans le cas notamment de la focalisation)
- au niveau syllabique (effets de l'accentuation et du rythme)
- au niveau du segment (ajustement des durées en fonction des contraintes contextuelles phonémiques).

4.1.2. Avantages du cadre syllabique

Campbell (1992 et suivantes) propose un modèle de l'organisation temporelle de la parole à deux niveaux :

- le niveau syllabique : les durées des syllabes sont tout d'abord calculées de façon à refléter l'organisation rythmique et structurale de l'énoncé ;
- le niveau des segments : les durées des segments sont calculées dans un deuxième temps, en s'accommodant au niveau immédiatement supérieur en fonction de leur élasticité propre.

La prééminence de la syllabe sur le segment confère à ce modèle deux avantages majeurs :

D'une part, si une erreur de prédiction de durée intervenait au niveau du segment, elle ne se répercuterait que sur une seule syllabe et le rythme en serait d'autant moins perturbé. Par l'effet de compensation de durée des segments à l'intérieur de l'unité syllabique, une erreur d'allongement excessif d'un segment par exemple serait compensée par le raccourcissement de la durée du phonème suivant dans la même unité rythmique. L'équilibre inter-syllabique est ainsi maintenu, et par conséquent le niveau rythmique immédiatement supérieur (le groupe accentuel ou pied métrique dans le sens d'Abercrombie, 1964) n'est pas perturbé.

D'autre part, le niveau syllabique présente une variabilité inter-locuteurs moins importante que le niveau segmental. Lorsqu'il compare l'organisation temporelle de

différents locuteurs, Campbell (sous presse) constate une importante variabilité dans la durée des phonèmes. Cette variabilité est considérablement réduite si l'on compare les durées au niveau syllabique. Les syllabes seraient donc des unités de prédiction beaucoup plus robustes que les segments et constitueraient en quelque sorte une unité de 'normalisation' inter-locuteurs.

Il semble donc que la syllabe soit l'unité de planification et de performance des durées segmentales.

4.2. Normalisation des durées segmentales

L'une des originalités du modèle de Campbell est la quête d'une méthode de calcul des durées permettant de faire la différence entre la variabilité temporelle déterminée linguistiquement, de celle due aux contraintes articulatoires de production des différents sons. A cette fin, il propose une méthode de normalisation des durées segmentales permettant de rendre compte des phénomènes d'allongement ou de raccourcissement des phonèmes, en fonction de leur élasticité (dispersion autour de la moyenne).

4.2.1. Pourquoi normaliser ?

Un argument décisif est lié à la notion de 'variabilité'. Rossi (1985a) distingue 2 types de variabilité :

- la variabilité aléatoire concerne la nature biologique de la parole. Les caractéristiques physiologiques de l'appareil phonatoire (masse et inertie des organes phonatoires) entraînent une grande part de variabilité incontrôlable, assimilable à du bruit. Ce type de variabilité est contingent et non pertinent sur le plan de la communication.
- la variabilité structurelle, dans laquelle on distingue :
 - ◇ la variabilité contextuelle, induite par les phénomènes de coarticulation. Elle peut être contingente ou pertinente, et seuls les phénomènes de coarticulation relevant du deuxième cas affleurent à la conscience.
 - ◇ la variabilité interindividuelle (liée aux différences physiologiques entre les locuteurs) qui, dès lors qu'elle dépasse les limites de variabilité automatique aléatoire et contextuelle, devient potentiellement significative.
 - ◇ la variabilité linguistique, enfin, qui concerne les variations sociolectales, idiolectales, dialectales etc. constitutives d'un système linguistique.

Toutes les formes de variabilité que Rossi distingue (mise à part la variabilité aléatoire) servent la communication, et ne doivent pas être réduites quantitativement, notamment dans le cadre de l'application à l'identification du locuteur et à la reconnaissance de la parole.

Néanmoins, selon le point de vue que l'on adopte, le niveau d'abstraction auquel on se situe, et les objectifs que l'on veut atteindre, on peut envisager de contrôler quantitativement une part de variabilité. Lorsque l'objectif de l'étude est de faire la part de l'allongement segmental induit par le contexte prosodique, il est bon de contrôler la variabilité de durée intrinsèque et co-intrinsèque. C'est précisément ce que permet le

modèle de Campbell : en ramenant l'élasticité propre de chaque type de segment (durée intrinsèque) à une valeur centrée et réduite, on s'affranchit de la variabilité intrinsèque des segments. Le principe d'ajustement des segments à l'intérieur de la syllabe permet d'autre part de s'affranchir des effets de coarticulation. De fait, la variabilité observée à la sortie de la normalisation sera imputable au contexte prosodique ; la comparaison entre les locuteurs ou les styles de parole permet ensuite de distinguer entre la variabilité constitutive du système noyau de la langue, de la variabilité imputable au locuteur ou au type de discours.

Certains chercheurs sont cependant farouchement opposés à toute forme de normalisation, arguant qu'elle détruit une partie de l'information linguistique.

4.2.2. La normalisation gomme-t-elle les variations de durée pertinentes linguistiquement ?

Nous avons vu que les phonèmes ont des durées intrinsèques variables, mais cependant reproductibles, permettant de les rassembler en macro-classes phonémiques. L'information fournie par la variabilité est donc pertinente linguistiquement. Certains phénomènes accentuels prendraient d'ailleurs appui sur la structure segmentale des syllabes. C'est le cas notamment des accents emphatiques qui semblent frapper préférentiellement les syllabes commençant par des consonnes sourdes, intrinsèquement longues et prégnantes sur le plan perceptif (Pasdeloup, 1990). Les locuteurs utiliseraient donc les caractéristiques de durées intrinsèques et les propriétés articulatoires à des fins linguistiques. Les détracteurs des procédés de normalisation évoquent cette tendance pour argumenter leur position.

Il nous semble cependant qu'un malentendu est à l'origine de cette discorde. En effet, les descriptions quantitatives et qualitatives des phénomènes temporels ne sont en aucun cas mutuellement exclusives. La normalisation en Z (cf. § « Normalisation par transformée Z », p. 177) permet de faire apparaître les allongements ou raccourcissements liés au contexte linguistique (accentuation, frontière, emphase etc.) en éliminant les allongements relatifs, liés à la durée intrinsèque de certains phonèmes. Elle permet donc, dans une démarche descriptive, de rendre compte précisément des phénomènes proprement rythmiques, en faisant apparaître les phénomènes d'allongement uniquement imputables aux effets de l'accentuation. Mais elle n'empêche pas de procéder à des comparaisons des structures syllabiques en fonction du contexte accentuel. Rien n'exclut en effet, dans une démarche de modélisation (ex : synthèse de la parole), de prendre en compte certaines caractéristiques phonotactiques, en majorant par exemple le poids des consonnes sourdes initiales, car le modèle de Campbell permet un aller-retour entre les niveaux syllabique et infra-syllabique.

4.2.3. Comment normaliser ?

Il existe essentiellement 3 méthodes de quantification des variations temporelles des sons de la parole : en termes de millisecondes (durées brutes), en rapport ou en pourcentage, ou bien en termes d'étendue de la distribution observée (Campbell, 1992).

La première méthode ne prend pas en compte les différences de durée intrinsèque des divers types de phonèmes, et néglige de prendre en considération le fait, par exemple, qu'une augmentation de durée de 50 ms sera relativement plus importante pour une voyelle dont la durée type est de 50 ms, que pour une autre dont la durée type est de 150 ms ; dans

le premier cas, il y a de fortes chances pour que l'augmentation de durée soit la conséquence d'un phénomène suprasegmental, alors que dans le deuxième cas, l'allongement n'est peut être que le résultat d'un contexte segmental allongeant.

La deuxième méthode, où l'allongement est exprimé en pourcentage, permet une certaine relativisation de la différence observée dans l'exemple ci-dessus, mais elle ne permet toujours pas une comparaison de l'allongement des différents types de phonèmes, à moins que ceux-ci aient une élasticité parfaitement similaire.

La troisième méthode semble être la plus sûre, car elle tient compte de la durée intrinsèque des phonèmes. Il s'agit d'exprimer l'allongement en termes d'écart type par rapport à la moyenne (normalisation en Z). Les durées brutes (ms), une fois normalisées en Z , vont se trouver normalement dans l'étendue de + ou - 3 écarts types. Cette mesure permet non seulement de comparer les phonèmes de différentes catégories entre eux, mais également de les comparer à l'intérieur d'une même catégorie, où une valeur positive de Z représente un allongement et une valeur négative un raccourcissement par rapport à la durée moyenne observée pour chaque item de cette catégorie dans la base de donnée.

4.2.4. Normalisation par transformée Z

Afin de rendre compte des phénomènes d'allongement ou de raccourcissement des phonèmes, en fonction de leur élasticité (dispersion autour de la moyenne), Campbell propose de convertir les durées de chaque individu dans la base de données segmentale en 'forme normale standard', i.e. en soustrayant la moyenne observée pour chaque individu à la valeur de cet individu et en la rapportant à l'écart type de l'individu.

Les moyennes (μ) et les variances (σ^2) sont donc calculées pour chaque phonème, et la durée segmentale individuelle résulte de la normalisation par transformée en Z , dont la formule est :

$$Z_{\text{item}} = (\text{durée}_{\text{item}} \text{ observée} - \mu_{\text{catégorie}}) / \sigma_{\text{catégorie}}$$

Les valeurs ainsi obtenue reflètent l'étendue de la variation en termes d'écart type autour de la moyenne pour chaque segment. Par conséquent, une valeur positive ou une valeur négative de Z représentent respectivement un allongement ou un raccourcissement par rapport à la durée moyenne observée pour toutes les occurrences d'un type de phonème dans la base de données. Les données qui ont une distribution normale tombent avec une probabilité de 0.998 dans une fourchette de plus ou moins 3. Lorsqu'on passe en valeurs normalisées en Z , on peut dès lors comparer les phonèmes entre eux à l'intérieur de chaque classe phonémique, mais également entre des phonèmes de types différents, car cette normalisation rapporte la moyenne de tous les segments à 0 et leur écart type à 1.

La transformée Z permet donc de se dédouaner de la variation de durées intrinsèques des phonèmes, ainsi que de la part de variabilité inter-locuteurs non pertinente linguistiquement. A partir de là, l'impact du contexte prosodique est plus facilement analysable.

« By these transforms we can factor out most of the articulatory-level effects and view lengthening independently from the phonetics of the message. Similarly, for any given database we need no longer be concerned with

speaker characteristics such as differences in overall speaking rate and can thus model more easily the extent to which all sounds are lengthened or shortened by contextual factors. » (Campbell, sous presse : 19).

L'argument majeur en faveur de ce type de normalisation est le suivant : l'analyse de l'allongement que subit un segment peut apporter plus d'information pertinente que la mesure de sa durée (brute).

En résumé, la valeur Z est une quantification de la position de chaque segment dans sa distribution ; elle peut être considérée comme une mesure de son état en fonction de son élasticité. Un segment ayant une grande variance (comme une voyelle nasale par exemple) et qui montre une différence considérable de durée en termes de mesure absolue en millisecondes peut être dans le même état relatif de compression ou d'expansion qu'un segment ayant une variance bien plus petite (comme une occlusive par exemple) qui semble varier moins en termes absolus.

4.2.4.1. Les principes d'élasticité segmentale et d'accommodation segmentale dans le cadre de la syllabe

Campbell & Isard (1991) proposent un algorithme permettant de rendre compte de la durée des segments, dans les rapports qu'ils entretiennent au sein de la syllabe. Pour le calcul de la durée syllabique, ils proposent d'appliquer à chaque segment un coefficient unique k , dans la formule suivante :

$$Durée_{seg} = \mu_{seg} + k\sigma_{seg} \quad \text{où} \quad \begin{array}{l} \mu_{seg} \text{ est la moyenne des durées des} \\ \text{réalisations du segment } seg. \\ \sigma_{seg} \text{ est l'écart type des durées des} \\ \text{réalisations du segment } seg \end{array}$$

Il s'agit de l'hypothèse forte de l'élasticité qui suppose que tous les segments dans une syllabe donnée se trouvent à la même place dans leurs distributions respectives.

Ce coefficient k mesure en fait l'étendue de l'écart type de chaque catégorie de phonèmes en relation avec les autres phonèmes. Ainsi, Campbell (sous presse) donne l'exemple suivant : les durées des phonèmes [a] et [t] à l'intérieur de la syllabe vont s'ajuster en fonction de leur élasticité comme suit :

$$Durée_{[a]+[t]} = \mu_{[a]} + \mu_{[t]} + k(\sigma_{[a]} + \sigma_{[t]})$$

Dans ce cas, [a] aura en sortie une durée x .

Si [a] était combiné au phonème [d], on aurait également :

$$Durée_{[a]+[d]} = \mu_{[a]} + \mu_{[d]} + k(\sigma_{[a]} + \sigma_{[d]})$$

mais dans ce cas là, [a] aurait une durée y , supérieure à la durée x en contexte [t], afin de compenser la durée et l'écart type moindres de [d].

L'élasticité décrit donc la notion d'allongement ou de raccourcissement des durées segmentales intrinsèques, alors que le principe d'accommodation décrit le processus d'allongement ou de raccourcissement que subissent les segments en fonction du contexte

phonémique pour entrer dans le cadre de la syllabe.

Le facteur k est dans un premier temps appliqué de manière égale sur les segments normalisés pour que leur somme atteigne la durée totale de la syllabe : il permet de rendre compte de l'élasticité temporelle des différents types de segments.

Dans une version ultérieure, Campbell (1993) propose une modification de son algorithme et implémente une variable d'affaiblissement ('decay variable') sur les segments des syllabes éloignés des frontières prosodiques. Il s'agit de l'hypothèse faible d'élasticité qui postule que les statistiques doivent être adaptées selon la position de la syllabe dans la phrase (finale vs. non finale), et selon la position des segments à l'intérieur de la syllabe (les segments composant l'Attaque vs. ceux composant la Rime) (cf. § «Tendances d'allongement différentiel des segments», p. 172). Cette version de l'algorithme prend également en compte les contraintes de bas niveau, notamment le contexte phonétique dans le cas des voyelles + occlusives.

4.3. Choix d'une méthode de normalisation adaptée à notre objet d'étude

Dans le chapitre présentant la méthodologie de notre travail (cf. pp. 145 à 151), nous avons montré que les caractéristiques de durées brutes des macro-classes phonémiques de notre corpus vont dans le sens des études réalisées sur des corpus de parole contrôlée (cf. § «Caractéristiques des durées segmentales en français», p. 170) :

- les voyelles basses sont significativement plus longues que les voyelles hautes et moyennes ;
- les voyelles nasales sont significativement plus longues que les voyelles orales ;
- les constrictives sont significativement plus longues que les autres types de consonnes ;
- et à l'intérieur des macro-classes consonantiques, les consonnes non voisées sont significativement plus longues que les consonnes voisées.

Or, nous avons précisé que ces mesures de durées sont réalisées sur un corpus non contrôlé phonétiquement, allant de la parole lue à la parole spontanée. De même, ces durées moyennes reflètent les tendances de durées intrinsèques des phonèmes quel que soit le contexte phonétique environnant, ou le contexte prosodique (accentué ou inaccentué).

Dans la perspective d'une normalisation de nos durées segmentales, la recherche d'une base de données des durées en français s'imposait, afin d'une part de contrôler la fiabilité de nos propres données, et d'autre part de puiser des durées et écarts types représentatifs. Or, les bases de données pour le français sont trop réduites pour qu'on puisse les utiliser de façon fiable comme valeurs de référence. La normalisation en Z nécessite d'avoir les valeurs moyennes et les écarts types des phonèmes. Cependant, les seules études qui proposent ces données ne correspondent pas à nos besoins : Di Cristo (1985) ne propose des valeurs que pour certaines voyelles, Barbosa & Bailly (1994) disposent des données d'un seul locuteur, issues d'un corpus de logatomes ; O'Shaughnessy (1984) et Bartkova (1988), enfin, proposent des données issues respectivement de la lecture d'un paragraphe et de la lecture de nombres. La base de données la plus appropriée aurait pu être celle fournie par O'Shaughnessy (1984), mais notre corpus comprend, outre de la parole lue, des styles de

parole plus spontanée.

Nous avons donc choisi de normaliser les durées des phonèmes de notre corpus en utilisant nos propres occurrences de phonèmes pour les raisons suivantes :

– nous avons vu que notre corpus, bien que n'étant pas contrôlé, est d'une part, représentatif des occurrences des phonèmes du français (cf. § « Base de données phonémique de notre corpus », p. 145), et d'autre part, présente des durées brutes moyennes en conformité avec les études décrivant les durées intrinsèques des phonèmes en français. Nous sommes cependant consciente des limites de la normalisation en Z, exprimées par Campbell (1992), selon lequel certains phonèmes ont tendance à apparaître majoritairement dans certains contextes phonémiques (par exemple, le phonème [w] en japonais, est très souvent en contexte allongeant /wa/, ce qui peut fausser la durée normalisée de /w/). Disposant pour notre part d'environ 1000 occurrences de phonèmes par locuteur, nous considérons que les variations contextuelles sont estompées par le nombre élevé d'occurrences, et que chaque catégorie de phonème a une distribution normale. Les phonèmes de notre base de données, rassemblés en macro-classes phonémiques, respectent d'ailleurs le critère d'homosédasticité, ce qui prouve que notre corpus est représentatif et robuste (cf. § 3.2.6, p. 146).

– notre étude est la seule, à notre connaissance, qui propose des valeurs de durées moyennes et d'écart types issus de corpus de parole spontanée pour le français.

Nous pouvions prendre en considération la totalité des occurrences phonémiques de notre corpus, quels que soient le locuteur ou le style de parole, comme base de données pour la normalisation. Or, une analyse de variance nous a montré que les locuteurs ont des stratégies de durées phonémiques intrinsèques propres (cf. les détails de l'Anova, p. 150). La normalisation en Z ne pouvait donc pas se faire sur une base de données commune à tous les locuteurs. D'autre part, un des buts de notre recherche étant de mettre en évidence l'influence des différents processus d'encodage (parole spontanée vs. parole lue) sur le rythme de la parole, la normalisation par locuteur présente l'avantage de clarifier la part de variabilité imputable au locuteur, voire au type de parole qu'il représente si les deux locuteurs d'un même style présentent des caractéristiques communes.

Il apparaît en effet que les variations de durée sont fonction des locuteurs, mais également des styles de parole. Campbell (1995a ; sous presse) compare les durées intrinsèques des phonèmes de divers types de discours, allant de la lecture de mots isolés, en passant par la lecture de mots dans des phrases porteuses isolées et des mêmes mots dans des phrases porteuses continues, jusqu'au monologue incontrôlé spontané. Il observe des stratégies de durées segmentales relativement distinctes selon le caractère spontané ou non de la parole. Les durées segmentales sont séparées de manière maximale (afin de distinguer les macro-classes phonémiques) lorsque la parole contient peu d'information contextuelle et que le locuteur tient à être clairement compris. En revanche, lorsque le discours devient plus naturel et que l'auditeur peut se fier à l'organisation prosodique pour aider à l'interprétation de la parole, on trouve alors plus de variance dans les durées phonémiques et moins de distinction des moyennes des macro-classes. Campbell exprime les différentes stratégies de gestion des durées segmentales par les styles de parole en termes de coefficients de variation (écart type rapporté à la moyenne pour chaque type de segment).

Les coefficients de variation pour la parole spontanée indiquent une importante dispersion des durées, alors que ceux reflétant la parole lue indiquent une distinction nette entre les catégories phonémiques.

Afin de rendre compte des différences de gestion des durées segmentales suivant le style de parole, nous avons calculé les coefficients de variations de notre propre corpus, par locuteur et par style de parole (pour une présentation détaillée, cf. Astésano, 1999). Il apparaît clairement que les locuteurs du style *Interview* présentent une dispersion plus importante de durée segmentale que les locuteurs du style *Lecture*. Nos résultats vont donc dans le sens des observations de Campbell, et nous invitent à pratiquer une normalisation interne à chaque locuteur.

5. ANALYSES EMPIRIQUES

5.1. Objectifs de recherche

Nous cherchons à rendre compte du système noyau du rythme du français à travers l'analyse de 3 styles de parole (cf. § « Choix des locuteurs et des types de discours », p. 135). Nous comparons les stratégies accentuelles des différents types de discours dans le cadre du modèle de Di Cristo & Hirst (1997) et Di Cristo (sous presse), qui prend en compte les tendances accentuelles complexes du français. Le choix de trois styles de parole nous permet de mettre en évidence l'invariance potentielle de l'organisation des durées dans divers contextes accentuels, ainsi que la part de variabilité imputable aux différents processus d'encodage (parole lue vs. spontanée). Le style *Journal* a été choisi pour illustrer un type de discours intermédiaire entre la parole lue et la parole spontanée, car il emprunte à la fois aux techniques d'encodage de l'un et l'autre style de parole. Les trois types de discours se situent donc sur une échelle allant de la parole spontanée 'improvisée' à la parole lue préparée.

Nos analyses concernant la durée des syllabes en français sont largement inspirées des travaux de Campbell (1992, 1993, 1995a, et sous presse) et de Campbell & Isard (1991). Par souci d'économie et de lisibilité, nous n'indiqueront pas les dates de publication de cet auteur à chaque fois que nous le citerons ; nous ne mentionnerons donc que son nom lorsque nous ferons référence à la méthode de normalisation ou à ses résultats.

A l'instar de Campbell, nous nous proposons donc de rendre compte des phénomènes d'ajustements temporels des différentes composantes syllabiques, à l'intérieur de la syllabe. Campbell a montré que l'utilisation des durées normalisées (transformées Z) permet de rendre compte de l'existence de la syllabe en tant qu'unité prosodique de niveau supérieur, à l'intérieur de laquelle s'opèrent des principes d'ajustements temporels segmentaux.

Les durées brutes (ms) ne permettent pas de rendre compte de ces stratégies d'allongements différentiels, car elles ne tiennent pas compte de l'élasticité spécifique de chaque type de phonème. De plus, elles sont le résultat de bien trop de facteurs interagissants (durée intrinsèque et co-intrinsèque, contexte prosodique etc.), et ne peuvent donc pas constituer une mesure fiable de l'allongement. C'est pourquoi la transformation en Z de nos données nous a paru pertinente pour rendre compte des distributions différentes de l'allongement en fonction de la catégorie accentuelle. D'autre part, la normalisation de nos

données segmentales nous permet également de procéder à une comparaison fiable des stratégies d'allongement des différents locuteurs de notre corpus, et de mettre en évidence plus précisément les phénomènes temporels directement liés aux différents processus d'encodage (parole lue, lecture interprétée ou parole spontanée) (cf. § « Choix d'une méthode de normalisation adaptée à notre objet d'étude », p. 179).

Par conséquent, tous les résultats auxquels nous ferons allusion sont exprimés en durées normalisées. Nulle part dans la suite de l'exposé ne ferons nous référence à la durée brute (ms), à moins que nous le spécifions explicitement. Le terme 'durée', utilisé seul, renvoie donc par défaut à la durée normalisée (en Z).

Nous proposons donc une analyse de l'ajustement temporel (allongement différentiel) des éléments constitutifs de la syllabe, que nous confronterons aux résultats de Campbell pour l'anglais, et de Fant et al. (1991a et b) pour le français, le suédois et l'anglais.

Cependant, nous analysons l'allongement induit par les phénomènes accentuels à différents niveaux de constituance : le niveau de la syllabe entière, le niveau infrasyllabique des éléments d'Attaque et de Rime, ainsi que le niveau infrasyllabique des Attaques, Noyaux et Codas. Nous émettons en effet l'hypothèse qu'il existe une certaine variabilité induite par les processus d'encodage différents (lu ou spontané), mais qu'une part d'invariabilité doit exister, quel que soit le type de discours, comme reflet du système noyau du rythme du français. La question est donc de savoir s'il existe un niveau privilégié qui rend compte de ce système noyau, et, si variabilité il y a, à quel niveau elle se situe. Nous insisterons davantage sur les résultats concernant le niveau des ANC. L'analyse du niveau intermédiaire des A/Rimes nous permet de vérifier dans quelle mesure la Rime est un constituant plus cohérent que ses parties Noyau et Coda, et plus robuste aux changements de locuteurs ou de styles de parole (Campbell, 1992).

5.2. Plan expérimental

Notre plan expérimental a été conçu pour répondre à trois questions :

1. Les styles de parole ont-ils des stratégies d'allongement spécifiques en fonction des catégories accentuelles ?
2. Si c'est le cas, quel niveau de constituance (syllabique ou infrasyllabique) est particulièrement sensible aux variations stylistiques ?
3. Aux niveaux de constituance infrasyllabique (Attaque + Rime, et Attaque + Noyau + Coda), est-ce que l'allongement se répartit de manière différente sur les constituants syllabiques, en fonction des catégories accentuelles ?

Afin de tester l'influence des styles de paroles

1. sur la durée syllabique globale et
2. sur la durée des constituants syllabiques,

nous avons procédé à des tests d'Analyse de la Variance (Anova)³. Dans le premier cas, les trois styles de parole et les 6 catégories accentuelles sont choisis comme variables indépendantes, et la durée normalisée des syllabes comme variable dépendante ; dans les autres cas, les styles de parole, les constituants syllabiques (ANC, puis A/Rime) et les catégories accentuelles sont choisis comme variables indépendantes, et la durée des constituants syllabiques comme variable dépendante.

Nous avons également procédé à des Anova à l'intérieur des styles, afin de voir si les deux locuteurs dans chaque style ont des stratégies comparables et cohérentes.

Enfin, nous avons procédé à des Anova opposant les styles de parole et les locuteurs pour chaque catégorie accentuelle, afin de mettre en évidence les types de catégories accentuelles qui sont particulièrement sensibles au locuteur et au style de parole.

5.2.1. *Limites méthodologiques*

La nature particulière de notre corpus (parole continue non contrôlée) conduit à deux limitations : d'une part, les occurrences de prééminences ne sont pas distribuées de manière homogène dans les catégories accentuelles, les accents finaux de syntagmes (AFS) étant largement plus représentés que les accents d'unités intonatives terminales (UI-T) ou les accents emphatiques (EMP) par exemple. D'autre part, le nombre de chaque type de constituant syllabique n'est pas le même : la population des Coda par exemple est nettement inférieure à celle des autres éléments syllabiques (pour une présentation détaillée de ces résultats, cf. Astésano, 1999). Cette tendance est probablement due à la spécificité des structures syllabiques du français, qui favorisent largement les syllabes ouvertes (CV) (Delattre, 1965 ; Wioland, 1985).

Le plus petit nombre d'éléments de Coda se reflète dans l'étendue de l'Intervalle de Confiance, plus importante pour ces constituants syllabiques. Nous avons décidé d'éliminer de l'interprétation des résultats, ainsi que des graphiques, les éléments syllabiques qui, combinés avec une catégorie accentuelle, ne représentaient pas plus de 5 items pour les deux locuteurs d'un même style. C'est ainsi que la Coda de l'accent emphatique EMP du style *Lecture* a été éliminée. Globalement et particulièrement pour les styles *Interview* et *Lecture*, les catégories accentuelles AI et EMP ont des populations de Coda très réduites (respectivement, 5 et 7 pour *Interview* ; 6 et 1 pour *Lecture*). Ce type de catégories accentuelles, se trouvant à l'initiale de mots, favorise en effet les constitutions syllabiques CV.

5.2.2. *Déroulement des analyses*

Nous proposons tout d'abord l'analyse des résultats des durées syllabiques. Dans un deuxième temps, nous tâchons d'expliquer les différences observées au niveau syllabique par les stratégies individuelles et stylistiques particulières développées aux niveaux de constituance immédiatement inférieurs (A/Rime et ANC).

Pour chaque niveau de constituance, nous proposons tout d'abord une analyse *globale*

³ Toutes les Analyses de Variance sont réalisées avec le logiciel StatView pour Windows, version 4.57.

des tendances observées pour l'ensemble des locuteurs, puis nous analysons les stratégies des locuteurs à l'intérieur d'un même *style de parole*, afin de distinguer entre la variabilité individuelle et stylistique. Seuls les tableaux et les graphes d'Anova concernant les styles de parole sont présentés dans le texte ; les tableaux et graphes d'Anova concernant chaque locuteur en particulier sont présentés en annexe. Ce choix de présentation est motivé par deux raisons principales : tout d'abord, nous ne voulions pas surcharger le document avec les tableaux et graphiques d'Anova par locuteur ; d'autre part, notre étude est centrée sur la caractérisation prosodique des styles de parole, et la part de variabilité imputable aux différents processus d'encodage (parole lue vs. parole spontanée). L'accent est donc mis sur les stratégies stylistiques, en tant que manifestations 'alloformes' du système noyau.

Nous envisageons l'exposé de nos analyses empiriques à trois niveaux de lecture :

1. un premier niveau qui permet de dégager les conclusions et les résultats principaux : ces parties là sont encadrées dans le texte pour faciliter l'accès aux informations importantes. Une synthèse des résultats les plus pertinents est donc proposée après l'analyse de chaque niveau de constituance, en insistant sur les particularités stylistiques (cf. paragraphes encadrés, pp. 192, 202, 212 et 223).
2. un deuxième niveau de lecture où nous présentons en détail les analyses statistiques ;
3. un troisième niveau de lecture, enfin, qui présente les résultats complémentaires, reportés en annexe.

Enfin, nous concluons par une discussion générale (p. 225) qui reprend les principaux résultats issus de nos analyses. Nous tâchons d'interpréter la pertinence des niveaux d'analyse, et nous insistons particulièrement sur l'intérêt que constitue l'analyse de divers styles de parole dans notre compréhension des facteurs complexes qui influencent les phénomènes de durée dans la parole.

5.3. Exposé des résultats par niveaux de constituance⁴

5.3.1. Niveau de constituance syllabique

5.3.1.1. Tendances générales

Comme nous l'avons indiqué dans la partie méthodologique (cf. p. 156), nous avons identifié les catégories accentuelles auditivement. Nous pouvons donc interpréter les allongements subits par les syllabes accentuées comme indices d'accentuation. Nous verrons d'ailleurs que sur la base de la distinction *fonctionnelle* de nos catégories accentuelles, les accents *sémantico-pragmatiques* se distinguent significativement des accents *rythmiques*. Ainsi, ces derniers ont globalement une durée normalisée moyenne inférieure à 1, alors que les premiers ont une durée de Z supérieure à 2. Campbell prend l'Ecart type comme étant une fourchette critique d'interprétation de l'allongement : $Z > 1$ ou $Z < -1$; toutes les syllabes dont la mesure moyenne est de +1 sont considérées comme

⁴ Ne sont présentés dans ce chapitre que les résultats principaux des durées segmentales. Pour une présentation détaillée des résultats par locuteur et par catégorie accentuelle (graphes et tableaux des tests post hoc des Anova), nous renvoyons le lecteur à Astésano, 1999.

longues, et celles < -1 comme courtes. Nos accents *sémantico-pragmatiques* entrent donc bien dans la catégorie des syllabes longues, alors que nos accents *rythmiques*, sans être courts, sont néanmoins de durée moyenne. Précisons que nos syllabes inaccentuées (INA) ont une valeur moyenne de Z d'environ $-0,5$. Selon les critères de Campbell, elles ne peuvent pas être considérées comme courtes. Nous pensons que nos résultats différents en termes de seuil d'allongement sont à imputer à la spécificité rythmique du français, pour lequel le contraste entre les syllabes inaccentuées et accentuées est moins important que pour l'anglais (cf. § « Rythmicité accentuelle vs. rythmicité syllabique », p. 32, pour une discussion).

Nous avons procédé à une analyse de la variance (Anova) pour rendre compte des tendances générales d'allongement syllabique en fonction des catégories accentuelles et des 6 locuteurs. Pour cela, les 6 locuteurs et les 6 catégories accentuelles ont été choisis comme variables indépendantes, et la durée comme variable dépendante.

Le Tableau 5 présente les résultats de l'Anova générale pour la durée des syllabes en fonction des catégories accentuelles, pour les 6 locuteurs.

Tableau 5 : Tableau d'Anova générale pour les durées normalisées des syllabes, par catégorie accentuelle, pour les 6 locuteurs. (DDL = Degré De Liberté).

Tableau d'ANOVA pour durées Normalisées					
	DDL	Somme des carrés	Carré moyen	Valeur de F	Valeur de p
Locuteurs	5	7,879	1,576	,571	,7222
Catégories Accentuelles	5	1994,360	398,872	144,555	<,0001
Locuteurs * Cat. Acc.	25	363,826	14,553	5,274	<,0001
Résidus	2577	7110,754	2,759		

Les tests d'Anova pour les durées moyennes des syllabes par catégories accentuelles, tous locuteurs (et par conséquent, tous styles) confondus montrent que les locuteurs ne se distinguent pas significativement sur la base des durées globales moyennes. Les différentes catégories accentuelles sont par contre significativement distinctes les unes des autres. Ce résultat indique que notre distinction fonctionnelle *a priori* des catégories accentuelles est pertinente.

La Figure 16 rend compte des tendances générales de l'allongement syllabique en fonction des catégories accentuelles, tous locuteurs confondus.

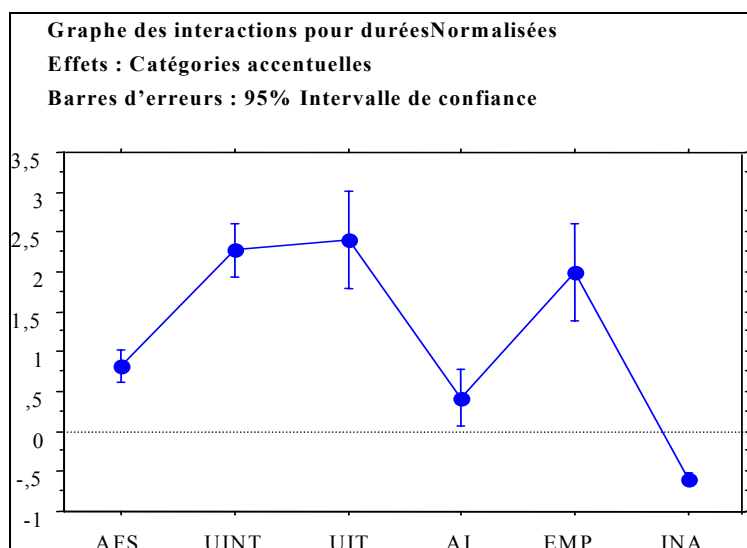


Figure 16 : Tendances générales de l'allongement syllabique en fonction des catégories accentuelles, tous locuteurs confondus. L'axe des abscisses représente les catégories accentuelles et l'axe des ordonnées indique les durées normalisées (Transformées Z). (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique).

D'une manière générale, les accents sémantico-pragmatiques (UI-NT, UI-T et EMP) ne se distinguent pas significativement sur la base de leurs durées moyennes ; ils diffèrent par contre significativement des accents rythmiques. Ces derniers (AFS, AI) sont également significativement différents l'un de l'autre, bien que, nous le verrons, ces deux types d'accents rythmiques ne se distinguent pas significativement l'un de l'autre lorsqu'on compare leurs durées respectives par locuteur. Enfin, les syllabes inaccentuées (INA) sont significativement distinguées des autres catégories accentuelles.

Les différences importantes et significatives de durée reflètent donc les distinctions fonctionnelles de l'accentuation.

Le tableau général de l'Anova indique enfin que l'interaction entre les locuteurs et les catégories accentuelles est également significative, ce qui laisse penser que chaque locuteur a ses propres stratégies d'allongement syllabique en fonction des catégories accentuelles.

La Figure 17 présente les stratégies personnelles des locuteurs en termes de durées syllabiques.

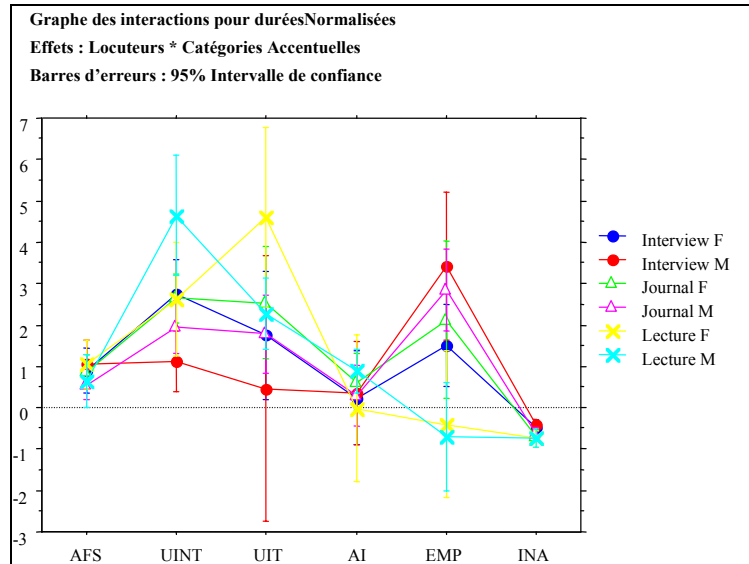


Figure 17 : Durées syllabiques normalisées (Z) par catégorie accentuelle, et par locuteur. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique)

On remarque une assez grande disparité dans les stratégies d'allongement syllabique des locuteurs, ce qui explique que l'interaction Locuteurs * Catégories Accentuelles est significative. Les principales distinctions que l'on peut observer entre les locuteurs se concentrent essentiellement sur les catégories *sémantico-pragmatiques* (UI-NT, UI-T et EMP). En effet, il semble que les stratégies d'allongement pour les accents *rythmiques* AI et AFS soient relativement stables, quel que soit le locuteur ou le style de parole. Il apparaît d'une part que les locuteurs du style *Lecture* allongent considérablement les accents finaux d'unités intonatives (UI-NT et UI-T) comparativement aux autres types d'accents, et comparativement aux UI-NT et UI-T des autres locuteurs. D'autre part, l'allongement des EMP pour ces deux locuteurs est considérablement moins important que pour les autres locuteurs.

Afin d'analyser plus précisément ce qui caractérise un style de parole sur le plan des stratégies d'allongement syllabique, et ce qui pourrait distinguer les locuteurs à l'intérieur d'un même style, nous proposons de présenter nos résultats par style de parole. Nous essaierons, dans un deuxième temps, d'expliquer les différences intra-stylistiques et inter-stylistiques observées pour les durées syllabiques, en élargissant notre champ de recherche aux stratégies d'allongement différentiel infrasyllabique en ANC et A/Rime. Sur la base de ces observations, nous tâcherons de proposer quel niveau d'analyse des durées (syllabique, ou en constituants syllabiques ANC ou A/Rime) est le plus pertinent pour la caractérisation des stratégies accentuelles des styles de parole, et plus généralement, nous déduirons de nos données le cadre d'analyse le plus pertinent pour l'analyse métrique.

5.3.1.2. Analyse des différences inter-locuteurs à l'intérieur des styles au niveau des durées syllabiques

Interview :

Le Tableau 6 présente les résultats de l'Anova générale pour le style *Interview*.

Tableau 6 : Tableau d'Anova générale pour les durées normalisées des syllabes, par catégorie accentuelle, pour les locuteurs du style *Interview*.

Tableau d'ANOVA pour duréesNormalisées					
	DDL	Somme des carrés	Carré moyen	Valeur de F	Valeur de p
Locuteurs	1	,732	,732	,220	,6394
Catégories Accentuelles	5	709,071	141,814	42,560	<,0001
Locuteurs * Cat. Acc.	5	75,086	15,017	4,507	,0005
Résidus	1014	3378,730	3,332		

Si les locuteurs *Interview M* et *Interview F* ne se distinguent pas significativement sur la base de leurs durées globales moyennes, en revanche l'interaction des variables indépendantes indique qu'ils ont des stratégies d'allongement syllabique propres.

En effet, la différence générale observée pour l'ensemble des locuteurs se retrouve également à l'intérieur du style *Interview* : les locuteurs semblent être plus cohérents dans l'allongement syllabique des accents *rythmiques* que des accents *sémantico-pragmatiques*. La

Figure 18 présente les résultats de durée normalisée syllabique des deux locuteurs.

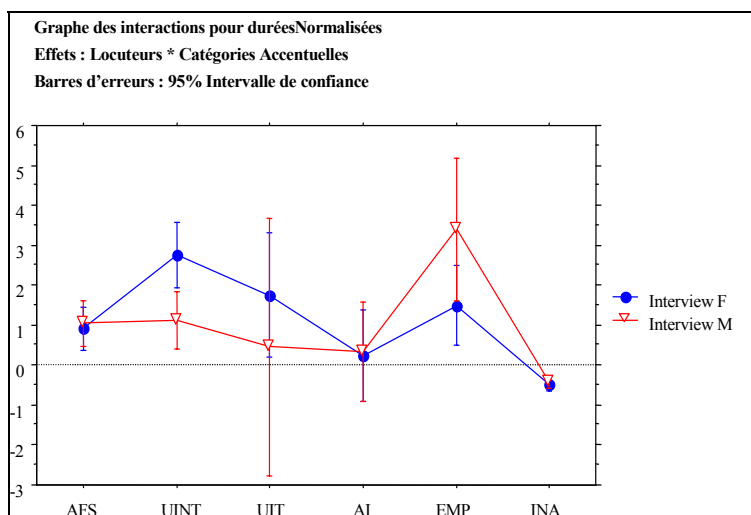


Figure 18 : Durées syllabiques normalisées par catégorie accentuelle, pour les locuteurs du style *Interview*. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique)

A l'intérieur de la macro-classe *sémantico-pragmatique*, il semble que la locutrice *Interview F* privilégie l'allongement des accents d'unités intonatives par rapport aux autres

types d'accents, alors que le locuteur *Interview M* réalise un allongement beaucoup plus substantiel pour les EMP par rapport aux autres accents. On peut faire le rapprochement avec les stratégies de densité accentuelle et de pauses propre à chaque locuteur, où la locutrice privilégie les accents de frontière et les pauses respiratoires, alors que le locuteur produit plus d'accents initiaux, notamment emphatiques (cf. Tableau 34, p. 275). Globalement, les résultats des tests post hoc montrent que le locuteur *Interview M* accentue les différences entre les accents initiaux (les accents initiaux emphatiques EMP étant significativement plus allongés que les accents rythmiques AI), alors que la locutrice *Interview F* privilégie la distinction entre les accents finaux (les accents d'unités intonatives UI-NT et UI-T étant notablement et significativement plus allongés que les accents finaux de syntagmes AFS).

Lecture :

Le Tableau 7 présente les résultats de l'Anova générale pour le style *Lecture*.

Tableau 7 : Tableau d'Anova générale pour les durées normalisées des syllabes, par catégorie accentuelle, pour les locuteurs du style *Lecture*

Tableau d'ANOVA pour duréesNormalisées					
	DDL	Somme des carrés	Carré moyen	Valeur de F	Valeur de p
Locuteurs	1	,015	,015	,007	,9351
Catégories Accentuelles	5	710,022	142,004	63,434	<,0001
Locuteurs * Cat.Acc.	5	51,830	10,366	4,630	,0004
Résidus	432	967,089	2,239		

Comme pour le style *Interview*, on remarque que les deux locuteurs du style *Lecture* ne se distinguent pas significativement sur la base de leurs durées globales moyennes. Par contre, ils semblent avoir des stratégies propres d'allongement syllabique par catégories accentuelles, car l'interaction Locuteurs * Catégories accentuelles est significative.

Plus particulièrement, les locuteurs *Lecture M* et *Lecture F* semblent avoir des stratégies différentes d'allongement syllabique des catégories accentuelles sémantico-pragmatiques. Contrairement au style *Interview* cependant, ces différences ne concernent que les accents d'unités intonatives (UI-NT et UI-T), les EMP étant identiques pour les deux locuteurs.

La Figure 19 présente les durées syllabiques des locuteurs du style *Lecture*.

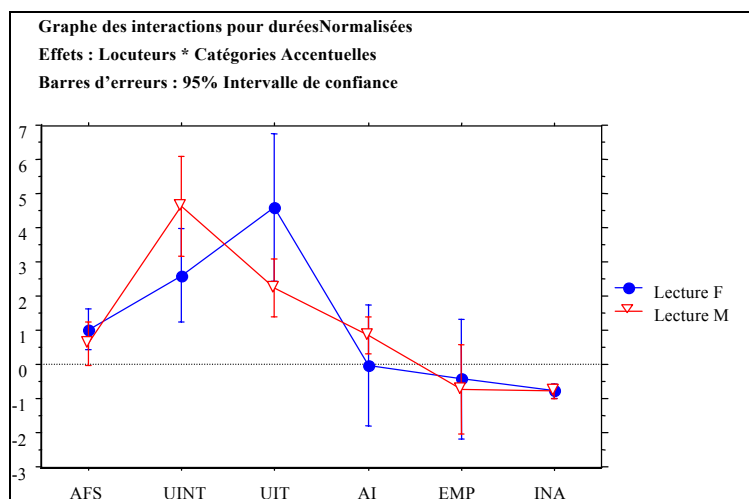


Figure 19 : Durées syllabiques normalisées par catégorie accentuelle, pour les locuteurs du style *Lecture*. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique)

Il apparaît que la locutrice privilégie l'allongement des UI-T, et le locuteur l'allongement des UI-NT. Néanmoins, les locuteurs marquent tous deux une nette différence (significative) entre les AFS et les accents d'unités intonatives (UI-NT et UI-T). Contrairement au style *Interview*, les deux locuteurs du style *Lecture* ont un allongement très réduit de l'accent emphatique, qui ne se distingue pas significativement de l'AI rythmique pour la locutrice *Lecture F*.

Leurs stratégies d'allongement pour les accents d'unités intonatives sont parfaitement symétriques et l'allongement de l'accent emphatique est identique, ce qui explique pourquoi ils ne se distinguent absolument pas sur la base de leur durée moyenne. Nous verrons que ces différences s'expliquent par des stratégies particulières d'allongement des constituants syllabiques pour chaque locuteur.

Journal :

Le Tableau 8 présente les résultats de l'Anova générale pour le style *Journal*.

Tableau 8 : Tableau d'Anova générale pour les durées normalisées des syllabes, par catégorie accentuelle, pour les locuteurs du style *Journal*

Tableau d'ANOVA pour durées Normalisées					
	DDL	Somme des carrés	Carré moyen	Valeur de F	Valeur de p
Locuteurs	1	3,942	3,942	1,613	,2044
Catégories Accentuelles	5	1134,970	226,994	92,852	<,0001
Locuteurs * Cat. Acc.	5	21,573	4,315	1,765	,1172
Résidus	1131	2764,935	2,445		

A la différence des deux autres styles de parole, les locuteurs du style *Journal* ne semblent pas avoir de stratégies d'allongement syllabique différentes, car l'interaction Locuteurs * Catégories Accentuelles n'est pas significative.

Comme on le voit sur la Figure 20, les deux locuteurs sont cohérents dans la gestion de leurs durées syllabiques.

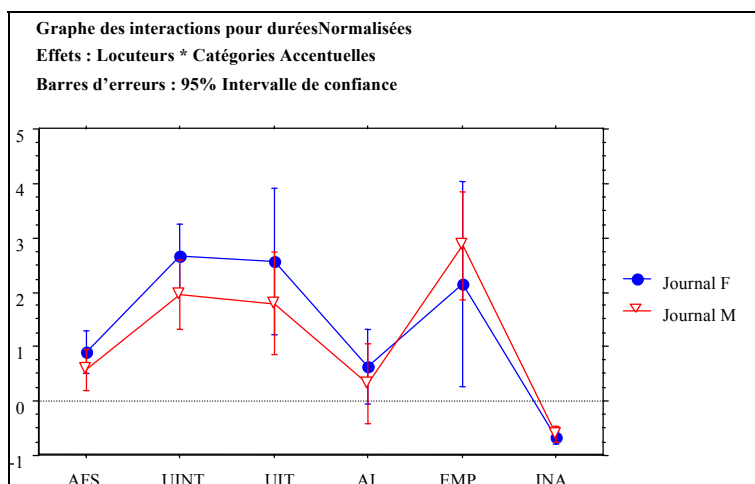


Figure 20 : Durées syllabiques normalisées par catégorie accentuelle, pour les locuteurs du style *Journal*. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique)

Comme pour les autres styles de parole, la différence inter-locuteurs se manifeste également pour les catégories accentuelles *sémantico-pragmatiques*, mais dans une bien moindre mesure. De même que pour le style *Interview*, la locutrice *Journal F* tend à produire des accents d'unités intonatives plus longs que le locuteur *Journal M*, ce dernier privilégiant l'allongement de l'EMP comparativement aux autres types d'accents. On peut faire la même remarque concernant les stratégies accentuelles globales des locutrices, en règle générale, par rapport aux locuteurs (cf. Tableau 34, p. 275).

Globalement, les deux locuteurs du style *Journal* marquent de manière optimale la distinction entre les macro-classes accentuelles (sémantico-pragmatiques vs. rythmiques) : les accents d'unités intonatives sont significativement plus allongés que les AFS, et les EMP sont significativement plus allongés que les AI.

5.3.1.3. Niveau de constituance syllabique : Conclusions partielles

Nous pouvons dorénavant observer une tendance générale commune à tous les styles de paroles.

La distinction fonctionnelle que nous avons postulée *a priori* entre les catégories accentuelles est en effet pertinente, quel que soit le style de parole (l'hypothèse de distinction fonctionnelle des prééminences est vérifiée). La macro-classe des catégories accentuelles sémantico-pragmatiques est caractérisée globalement par un allongement syllabique plus important que la macro-classe des catégories accentuelles rythmiques. Certaines limitations stylistiques peuvent néanmoins être observées, notamment pour le style *Lecture* où l'accent initial emphatique EMP est particulièrement non marqué par l'allongement, et pour le style *Interview*, pour lequel l'accent final d'unité intonative terminale (UI-T) n'est pas significativement plus allongé que l'accent final de syntagme (AFS).

Les différences entre les styles de parole peuvent se résumer ainsi : le style *Lecture* marque particulièrement les accents finaux d'unités intonatives (UI-NT et UI-T), alors que le style *Journal* marque la différence entre les accents sémantico-pragmatiques (UI-NT, UI-T et EMP) et les accents rythmiques (AFS et AI) de manière optimale. Le style *Interview*, qui présente le moins de cohérence intra-stylistique au niveau de constituance syllabique, se situe entre les deux tendances.

L'analyse des stratégies d'allongement différentiel aux niveaux infrasyllabiques devrait nous permettre d'apporter un éclairage supplémentaire sur les stratégies d'allongement propre à chaque style de parole. Nous verrons en effet que les remarques générales concernant la gestion des durées par les différents styles de parole au niveau syllabique doivent être relativisées par une analyse des stratégies d'allongement différentiel.

5.3.2. Niveaux infrasyllabiques : remarques méthodologiques préliminaires

D'un point de vue purement statistique, il faut noter que les tests post hoc de Fisher indiquant les relations de distinction entre deux catégories accentuelles peuvent présenter certaines différences, selon que l'on se situe au niveau syllabique global ou aux niveaux infrasyllabiques, et ce pour deux raisons principales : d'une part, le nombre de degrés de liberté varie, ce qui influe sur les résultats de l'Anova générale et des tests post hoc ; d'autre part, une différence de relation, entre deux catégories accentuelles au niveau syllabique et au niveau infrasyllabique, indique que ces deux catégories accentuelles peuvent avoir des caractéristiques communes à un niveau de constituance, mais des différences notables à un autre niveau de constituance. Pour ne prendre qu'un exemple, la différence entre les UI-NT et les UI-T pour la locutrice *Interview F* n'est pas significative au niveau syllabique (valeur de $p = ,1511$), alors qu'elle l'est au niveau infrasyllabique des ANC (valeur de $p = ,0135$), les deux catégories accentuelles ayant un allongement différentiel propre. Cependant, de telles différences de relations hiérarchiques à divers niveaux de constituance sont relativement rares.

De plus, afin de pouvoir rendre compte explicitement des résultats de significativité dans les relations entre les composantes syllabiques (A/Rime d'une part, et ANC de l'autre) à l'intérieur de chaque catégories accentuelles, nous avons dû concaténer dans une même colonne les informations de constituance syllabique et celles de catégories accentuelles ; chaque constituant syllabique est ainsi associé directement à la catégorie accentuelle à laquelle il appartient, ce qui permet une opposition terme à terme détaillée dans les tests post hoc de Fisher. Ces tableaux de comparaisons binaires comportent des centaines de lignes pour chaque locuteur. Pour des raisons de lisibilité, nous avons rassemblé dans un tableau récapitulatif (Tableau 9) , les résultats des tests binaires de Fisher les plus pertinents, permettant de visualiser les relations hiérarchiques entre les composantes infrasyllabiques par catégories accentuelles, par locuteur et par style, au niveau des ANC et des A/Rimes.

Le Tableau 9 présente donc le résumé des oppositions terme à terme des composantes syllabiques (ANC et A/Rime), en fonction des catégories accentuelles. Cependant, seules les oppositions à l'intérieur d'une même catégorie accentuelle sont proposées, par souci de lisibilité et de pertinence. Les oppositions *entre* les catégories seront éventuellement présentées au fur et à mesure des discussions, dans le cas où elles sont pertinentes dans la distinction fonctionnelle des catégories. Ce résumé est proposé par locuteur, ainsi que par style de parole.

Tableau 9 : Résumé des résultats des tests Anova : les rapports entre les différentes composantes syllabiques (ANC ; et A/Rime) sont présentés par catégorie accentuelle et par locuteur. Une synthèse par style est proposée. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique)

	AFS	UI-NT	UI-T	AI	EMP
Interview M	(A=N)>C	A<N>C	A=N=C	A=N=C	A=N=C
	A=R	A<R	A=R	A=R	A>R
Interview F	(A=N)>C	(A<N)>C	A=N=C	A=N	A>(N=C)
	A=R	A<R	A=R	A=R	A>R
Interview	(A=N)>C	A<N>C	A=N=C	A>N	A>(N=C)
	A=R	A<R	A=R	A=R	A>R
Lecture M	A=(N>C)	(A<N)=C	A=N=C	A>(N=C)	A=N
	A=R	A<R	A<R	A>R	A=R
Lecture F	A=(N>C)	(A<N)>C	(A=N)<C	A=N=C	A=N=C
	A=R	A<R	A<R	A=R	A=R
Lecture	A=(N>C)	A<N>C	(A=N)<C	A>(N=C)	A=N
	A=R	A<R	A<R	A>R	A=R
Journal M	A<N>C	A<N>C	A<(N=C)	(A>N)=C	A>(N=C)
	A<R	A<R	A<R	A>R	A=R
Journal F	(A=N)>C	A<N>C	A=N=C	A>(N=C)	A=N=C [A<C]
	A=R	A<R	A=R	A>R	A>R
Journal	A<N>C	A<N>C	A<(N=C)	A>(N=C)	A>(N=C)
	A<R	A<R	A<R	A>R	A>R

Dans le tableau ci-dessus, les différences de durées normalisées entre les composantes syllabiques sont résumées comme suit :

- Le signe '=' signifie que les différences entre les composantes syllabiques ne sont pas significatives.
- Les signes '>' et '<' signifient qu'une composante syllabique présente un allongement significativement plus important ou significativement moins important qu'une autre composante.
- Les parenthèses indiquent que deux des trois composantes syllabiques entretiennent le même rapport avec la troisième composante. Par exemple : (A=N)>C signifie que A ne se distingue pas significativement de N sur le plan de l'allongement, et que A et N sont l'un et l'autre significativement plus allongés que C.

Au cours des prochains paragraphes, les comparaisons de stratégies entre les composants syllabiques que nous présenterons seront issues de ce Tableau 9.

5.3.3. Niveau de constituance infrasyllabique en A/Rime

5.3.3.1. Tendances générales

Ce niveau de constituance permet essentiellement de distinguer les catégories accentuelles en fonction de leur *distribution*. Les accents *initiaux* ont tendance à recevoir

l'allongement sur l'Attaque (A), alors que les accents *finaux* ont une Rime généralement plus allongée. Cette dernière tendance n'est cependant significative que pour les UI-NT, et non significative pour les AFS, pour tous les locuteurs. Nous verrons ultérieurement qu'il existe des différences de styles de parole pour ce qui concerne notamment les UI-T, les AI et les EMP.

La Figure 21 présente les tendances d'allongement différentiel sur l'Attaque et la Rime en fonction des catégories accentuelles, et tous locuteurs confondus.

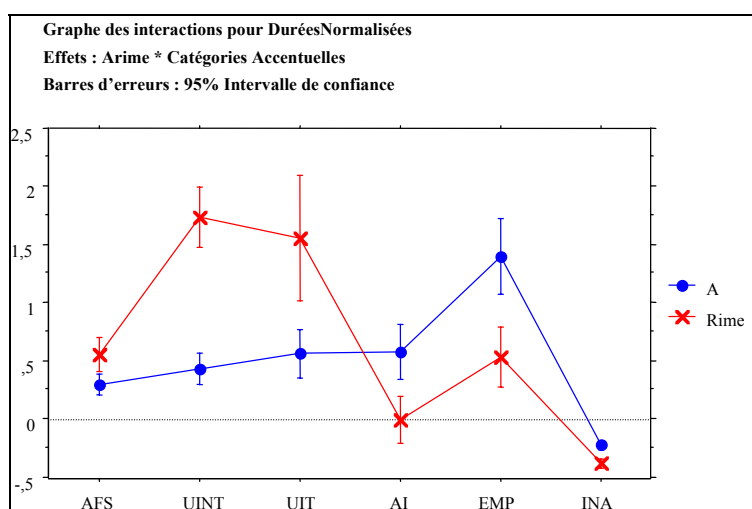


Figure 21 : Tendances générales observées pour la hiérarchie des constituants syllabiques sur la base des durées normalisées par catégorie accentuelle, tous locuteurs et styles de parole confondus. L'axe des abscisses représente les catégories accentuelles ; sur l'axe des ordonnées se trouvent les durées normalisées (Transformée Z) ; A = Attaque. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique)

On a vu que les locuteurs se distinguaient significativement sur la base des durées syllabiques en fonction des catégories accentuelles (cf. Tableau 5).

Si on entre dans le détail de la gestion par chaque locuteur des durées infrasyllabiques (A/Rime), que peut-on observer ?

Le Tableau 10 donne les résultats de l'Anova générale pour la durée des A/Rime en fonction des catégories accentuelles, pour tous les locuteurs. Les locuteurs, les constituants syllabiques (Attaque et Rime) et les catégories accentuelles constituent les variables indépendantes dans le calcul de l'Anova, la durée des constituants syllabiques étant la variable dépendante.

Tableau 10 : Tableau d'Anova générale pour les durées normalisées des A/Rimes, par catégorie accentuelle, pour les 6 locuteurs

Tableau d'ANOVA pour DuréesNormalisées					
	DDL	Somme des carrés	Carré moyen	Valeur de F	Valeur de p
Locuteurs	5	9,285	1,857	1,917	,0882
ARime	1	5,068	5,068	5,231	,0222
Locuteurs * ARime	5	18,040	3,608	3,724	,0023
Catégories Accentuelles	5	948,008	189,602	195,700	<,0001
Locuteurs * Cat. Acc.	25	150,754	6,030	6,224	<,0001
ARime * Cat. Acc.	5	192,985	38,597	39,839	<,0001
Locuteurs * AR * Cat. Acc.	25	51,821	2,073	2,140	,0008
Résidus	4964	4809,301	,969		

A la différence de ce qu'on a pu observer au niveau syllabique, les locuteurs ne se distinguent pas significativement sur la base de leurs durées globales moyennes. Par contre, ils semblent avoir des stratégies propres d'allongement différentiel, car l'interaction Locuteurs * A/Rime est significative. Ces différences inter-locuteurs sont présentées dans la Figure 22, qui donne les tendances générales des allongements différentiels par locuteur, toutes catégories accentuelles confondues.

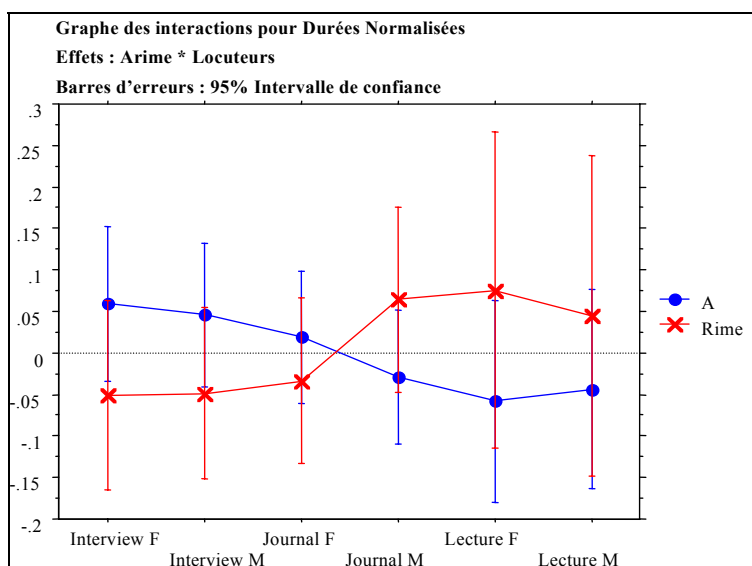


Figure 22 : Durée moyenne des Attaques et des Rimes (A/Rimes) par locuteur, toutes catégories accentuelles confondues. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique)

Deux tendances majeures sont remarquables :

- d'une part, les locuteurs des styles *Interview* et *Lecture* sont relativement cohérents à l'intérieur de leur style de parole respectifs. Les styles *Interview* et *Lecture* ont des stratégies cependant totalement opposées, à savoir que les locuteurs d'*Interview* privilégient en moyenne l'allongement de l'Attaque, alors que ceux de *Lecture* privilégient l'allongement de la Rime.

– d’autre part, les locuteurs du style *Journal* ne sont pas du tout cohérents : la locutrice privilégie l’allongement de l’Attaque, à l’instar des locuteurs du style *Interview*, alors que le locuteur *Journal M* allonge davantage la Rime, comme les locuteurs de *Lecture*. Il est intéressant de noter que cette différence d’allongement différentiel est compensée au niveau de la durée syllabique entière, comme nous l’avons vu précédemment (cf. Tableau 8, p. 190).

Comme pour les résultats syllabiques, l’Anova générale montre que les locuteurs ont, sur un plan général, des stratégies personnelles d’allongement par catégories accentuelles. Ceci se reflète également dans la valeur p de l’interaction entre Locuteurs * A/Rime * Catégories Accentuelles.

Afin de clarifier l’exposé, nous allons présenter et commenter ces tendances avec plus de détails, notamment en analysant les stratégies d’allongement en fonction des catégories accentuelles, par styles de parole.

5.3.3.2. Analyse des différences inter-locuteurs à l’intérieur des styles au niveau des constituants syllabiques A/Rime

Pour tous les styles de parole, nous proposons en annexe un graphe des durées infrasyllabique en A/Rime par locuteur, afin de clarifier la présentation.

Interview :

Le Tableau 11 présente les tests d’Anova générale pour les durées infrasyllabiques A/Rime en fonction des catégories accentuelles, pour les locuteurs du style *Interview*.

Tableau 11 : Tableau d’Anova générale pour les durées normalisées des A/Rimes, par catégorie accentuelle, pour les locuteurs du style *Interview*

Tableau d’ANOVA pour DuréesNormalisées					
	DDL	Somme des carrés	Carré moyen	Valeur de F	Valeur de p
A/Rime	1	1,611	1,611	1,624	,2027
Catégories Accentuelles	5	342,737	68,547	69,085	<,0001
A/Rime * Cat. Acc.	5	48,589	9,718	9,794	<,0001
Locuteurs	1	1,815	1,815	1,829	,1764
A/Rime * Locuteurs	1	,271	,271	,273	,6016
Cat. Acc. * Locuteurs	5	38,442	7,688	7,749	<,0001
AR * Cat. Acc. * Locuteurs	5	1,157	,231	,233	,9480
Résidus	1887	1872,307	,992		

Les locuteurs du style *Interview* ne se distinguent pas sur la base des durées globales des A/Rimes. Ils semblent avoir également la même stratégie globale d’allongement des A/Rimes (p = ,6016), et plus remarquable encore, ils semblent avoir les mêmes stratégies d’allongement différentiel à travers les catégories accentuelles (p = ,9480). En effet, l’interaction Catégories Accentuelles * Locuteurs est hautement significative, ce qui indique qu’ils ont des stratégies d’allongement globales différentes. Ce résultat est d’ailleurs pleinement cohérent eu égard à la valeur p obtenu pour les durées des syllabes entières (cf. Tableau 6, p. 188, valeur de p pour l’interaction = ,0005).

Il apparaît donc que les différences de stratégies globales (niveau syllabique) d’allongement sont compensées par des stratégies d’allongement différentiel similaires au niveau infrasyllabique (A/Rime).

La Figure 23 présente la répartition de l'allongement sur l'Attaque (A) et la Rime (R) pour le style *Interview*, en fonction des catégories accentuelles.

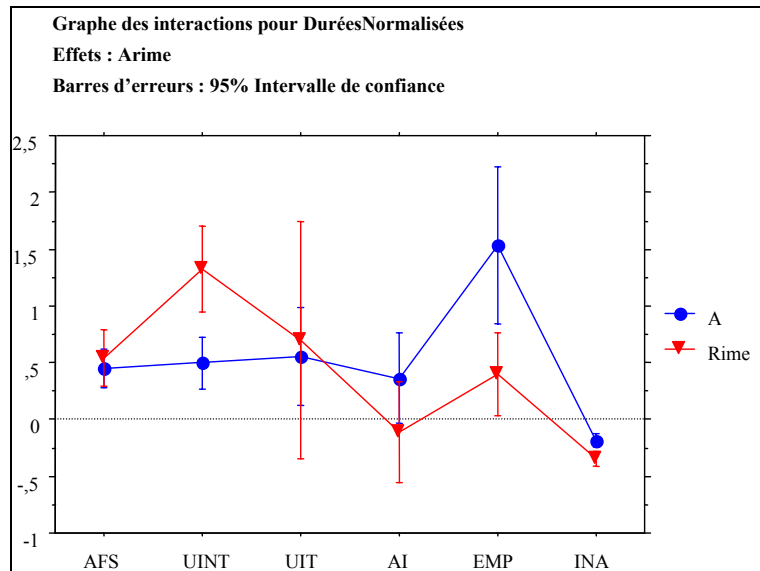


Figure 23 : Durée moyenne des Attaques et des Rimes (A/Rimes) pour le style *Interview*, en fonction des catégories accentuelles. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique)

Les stratégies d'allongement différentiel des deux locuteurs sont très cohérentes. Pour chaque locuteur en effet, seul UI-NT présente une Rime significativement supérieure à l'Attaque. Également, seul EMP présente une Attaque significativement supérieure à la Rime. Cependant, les deux locuteurs respectent les tendances générales des autres locuteurs, selon lesquelles les accents finaux ont des Rimes plus allongées que les Attaques, et les accents initiaux des Attaques plus allongées que les Rimes.

Lecture :

Le Tableau 12 présente les tests d'Anova générale pour les durées infrasyllabiques A/Rime en fonction des catégories accentuelles, pour les locuteurs du style *Lecture*.

Tableau 12 : Tableau d'Anova générale pour les durées normalisées des A/Rimes, par catégorie accentuelle, pour les locuteurs du style Lecture

Tableau d'ANOVA pour DuréesNormalisées					
	DDL	Somme des carrés	Carré moyen	Valeur de F	Valeur de p
ARime	1	9,156	9,156	10,214	,0014
Catégories Accentuelles	5	342,637	68,527	76,448	<,0001
ARime * Cat. Acc.	5	77,820	15,564	17,363	<,0001
Locuteurs	1	,014	,014	,016	,8997
ARime * Locuteurs	1	,037	,037	,041	,8390
Cat. Acc. * Locuteurs	5	26,373	5,275	5,884	<,0001
AR * Cat. Acc. * Locuteurs	5	10,250	2,050	2,287	,0444
Résidus	826	740,415	,896		

Comme pour le style *Interview*, les locuteurs de *Lecture* ont des stratégies globales d'allongement similaires. Par contre, l'interaction significative entre les trois variables indépendantes indique que les deux locuteurs ont des stratégies d'allongement différentiel par catégories accentuelles différentes.

La Figure 24 présente la répartition de l'allongement sur l'Attaque (A) et la Rime (R) pour le style *Lecture*, en fonction des catégories accentuelles.

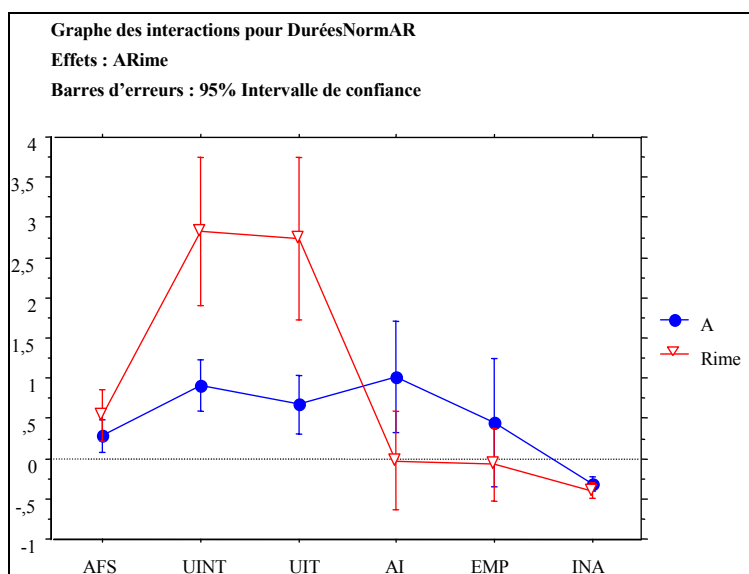


Figure 24 : Durée moyenne des Attaques et des Rimes (A/Rimes) pour le style Lecture, en fonction des catégories accentuelles. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique)

Les deux locuteurs ont les mêmes stratégies d'allongement en ce qui concerne les accents finaux : les UI-NT et UI-T ont une Rime significativement plus allongée que l'Attaque, alors que pour l'AFS, la Rime n'est pas significativement différente de l'Attaque. Pour l'EMP également, la Rime n'est pas significativement différente de l'Attaque. La différence observée dans les résultats de l'Anova générale est sans doute à imputer à l'AI rythmique, pour lequel *Lecture M* présente une Attaque significativement

supérieure à la Rime, alors que ce n'est pas le cas pour la locutrice *Lecture F*. Cette différence s'estompe lorsque le style *Lecture* est pris globalement comme variable indépendante.

En général, les locuteurs de *Lecture* respectent les tendances globales selon lesquelles l'Attaque est supérieure à la Rime pour les accents initiaux, et la Rime est supérieure à l'Attaque pour les accents finaux. Cependant, ils marquent nettement plus les accents finaux d'unités intonatives (UI-NT et UI-T).

***Journal* :**

Le Tableau 13 présente les tests d'Anova générale pour les durées infrasyllabiques A/Rime en fonction des catégories accentuelles, pour les locuteurs du style *Journal*.

Tableau 13 : Tableau d'Anova générale pour les durées normalisées des A/Rimes, par catégorie accentuelle, pour les locuteurs du style *Journal*

Tableau d'ANOVA pour DuréesNormalisées					
	DDL	Somme des carrés	Carré moyen	Valeur de F	Valeur de p
ARime	1	3,312	3,312	3,394	,0656
Catégories Accentuelles	5	485,811	97,162	99,569	<,0001
ARime * Cat. Acc.	5	148,816	29,763	30,501	<,0001
Locuteurs	1	,585	,585	,599	,4390
ARime * Locuteurs	1	6,863	6,863	7,033	,0081
Cat. Acc. * Locuteurs	5	7,274	1,455	1,491	,1895
AR * Cat. Acc. * Locuteurs	5	7,803	1,561	1,599	,1569
Résidus	2251	2196,579	,976		

L'interaction Catégories Accentuelles * A/Rime * Locuteurs n'est pas significative, alors qu'on a vu (cf. Figure 22, p. 196) qu'ils avaient des stratégies globales de A/Rimes relativement différentes (valeur de p = ,0081). Cela indique qu'ils ont sans doute une hiérarchie similaire entre constituants à l'intérieur des catégories accentuelles.

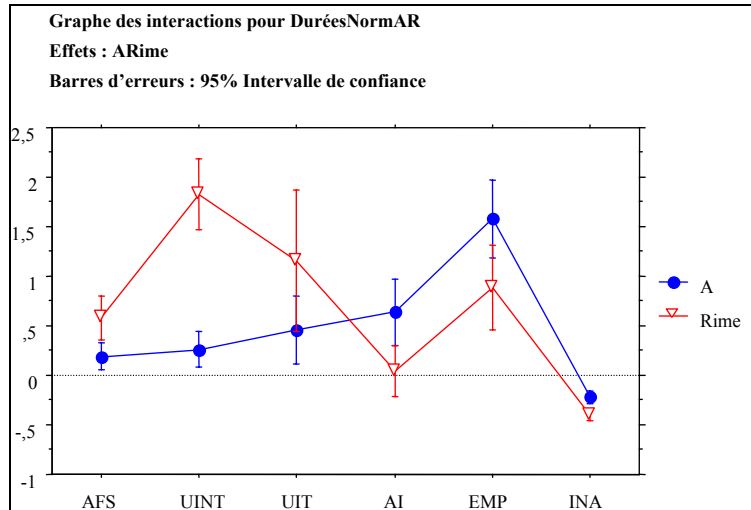


Figure 25 : Durée moyenne des Attaques et des Rimes (A/Rimes) pour le style *Journal*, en fonction des catégories accentuelles. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique)

Globalement, les deux locuteurs présentent les mêmes tendances générales : les Attaques sont plus allongées que les Rimes pour les accents initiaux, alors que l'inverse est vrai pour les accents finaux. Ce qui les distingue est la significativité des relations hiérarchiques entre constituants. Pour le locuteur *Journal M*, tous les accents finaux, y compris AFS, ont des Rimes significativement plus allongées que les Attaques, alors que pour la locutrice *Journal F*, à l'instar des locuteurs du style *Interview*, seul l'accent UI-NT présente une Rime significativement plus allongée que l'Attaque. En revanche, *Journal F* présente des Attaques significativement plus allongées pour les EMP et AI initiaux, alors que le locuteur *Journal M*, à l'instar du style *Lecture*, ne distingue pas significativement l'Attaque de la Rime pour EMP, la Rime étant également fortement allongée.

Cependant, ces différences de degré de hiérarchie s'estompent au niveau du style *Journal* global, qui marque de manière optimale les différences distributionnelles entre les accents finaux et les accents initiaux.

5.3.3.3. Niveau de constituance infrasyllabique en A/Rime : Conclusions partielles

Si le niveau syllabique permet de rendre compte de la distinction *fonctionnelle* entre les catégories accentuelles (macro-classes sémantico-pragmatique vs. rythmique), le niveau de constituance infrasyllabique en A/Rime permet de mettre en évidence la distinction *distributionnelle* entre les catégories accentuelles (accents finaux vs. initiaux).

En effet, la tendance générale observée pour tous les locuteurs et tous les styles de parole, montre que les accents finaux (AFS, UI-NT et UI-T) sont caractérisés par une Rime plus allongée que l'Attaque, alors que l'inverse est vrai pour les accents initiaux (AI et EMP).

Cependant, seul l'accent final d'unité intonative non terminale (UI-NT) présentant une Rime significativement plus allongée que l'Attaque, fait l'unanimité à travers l'ensemble des locuteurs et des styles (cf. Tableau 9, p. 194). Les autres types d'accents présentent des variations inter-individuelles et stylistiques dans le *degré* de hiérarchisation des composantes syllabiques, même si les tendances générales sont similaires.

Les styles *Lecture* et *Interview* présentent une forte cohérence intra-stylistique dans leurs tendances respectives d'allongement global (cf. Figure 22, p. 196), ainsi qu'à l'intérieur des catégories accentuelles. Les locuteurs du style *Journal*, en revanche, présentent des degrés de hiérarchie des constituants syllabiques assez différents à travers les types d'accents. Ceci se reflète également dans la tendance pour la locutrice *Journal F* à présenter un allongement global similaire aux locuteurs du style *Interview* (Attaques globalement plus longues que Rimes, et tendance à marquer plus fortement les accents initiaux), alors que le locuteur *Journal M* se rapproche davantage des locuteurs du style *Lecture* (Rimes globalement plus longues que Attaques, et tendance à marquer plus fortement les accents finaux).

Cependant, la comparaison des stratégies d'allongement syllabique et infrasyllabique indique que le niveau de constituance A/Rime présente une variabilité intra-stylistique moindre que le niveau syllabique.

Une analyse détaillée des stratégies d'allongement sur les composantes de la Rime (Noyau et Coda) est maintenant proposée, afin de rendre compte d'éventuelles stratégies communes à l'intérieur et à travers les styles de parole.

5.3.4. Niveau de constituance syllabique en A, N et C

5.3.4.1. Tendances générales

Il est intéressant de procéder à une double comparaison en ce qui concerne les stratégies d'allongement des constituants syllabiques : d'une part, en fonction de la *distribution* des catégories accentuelles (accents finaux vs. accents initiaux), et d'autre part, en fonction de la *fonction* des catégories accentuelles (sémantico-pragmatique vs. rythmique).

- *Distribution des catégories accentuelles : Accents finaux vs. initiaux*

L'analyse des constituants de la Rime (Noyau et Coda) révèle des stratégies d'allongement qui distinguent les accents finaux entre eux, et les accents initiaux entre eux.

Globalement, on s'aperçoit que, dans une comparaison interne à la catégorie accentuelle, les *Accents finaux* (AFS, UI-NT et UI-T) dans leur ensemble ont tendance à avoir des Noyaux plus allongés que les autres composantes syllabiques. Cependant, ce résultat n'est significatif que pour la catégorie UI-NT ; l'allongement du Noyau n'est significativement plus long que par rapport à la Coda pour les accents finaux de syntagmes (AFS). Lorsqu'on compare l'allongement différentiel à travers les catégories accentuelles, on remarque que le Noyau des catégories UI-NT est significativement plus long que les Noyaux des AFS et des UI-T pour l'ensemble des locuteurs. Également, les UI-T présentent des Coda (C) plus allongées que les autres composantes, mais nous verrons ultérieurement que cette tendance n'est pas significative pour tous les styles de parole.

Pour ce qui est des *Accents Initiaux* (AI et EMP), on constate qu'à l'intérieur d'une même catégorie accentuelle, l'Attaque (A) est largement plus allongée que le Noyau et la Coda. Les Attaques (A) des accents initiaux emphatiques (EMP) sont comparativement plus allongées que les Attaques des accents initiaux rythmiques (AI), mais la hiérarchie des constituants syllabiques est la même que pour AI. On remarque que l'allongement différentiel touche toutes les composantes syllabiques pour les EMP, mais uniquement les Attaques pour les AI, les Noyaux et les Codas étant globalement non allongés.

- *Fonction des catégories accentuelles : sémantico-pragmatiques vs. rythmiques*

L'analyse du niveau syllabique a permis de mettre en évidence l'allongement plus important des catégories sémantico-pragmatiques par rapport aux catégories rythmiques. L'analyse des constituants ANC révèle des distinctions supplémentaires à l'intérieur des macro-classes.

Si l'on procède à une comparaison globale entre les catégories accentuelles suivant leurs fonctions, on remarque que les accents *sémantico-pragmatiques* se distinguent les uns des autres par la répartition différente de durée sur les constituants syllabiques. En règle générale, les UI-NT privilégient l'allongement du Noyau, alors que les UI-T tendent à avoir une Coda plus longue. Les EMP enfin, présentent une Attaque beaucoup plus importante que les autres constituants syllabiques.

Les accents *rythmiques* présentent également des stratégies d'allongement différentes : l'allongement est réparti de manière similaire sur les constituants syllabiques pour les AFS, alors qu'il est essentiellement porté sur l'Attaque pour les AI, ce dernier ayant un Noyau et

une Coda de durée normalisée nulle.

On voit donc que les similitudes d'allongement global observées au niveau syllabique à l'intérieur des macro-classes accentuelles doivent être relativisées par des différences de stratégies d'allongement différentiel au niveau des ANC.

La Figure 26 donne un premier aperçu des tendances générales observées pour la hiérarchie des constituants syllabiques sur la base des durées normalisées par catégories accentuelles, tous styles de parole confondus.

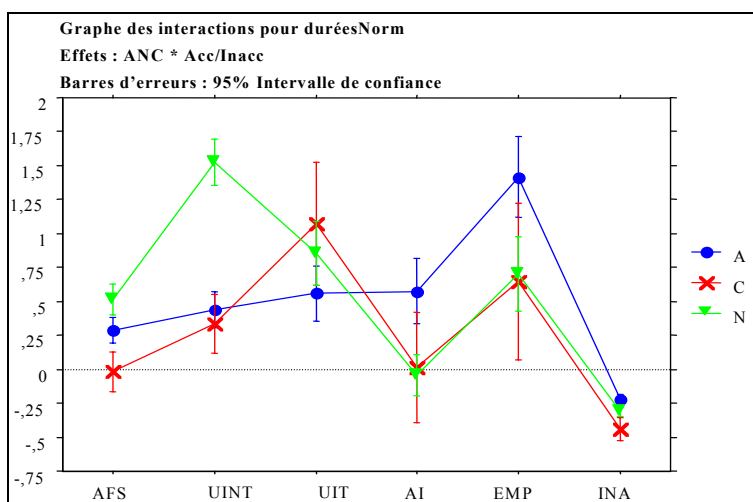


Figure 26 : Tendances générales observées pour la hiérarchie des constituants syllabiques sur la base des durées normalisées par catégorie accentuelle, tous locuteurs et styles de parole confondus. L'axe des abscisses représente les catégories accentuelles ; sur l'axe des ordonnées se trouvent les durées normalisées (Transformée Z) ; (A = Attaque ; N = Noyau ; C = Coda). (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique)

Sur la base des durées différentielles en A/Rime, on a vu que les trois styles pouvaient être distingués facilement, les styles *Interview* et *Lecture* ayant notamment des stratégies globales d'allongement opposées. Une analyse plus détaillée du comportement des éléments de la Rime (Noyau et Coda) pourrait sans doute affiner davantage nos connaissances des stratégies différentes des locuteurs.

Le Tableau 14 donne les résultats de l'Anova générale pour la durée des A, N et C en fonction des catégories accentuelles, pour les 6 locuteurs.

Tableau 14 : Tableau d'Anova générale pour les durées normalisées des A N C, par catégorie accentuelle, pour les 6 locuteurs

Tableau d'ANOVA pour duréesNorm					
	DDL	Somme des carrés	Carré moyen	Valeur de F	Valeur de p
Locuteurs	3	5,239	1,746	2,254	,0800
ANC	1	,223	,223	,288	,5918
Locuteurs * ANC	8	8,481	1,060	1,368	,2050
Catégories Accentuelles	3	595,442	198,481	256,191	<.0001
Locuteurs * Cat. Acc.	23	91,533	3,980	5,137	<.0001
ANC * Cat. Acc.	8	134,204	16,776	21,653	<.0001
Locuteurs * ANC * Cat. Acc.	48	44,190	,921	1,188	,1756
Résidus	5706	4420,649	,775		

Il apparaît que les locuteurs ne se distinguent pas significativement sur la base de la durée globale moyenne de l'allongement différentiel en ANC (valeur de $p = ,0800$ pour l'effet 'Locuteurs'). Par contre, à la différence du niveau de constituance A/Rime, les locuteurs ne sont pas distingués sur la base des stratégies d'allongement différentiel ANC, toutes catégories confondues (valeur de $p = ,2050$). La Figure 27 illustre l'interaction globale Locuteurs * ANC.

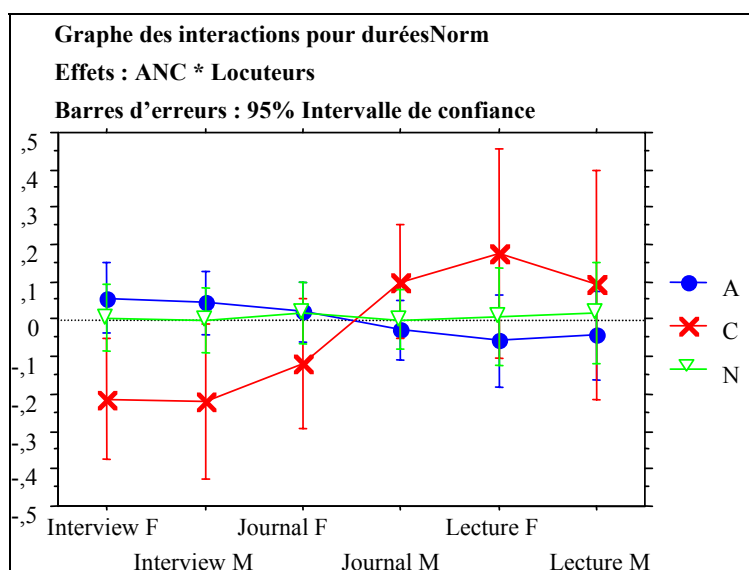


Figure 27 : Durée moyenne des Attaques, Noyaux et Codas (ANC) par locuteur, toutes catégories accentuelles confondues

Il apparaît clairement sur cette figure que l'élément qui distingue le plus les styles de parole entre eux est l'allongement de la Coda. Cette dernière est globalement très allongées en *Lecture*, même plus que les autres composantes syllabiques, alors que c'est exactement le contraire pour le style *Interview*. Ces résultats vont dans le même sens que ceux trouvés pour les constituants A/Rime. Cependant, il est intéressant de noter que les locuteurs ont une stratégie d'allongement comparable du Noyau. Cette tendance explique pourquoi les locuteurs ne se distinguent pas significativement sur la base de l'allongement différentiel global (toutes catégories accentuelles confondues), alors que le niveau de la Rime,

largement influencée par la Coda, distinguait les locuteurs entre eux. Une analyse détaillée des constituants de la Rime s'avérait donc du plus grand intérêt.

On voit que les locuteurs du style *Interview* sont très cohérents au niveau des constituants ANC, comme ils l'étaient pour la Rime. Par contre, l'analyse plus détaillée des constituants révèle que les locuteurs du style *Lecture* ont une stratégie légèrement différente d'allongement différentiel, la locutrice *Lecture F* privilégiant nettement l'allongement de la C par rapport au locuteur *Lecture M*.

Enfin, il apparaît clairement que les locuteurs du style *Journal* se distinguent essentiellement par leur stratégie personnelle d'allongement global de la Coda, le locuteur *Journal M* allongeant sensiblement plus la Coda que la locutrice *Journal F*. La différence observée entre ces deux locuteurs au niveau A/Rime s'explique donc essentiellement par des différences d'allongement des éléments de Coda.

Si les 6 locuteurs ont des stratégies personnelles d'allongement différentiel moyen (i.e. toutes catégories accentuelles confondues), comme on peut le voir sur la Figure 27 (p. 205), l'interaction 'Locuteurs * ANC * Catégories Accentuelles' non significative (valeur de $p = ,1756$) indique que les stratégies d'allongement différentiel par catégories accentuelles est relativement similaire. C'est ce que nous nous proposerons d'analyser dans le détail dans les paragraphes suivants, consacrés à la comparaison inter-locuteurs à l'intérieur des styles.

5.3.4.2. Analyse des différences inter-locuteurs à l'intérieur des styles au niveau des constituants syllabiques A N C

Nous présentons maintenant le détail des résultats par style de parole. Cependant, nous avons été obligé de procéder un peu différemment dans le calcul de l'Anova, par rapport aux analyses de l'allongement syllabique et en A/Rime. En effet, les subdivisions en ANC * catégories accentuelles * locuteurs ($3 * 6 * 2$) sont trop nombreuses et les combinaisons de paramètres trop importantes pour permettre le calcul de l'Anova générale, car nos données ne sont visiblement pas en nombre suffisant⁵. La concaténation des données ANC et Catégories Accentuelles nous permet de pallier ce problème. Par contre, cela ne nous permet pas d'analyser séparément les effets d'interaction entre ANC et Locuteurs, ou Catégories Accentuelles et Locuteurs, ainsi que l'interaction globale Locuteurs * ANC * Catégories Accentuelles. Nous ne pourrions donc procéder qu'à une comparaison partiellement quantifiée des stratégies des locuteurs par style de parole : nous nous contenterons d'une comparaison qualitative des stratégies personnelles des locuteurs à l'intérieur d'un même style de parole. L'information quantitative concernant la similarité ou non des stratégies d'allongement différentiel sur les ANC des deux locuteurs par catégorie accentuelle ne nous est pas fournie.

⁵ Le nombre des éléments syllabiques, divisés en 3 catégories principales (A, N et C) s'élève à 5801, tous locuteurs confondus, soit environ 1000 occurrences par locuteur et 2000 occurrences par style. Le nombre de combinaisons de paramètres, à l'intérieur d'un style de parole (2 locuteurs) s'élève à 36, certaines combinaisons étant de surcroît représentées par très peu d'items (< 5 items, notamment les éléments de Coda). Statview ne permet pas de calculer des Anovas présentant un nombre aussi élevé de combinaisons.

Les tests d'Anova générale et des tests post-hocs par locuteur sont indiqués en annexe. Il faut noter qu'en procédant par locuteur, la concaténation des colonnes ANC et Catégories Accentuelles n'a pas été nécessaire, car les subdivisions des variables indépendantes étaient moindres.

Interview :

Avec les limitations que nous venons d'indiquer ci dessus, nous obtenons le tableau d'Anova générale suivant pour les locuteurs du style *Interview* (Tableau 15) :

Tableau 15 : Tableau d'Anova générale pour les durées normalisées des ANC, par catégorie accentuelle * type de constituant, pour les locuteurs du style *Interview*

Tableau d'ANOVA pour durées Normalisées					
	DDL	Somme des carrés	Carré moyen	Valeur de F	Valeur de p
ANC/Cat.Acc.	16	383,042	23,940	29,255	<.0001
Locuteurs	1	,061	,061	,074	,7851
ANC/Cat. Acc. * Locuteurs	16	39,264	2,454	2,999	<.0001
Résidus	2145	1755,329	,818		

En toute logique, l'ensemble ANC * Catégories Accentuelles est significatif car les catégories accentuelles sont distinguées sur une base fonctionnelle. Nous avons vu d'ailleurs qu'elles sont distinguées à tous les niveaux de constituance syllabique. L'interaction avec l'effet Locuteurs est également significative, indiquant que les locuteurs ont des stratégies personnelles d'allongement différentiel en fonction des catégories accentuelles, alors que l'effet simple Locuteurs est en moyenne fortement non significatif. Nous verrons dans le détail ce qui distingue les locuteurs, lors d'une comparaison qualitative de leurs stratégies d'allongement différentiel respectives.

La Figure 28 présente la tendance d'allongement différentiel pour les 2 locuteurs du style *Interview* confondus.

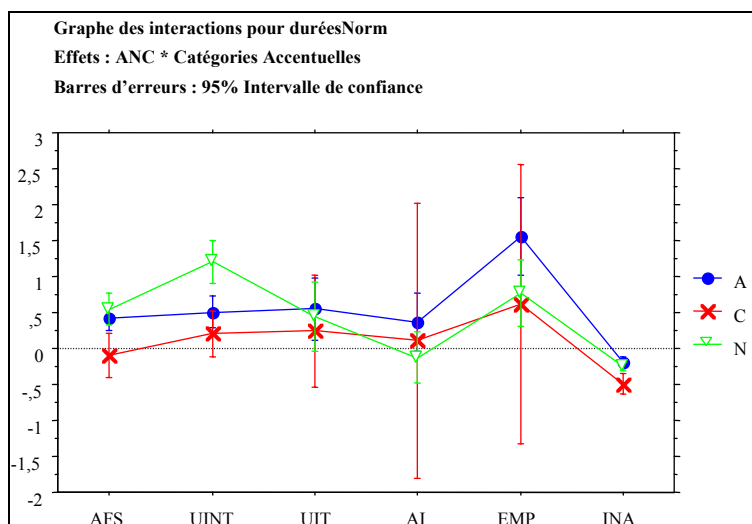


Figure 28 : Durée moyenne des Attaques, Noyaux et Codas (ANC) par catégorie accentuelle, pour les locuteurs du style *Interview*. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique)

Globalement, la tendance est la même que pour l'ensemble des locuteurs (cf. Figure 26, p. 204) : les accents finaux sont marqués par un allongement plus important du Noyau, et les accents initiaux par un allongement plus marqué proportionnellement de l'Attaque. La Coda est légèrement plus longue pour les accents à la frontière des unités intonatives, mais cette tendance n'est pas significative.

En nous reportant au Tableau 9 (p. 194), on s'aperçoit que ce qui distingue essentiellement les deux locuteurs du style *Interview* est la hiérarchie des ANC à l'intérieur de la catégorie UI-NT : la locutrice *Interview F* allonge non seulement considérablement plus chaque constituant syllabique que *Interview M*, mais les 3 éléments syllabiques sont de surcroît significativement distincts les uns des autres, le Noyau étant le plus allongé. En revanche, *Interview M* allonge plus les 3 constituants syllabiques de EMP, ces derniers ne se distinguant pas les uns des autres, alors que *Interview F* allonge significativement l'Attaque par rapport aux Noyau et Coda.

Ce constat nous permet de confirmer la différence intra-stylistique observée au niveau syllabique, par des stratégies d'allongement différentes sur les éléments syllabiques (pour les catégories accentuelles UI-NT et EMP).

Journal :

Nous obtenons le tableau d'Anova générale suivant pour les locuteurs du style *Journal* (Tableau 16) :

Tableau 16 : Tableau d'Anova générale pour les durées normalisées des A N C, par catégorie accentuelle * type de constituant, pour les locuteurs du style Journal

Tableau d'ANOVA pour duréesNormalisées					
	DDL	Somme des carrés	Carré moyen	Valeur de F	Valeur de p
ANC/Cat. Acc.	17	600,966	35,351	46,362	<,0001
Locuteurs	1	1,595	1,595	2,092	,1482
ANC/Cat. Acc. * Locuteurs	17	25,255	1,486	1,948	,0113
Résidus	2589	1974,086	,762		

Comme pour le style *Interview*, nous constatons que l'ensemble ANC/Catégories Accentuelles est significatif. Les locuteurs ne sont pas distingués significativement sur le plan des durées globales, mais la valeur de p de l'interaction indique qu'ils ont des stratégies d'allongement différentiel personnelles. Nous verrons également plus en détail quelles sont les catégories accentuelles, et les constituants syllabiques qui sont à l'origine de cette différence inter-individuelle.

La Figure 29 présente la tendance d'allongement différentiel pour les 2 locuteurs du style *Journal* confondus.

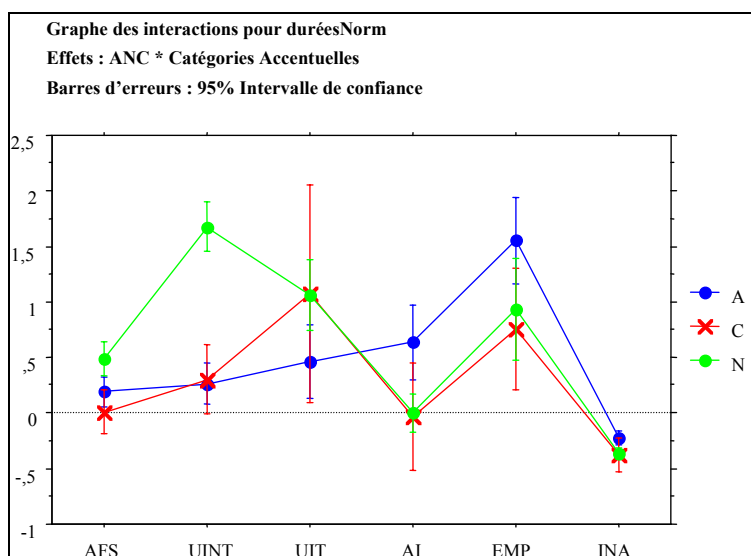


Figure 29 : Durée moyenne des Attaques, Noyaux et Codas (ANC) par catégorie accentuelle, pour les locuteurs du style Journal. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique)

La tendance générale est, comme pour *Interview*, la même que tous styles et locuteurs confondus. Mais les différences sont plus marquées que pour le style *Interview*. L'allongement du Noyau pour les accents finaux est plus important, et l'allongement de la Coda devant une frontière d'unité intonative est également plus marqué, notamment pour les accents à la frontière d'unités intonatives terminales (UI-T). Également, l'allongement de l'Attaque pour les accents initiaux est plus fort que pour le style *Interview*, notamment pour les accents initiaux rythmiques où le contraste avec les autres composantes syllabiques est plus important.

Le locuteur *Journal M* se distingue de *Journal F* surtout par la tendance à allonger significativement et de manière plus importante les éléments Noyaux par rapport aux autres constituants dans les accents finaux (UI-NT, UI-T et AFS). *Journal M* marque également plus l'allongement de l'Attaque pour l'accent initial emphatique EMP, malgré l'allongement concomitant des autres éléments syllabiques.

La différence entre les deux locuteurs se situe cependant plus au niveau global, où l'Attaque est plus allongée pour la locutrice *Journal F* et la Coda plus allongée pour *Journal M*. Cette tendance transparaît davantage au niveau de constituance A/Rime.

Lecture :

Nous obtenons le tableau d'Anova générale suivant pour les locuteurs du style *Lecture* (Tableau 17) :

Tableau 17 : Tableau d'Anova générale pour les durées normalisées des A N C, par catégorie accentuelle * type de constituant, pour les locuteurs du style *Lecture*

Tableau d'ANOVA pour durées Normalisées					
	DDL	Somme des carrés	Carré moyen	Valeur de F	Valeur de p
ANC/Cat.Acc.	16	286,922	17,933	25,463	<.0001
Locuteurs	1	,042	,042	,059	,8076
ANC/Cat. Acc. * Locuteurs	16	27,366	1,710	2,429	,0013
Résidus	968	681,731	,704		

Les mêmes remarques peuvent être faites concernant les résultats de l'Anova générale que pour les styles *Interview* et *Journal*.

La Figure 30 présente la tendance d'allongement différentiel pour les 2 locuteurs du style *Lecture* confondus.

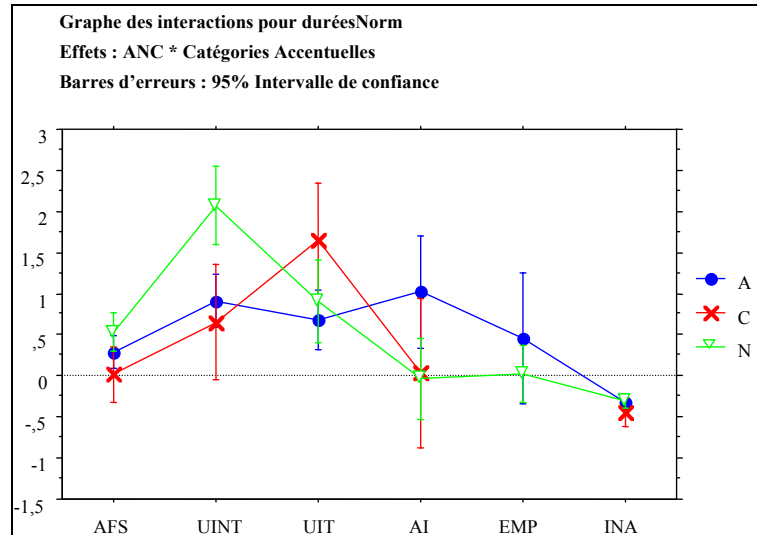


Figure 30 : Durée moyenne des Attaques, Noyaux et Codas (ANC) par catégorie accentuelle, pour les locuteurs du style *Lecture*. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique)

Comme pour les deux autres styles de parole, les tendances générales sont un Noyau globalement plus allongé que les autres composantes syllabiques pour les accents finaux et une Coda allongée plus particulièrement devant les frontières d'unités intonatives terminales (UI-T). Cette dernière tendance n'est cependant significative que pour la locutrice *Lecture F*. La tendance pour les accents initiaux à présenter un allongement de l'Attaque se retrouve également pour le style *Lecture*, et bien que la syllabe portant l'accent emphatique (EMP) soit notablement abrégée par rapport aux autres catégories accentuelles, et par rapport aux autres styles, la hiérarchie des constituants syllabiques est maintenue.

Il est toutefois remarquable que l'Attaque de l'EMP, à durée de Noyau égale avec celle de l'AI, est notablement plus courte que celle de l'AI. Pour les deux locuteurs d'ailleurs, l'Attaque de EMP n'est pas significativement plus allongée. Par contre, l'Attaque de AI est significativement plus allongée pour le locuteur *Lecture M*.

5.3.4.3. Niveau de constituance infrasyllabique en ANC : Conclusions partielles

L'analyse de l'allongement différentiel des constituants syllabiques ANC a révélé une variabilité interlocuteurs et intra-stylistique plus importante. Cette variabilité est essentiellement due à l'allongement des éléments de Coda, qui s'avèrent très sensibles aux changements de catégories accentuelles et de locuteurs. Bartkova avait déjà noté, en 1988, une variabilité inter-locuteurs plus importante pour les consonnes en position pré-pausales. Cette variabilité peut également s'expliquer par le faible nombre d'individus de cet élément syllabique.

Pour les styles *Lecture* et *Interview*, les différences de stratégies d'allongement des ANC sont compensées par des stratégies intra-stylistiques similaires au niveau des A/Rimes. Pour le style *Journal* en revanche, les tendances divergentes des deux locuteurs en ce qui concerne l'allongement global des éléments d'Attaque et de Coda confirment les résultats observés au niveau A/Rime. Le style *Journal* compense par contre les différences intra-stylistiques observées aux niveaux infrasyllabiques par des stratégies d'allongement syllabique global fortement similaires.

Le niveau de constituance ANC nous a cependant permis de rendre compte de tendances d'allongement différentiel en français, déjà observées pour des langues comme l'anglais, le suédois ou le hollandais (Campbell, 1993 et suivantes ; Fant et al., 1991a et b ; Hofhuis, 1993) : un allongement plus important des éléments Noyaux pour les accents finaux, et également des éléments de Codas, notamment pour les accents finaux d'unités intonatives non terminales ; un allongement plus important des éléments d'Attaque pour les accents initiaux emphatiques et rythmiques.

Mais les études réalisées sur l'anglais, le suédois et le hollandais ne portent en général que sur un type de discours particulier (lecture de phrases connectées ou isolées) et sur un nombre limité de locuteurs par langue. Nous avons montré qu'une certaine variabilité dans les stratégies d'allongement différentiel est à imputer aux différents types de processus d'encodage (parole lue vs. interprétée vs. spontanée), mais également aux types de catégories accentuelles dont le modèle de Di Cristo & Hirst (1997) et Di Cristo (sous presse) rend compte pour le français, les deux facteurs (styles et catégories accentuelles) se combinant généralement.

Le nombre de combinaisons des analyses (composantes syllabiques, locuteurs et catégories accentuelles) ne nous permettait pas de présenter l'analyse de l'allongement différentiel en fonction des catégories accentuelles et des 3 styles sur un même graphique. C'est pourquoi nous proposons de procéder à des analyses *par catégories accentuelles*, afin de mettre plus précisément en évidence les types d'accents à l'origine des distinctions inter-stylistiques et intra-stylistiques.

Nous pensons que, pour des applications à la synthèse vocale notamment, il est nécessaire de tenir compte des différences et similitudes de marquage prosodique des énoncés de parole continue produite lors de situations de discours différentes.

5.3.5. Comparaison des stratégies individuelles et stylistiques par catégories accentuelles au niveau infrasyllabique des ANC

Nous avons vu que les distinctions entre locuteurs au niveau syllabique semblaient être dues essentiellement aux catégories accentuelles sémantico-pragmatiques (UI-NT, UI-T et EMP) (cf. Figure 17, p. 187). Nous proposons d'analyser les stratégies d'allongement différentiels par locuteur et par style de parole, en insistant sur les deux macro-classes accentuelles (sémantico-pragmatiques vs. rythmiques). Les résultats pour chaque catégorie accentuelle sont analysés par une Anova, pour laquelle les locuteurs (ou les styles) et les constituants syllabiques (ANC) sont les variables indépendantes, et la durée constitue la variable dépendante. Par souci de clarté, nous ne présentons dans le texte que les tableaux et figures d'Anova par style de parole (cf. note de bas de page n°4, p. 184).

Les commentaires de détail sur la significativité des oppositions 2 à 2 entre les constituants syllabiques (ANC) renvoient au Tableau 9 p. 194 qui récapitule les tendances d'allongement différentiel par locuteur et par style de parole pour chaque catégorie accentuelle.

5.3.5.1. Macro-classe rythmique : accent final de syntagme (AFS) et accent initial secondaire (AI)

5.3.5.1.1. Accent final de syntagme (AFS)

Le Tableau 18 donne les résultats de l'Anova générale pour la catégorie AFS.

Tableau 18 : Tableau d'Anova générale pour les durées normalisées des ANC pour la catégorie accentuelle AFS, pour les trois styles

Tableau d'ANOVA pour durées Normalisées					
	DDL	Somme des carrés	Carré moyen	Valeur de F	Valeur de p
Styles	2	,718	,359	,318	,7275
ANC	2	33,816	16,908	14,995	<,0001
Styles * ANC	4	3,278	,820	,727	,5737
Résidus	1024	1154,627	1,128		

Les styles ne se distinguent pas en termes de durée globale pour cette catégorie. Les ANC sont significativement différentes les unes des autres. Par contre, l'interaction entre les deux variables indépendantes n'est pas significative, ce qui indique que les styles ont des stratégies d'allongement différentiel pour cette catégorie fortement similaires.

La Figure 31 présente la tendance d'allongement différentiel pour la catégorie AFS et pour les 3 styles de parole.

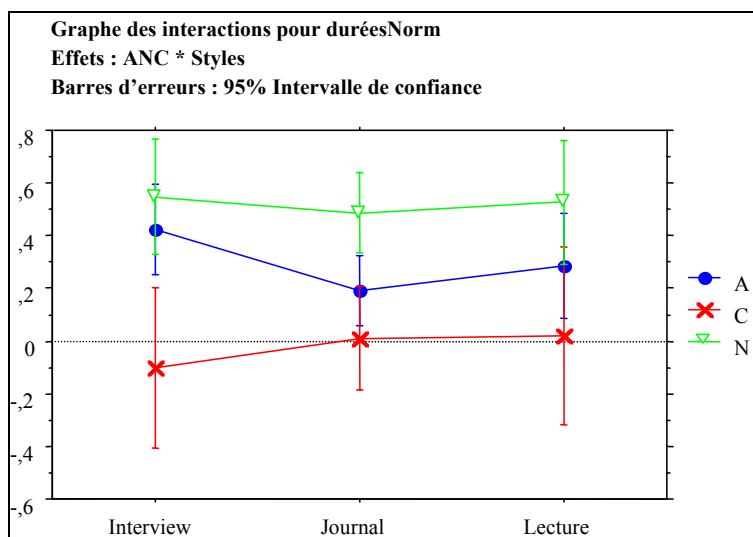


Figure 31 : Durée moyenne des Attaques, Noyaux et Codas (ANC) pour la catégorie accentuelle AFS, et pour les 3 styles de parole

Ces résultats sont confortés par le comportement de chaque locuteur pris indépendamment. Les résultats de l'Anova générale montrent en effet que le facteur Locuteurs ainsi que l'interaction entre les ANC et les locuteurs sont non significatives (respectivement, $F_{(5, 1015)} = ,682$; $p = ,6375$; et $F_{(10, 1015)} = ,467$; $p = ,9116$), suggérant que tous les locuteurs ont des stratégies d'allongement différentiel similaires.

Globalement, les Noyaux sont plus longs que les Attaques : cette différence n'est cependant significative que pour le locuteur *Journal M*. Les Noyaux sont également largement plus longs que les Codas, ce qui est significatif pour tous les locuteurs. Enfin, la Coda est plus courte que l'Attaque, mais cette différence n'est significative que pour les locuteurs du style *Interview* et la locutrice *Journal F*.

En résumé, la tendance pour le Noyau à être significativement plus long que la Coda est significative pour les trois styles.

Ces résultats sont quelque peu en contradiction avec les résultats de Campbell, qui trouvait un facteur d'allongement identique sur les différentes composantes syllabiques, pour les accents internes de syntagme. Or, nos résultats suggèrent que le Noyau subit un allongement, même réduit, plus important que les autres composantes syllabiques. Il se peut que cette différence soit due à la proximité d'une frontière mineure pour ce type d'accent en français : l'allongement des éléments de la Rime semble être en effet propre aux accents situés à des frontières prosodiques.

5.3.5.1.2. Accent initial secondaire (AI)

Le Tableau 19 donne les résultats de l'Anova générale pour la catégorie AI et pour les 3 styles.

Tableau 19 : Tableau d'Anova générale pour les durées normalisées des A N C pour la catégorie accentuelle AI, pour les trois styles

Tableau d'ANOVA pour duréesNormalisées					
	DDL	Somme des carrés	Carré moyen	Valeur de F	Valeur de p
Styles	2	,730	,365	,381	,6836
ANC	2	16,737	8,368	8,738	,0002
Styles * ANC	4	1,672	,418	,437	,7821
Résidus	208	199,205	,958		

Les trois styles de parole ne se distinguent pas significativement sur la base de la durée globale. Les ANC sont significativement distinguées, mais l'interaction entre styles et ANC est non significative.

La Figure 32 présente les tendances d'allongement différentiel pour la catégorie AI et pour les 3 styles de parole.

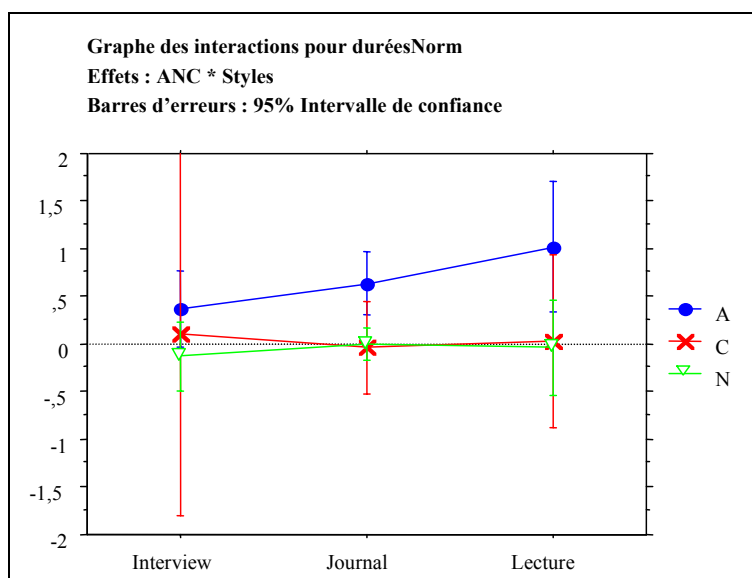


Figure 32 : Durée moyenne des Attaques, Noyaux et Codas (ANC) pour la catégorie accentuelle AI, et pour les 3 styles de parole

Ces résultats sont confortés par l'Anova générale opposant les 6 locuteurs, où l'interaction entre les effets 'locuteurs' et 'ANC' est non significative ($F_{(9, 200)} = ,549$; $p = ,8374$).

La tendance générale est à un allongement plus important de l'élément d'Attaque par rapport aux autres composantes syllabiques, ces dernière ayant une valeur de Z en moyenne nulle. Cette tendance est cependant non significative pour les 2 locuteurs du style *Interview*. Pour le style *Journal* en revanche, l'Attaque est significativement supérieure au Noyau pour les deux locuteurs, et également significativement supérieure à la Coda pour la locutrice *Journal F*. Pour le style *Lecture* par contre, si les deux locuteurs présentent la même hiérarchie, l'Attaque n'est pas significativement supérieure aux Noyau et Coda pour la

locutrice *Lecture F*, alors qu'elle l'est pour le locuteur *Lecture M*.

Cependant, au niveau des styles, les trois styles présentent la même hiérarchie, pour laquelle l'Attaque est significativement plus allongée que les autres composantes syllabiques.

5.3.5.2. Macro-classe sémantico-pragmatique : accent final d'unité intonative non terminale (UI-NT), accent final d'unité intonative terminale (UI-T) et accent initial emphatique (EMP)

5.3.5.2.1. Accent final d'unité intonative non terminale (UI-NT)

Le Tableau 20 donne les résultats de l'Anova générale pour la catégorie UI-NT et pour les 3 styles.

Tableau 20 : Tableau d'Anova générale pour les durées normalisées des A N C pour la catégorie accentuelle UI-NT, pour les trois styles

Tableau d'ANOVA pour durées Normalisées					
	DDL	Somme des carrés	Carré moyen	Valeur de F	Valeur de p
Styles	2	15,630	7,815	6,792	,0012
ANC	2	108,444	54,222	47,122	<,0001
Styles * ANC	4	10,525	2,631	2,287	,0592
Résidus	476	547,718	1,151		

Les styles sont distingués significativement sur la base des durées globales. Les ANC sont fortement distinctes. Par contre, l'interaction entre les deux variables indépendantes n'est pas significative, ce qui laisse penser que les styles de parole ont des stratégies d'allongement différentiel relativement semblables, qui compensent leurs différences au niveau global.

Des tests post hocs montrent que le style *Lecture* se différencie significativement des deux autres styles de parole, ces deux derniers n'étant pas distincts globalement (valeur de $p = ,4198$). Le style *Lecture* allonge en effet considérablement plus ce type de catégorie accentuelle.

La Figure 33 présente les tendances d'allongement différentiel pour la catégorie UI-NT et les 3 styles de parole.

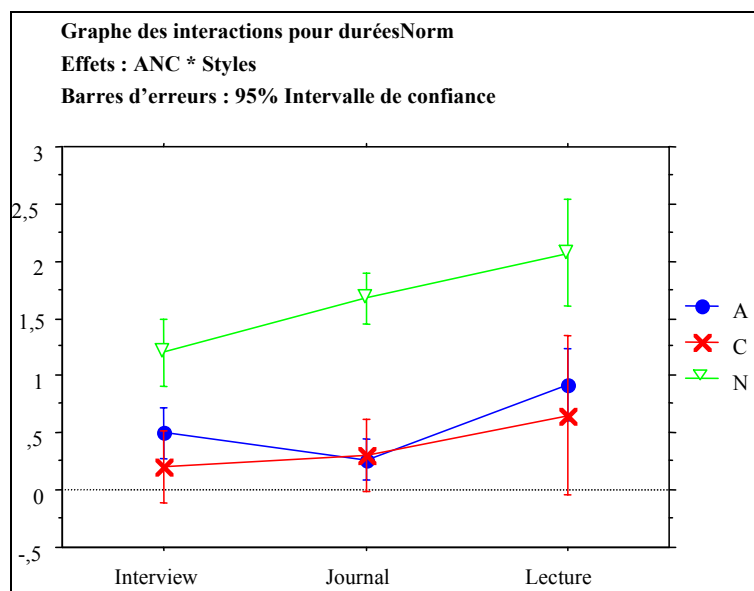


Figure 33 : Durée moyenne des Attaques, Noyaux et Codas (ANC) pour la catégorie accentuelle UI-NT, et pour les 3 styles de parole

Les tests d'Anova générale confrontant les différents locuteurs ne confirment pas les résultats dégagés pour les styles de parole. Ainsi, l'effet 'locuteurs' est significatif ($F_{(5, 467)} = 8,746$; $p = <,0001$), et l'interaction entre les locuteurs et les ANC est également significative ($F_{(10, 467)} = 2,425$; $p = ,0080$). De plus, les tests post hoc montrent que les locuteurs d'un même style de parole présentent des différences significatives : pour le style *Interview*, les 2 locuteurs sont significativement différents (valeur de $p = <,0001$) ; pour le style *Journal*, c'est également le cas (valeur de $p = ,0200$), ainsi que pour le style *Lecture* (valeur de $p = ,0005$).

L'analyse graphique des durées des ANC révèle que ces différences ne sont pas à imputer aux mêmes composantes syllabiques suivant le style : ainsi, pour le style *Interview*, il semblerait que la locutrice *Interview F* allonge beaucoup plus les éléments de Noyau et d'Attaque que le locuteur ; pour le style *Journal*, la différence se porterait essentiellement sur le Noyau, nettement plus long pour la locutrice *Journal F* ; pour *Lecture* enfin, la différence entre les locuteurs ne semble tenir qu'à la Coda, extrêmement plus longue pour *Lecture M* que pour *Lecture F*.

Cependant, ces différences sont compensées par des stratégies assez comparables des locuteurs, i.e. que les proportions entre les composantes syllabiques sont relativement semblables. En règle générale, les Noyaux sont plus allongés que les Attaques (significatif pour tous les locuteurs) et plus allongés que les Codas (significatif pour tous les locuteurs, sauf *Lecture M*). Enfin, les Attaques et les Codas ne se distinguent pas sur la base de leurs durées, sauf pour les locutrices *Interview F* et *Lecture F*, où les Codas sont sensiblement inférieures aux Attaques.

Au niveau des styles pourtant, il existe une certaine homogénéité, les Noyaux étant

significativement plus longs que les Attaques et les Codas pour les trois styles de parole.

On voit donc que cette catégorie accentuelle est particulièrement peu homogène à travers les locuteurs, y compris (et même surtout) à l'intérieur d'un même style. On pourrait rapprocher cela des observations de Vaissière (1983) selon laquelle le rôle de la f0 et de la durée varie en fonction du locuteur dans les tâches de segmentation de la parole en unités prosodiques. La fonction sémantico-pragmatique de la catégorie UI-NT la rend sans doute particulièrement sensible aux stratégies individuelles de marquage prosodique. Une comparaison avec les tendances individuelles de marquage prosodique de cette catégorie accentuelle par le paramètre de f0 serait intéressante, et permettrait éventuellement de mettre en évidence des stratégies de compensation entre la durée et la f0.

5.3.5.2.2. *Accent final d'unité intonative terminale (UI-T)*

Le Tableau 21 donne les résultats de l'Anova générale pour la catégorie UI-T et pour les 3 styles.

Tableau 21 : Tableau d'Anova générale pour les durées normalisées des ANC pour la catégorie accentuelle UI-T, pour les trois styles

Tableau d'ANOVA pour durées Normalisées					
	DDL	Somme des carrés	Carré moyen	Valeur de F	Valeur de p
Styles	2	11,060	5,530	4,981	,0080
ANC	2	4,764	2,382	2,146	,1204
Styles * ANC	4	9,125	2,281	2,055	,0893
Résidus	157	174,285	1,110		

Les styles sont significativement différents les uns des autres. Les ANC ne sont pas distinguées, et l'interaction entre les styles et les ANC n'est pas significative.

Des tests post hoc montrent que le style *Lecture* se distingue significativement du style *Interview*, ce dernier allongeant globalement beaucoup moins cette catégorie accentuelle.

La Figure 34 présente les tendances d'allongement différentiel pour la catégorie UI-T et les 3 styles de parole.

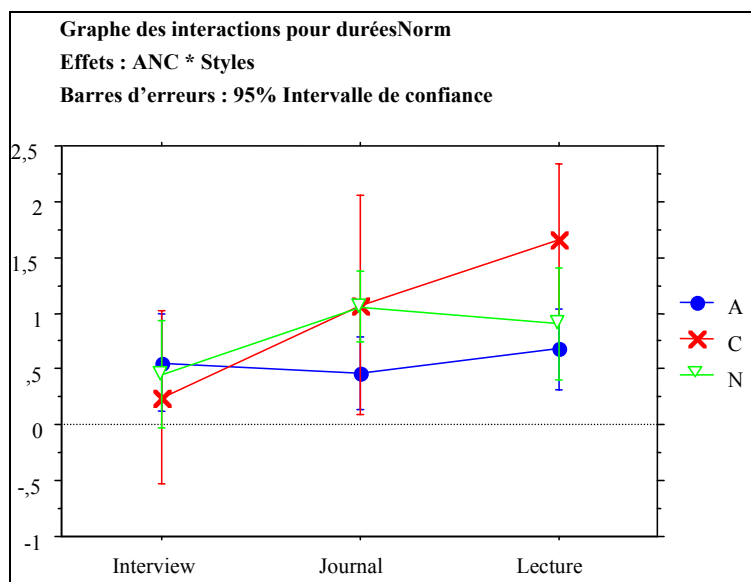


Figure 34 : Durée moyenne des Attaques, Noyaux et Codas (ANC) pour la catégorie accentuelle UI-T, et pour les 3 styles de parole

Les résultats pour les styles de parole sont en partie confirmés par l'Anova générale confrontant les 6 locuteurs, mais l'interaction entre les locuteurs et les ANC est non significative ($F_{(10, 148)} = 1,027$; valeur de $p = ,4233$). Les tests post hocs indiquent que les locuteurs à l'intérieur de chaque style ne se distinguent pas, sauf les locuteurs du style *Lecture* (valeur de $p = ,0153$).

Globalement, la Coda est plus allongée que le Noyau et l'Attaque. Cette tendance n'est pas significative pour les locuteurs du style *Interview*, qui présentent un allongement identique des composants syllabiques. Pour le style *Journal*, la Coda est significativement supérieure à l'Attaque pour les deux locuteurs, mais elle n'est significativement supérieure au Noyau que pour le locuteur *Journal M*. Enfin, pour *Lecture*, seule la locutrice *Lecture F* présente des Coda significativement plus allongées que le Noyau et l'Attaque.

Il apparaît donc globalement que la Coda joue un rôle important pour cette catégorie accentuelle pour le style *Lecture* où elle est significativement plus longue que l'Attaque et le Noyau, alors qu'il s'agit plutôt de l'ensemble Noyau et Coda pour le style *Journal*. Pour le style *Interview* enfin, l'allongement est réparti de manière homogène sur les trois constituants syllabiques.

5.3.5.2.3. Accent initial emphatique (EMP)

Le Tableau 22 donne les résultats de l'Anova générale pour la catégorie EMP et pour les 3 styles.

Tableau 22 : Tableau d'Anova générale pour les durées normalisées des A N C pour la catégorie accentuelle EMP, pour les trois styles

Tableau d'ANOVA pour duréesNormalisées					
	DDL	Somme des carrés	Carré moyen	Valeur de F	Valeur de p
Styles	2	11,877	5,939	4,202	,0167
ANC	2	14,014	7,007	4,958	,0082
Styles * ANC	4	1,370	,343	,242	,9139
Résidus	152	214,807	1,413		

Les styles sont significativement différents, ce qui montre qu'ils ont des durées globales différentes. Les ANC sont également significativement distinguées. Cependant, l'interaction entre les styles de parole et les constituants syllabiques est fortement non significative, indiquant que les 3 styles ont des stratégies d'allongement différentiel similaires.

Des tests post hoc montrent que le style *Lecture* se distingue significativement des deux autres styles, ces derniers n'étant pas distincts sur la base des durées globales (valeur de $p = ,8663$). Cette différence tient au fait que le style *Lecture* allonge considérablement moins cette catégorie accentuelle.

La Figure 35 présente les tendances d'allongement différentiel pour la catégorie EMP et les 3 styles de parole. Nous rappelons que nous avons éliminé du graphique l'élément de Coda de cette catégorie accentuelle pour le style *Lecture*, car la population de cette combinaison est inférieure à 5 items.

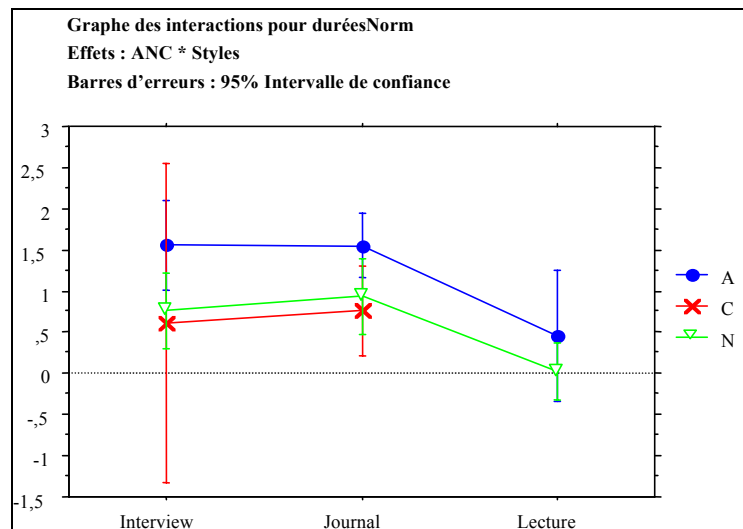


Figure 35 : Durée moyenne des Attaques, Noyaux et Codas (ANC) pour la catégorie accentuelle EMP, et pour les 3 styles de parole

L'Anova générale confrontant les 6 locuteurs confirme ces résultats, l'interaction entre les ANC et les locuteurs étant non significative ($F_{(9, 144)} = ,206$; $p = ,9932$). Des tests post hoc montrent que les locuteurs à l'intérieur d'un même style ne se distinguent pas significativement sur la base des durées des ANC, sauf les locuteurs du style *Interview*

(valeur de $p = ,0101$).

Globalement, l'allongement est plus important sur l'élément d'Attaque que sur les autres composantes syllabiques. Dans le style *Interview*, cette différence est significative pour la locutrice *Interview F*, mais pas pour le locuteur *Interview M*, ce dernier allongeant également fortement les Noyaux et Codas. L'Attaque n'est pas significativement plus allongée pour les 2 locuteurs du style *Lecture*, même s'ils respectent la même hiérarchie que les autres locuteurs. Enfin, les deux locuteurs du style *Journal* présentent une Attaque significativement plus allongée que les 2 autres composantes syllabiques.

Au niveau des styles, les styles *Interview* et *Journal* présentent un allongement significatif de l'Attaque sur les autres composantes, alors qu'il n'en est rien pour le style *Lecture*.

5.3.5.3. Syllabes inaccentuées (INA)

Le Tableau 23 donne les résultats de l'Anova générale pour la catégorie EMP et pour les 3 styles.

Tableau 23 : Tableau d'Anova générale pour les durées normalisées des A N C pour la catégorie accentuelle INA, pour les trois styles

Tableau d'ANOVA pour duréesNormalisées					
	DDL	Somme des carrés	Carré moyen	Valeur de F	Valeur de p
Styles	2	,571	,285	,480	,6188
ANC	2	9,980	4,990	8,392	,0002
Styles * ANC	4	4,931	1,233	2,073	,0817
Résidus	3741	2224,523	,595		

L'Anova générale indique que les locuteurs ne se distinguent pas globalement, et qu'ils ont des stratégies d'allongement différentiel semblables (valeur p de l'interaction).

Des tests post hocs indiquent que le style *Interview* se distingue significativement du style *Lecture* (valeur de $p = ,0326$). Globalement en effet, le style *Interview* présente des durées plus longues que les autres styles, le style *Lecture* ayant les durées de syllabes inaccentuées les plus courtes.

La Figure 36 illustre les tendances d'allongement différentiel pour la catégorie INA et les 3 styles de parole.

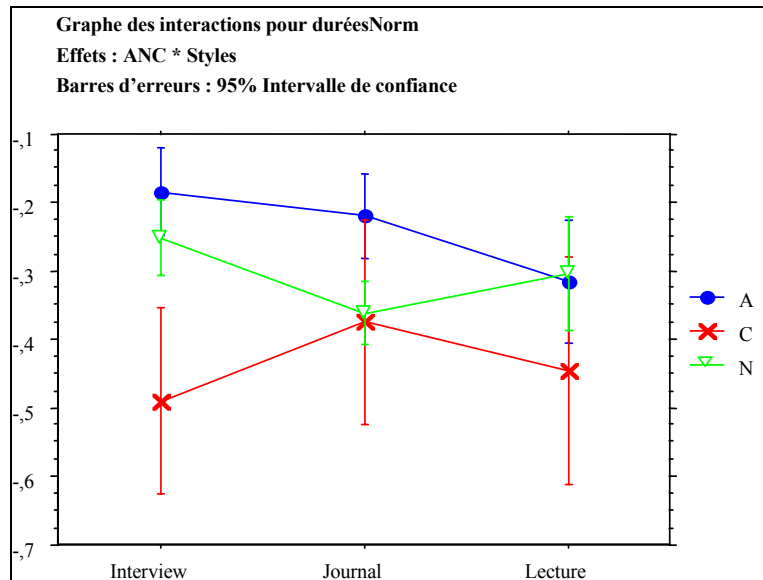


Figure 36 : Durée moyenne des Attaques, Noyaux et Codas (ANC) pour la catégorie accentuelle INA, et pour les 3 styles de parole

Les tendances par style sont corroborées par l'Anova pour les locuteurs, l'interaction entre les locuteurs et les ANC étant non significative ($F_{(10, 3732)} = 1,056$; $p = ,3933$). Les tests post hoc montrent que, si les locuteurs des styles *Journal* et *Lecture* sont homogènes à l'intérieur de leur style respectif, les locuteurs du style *Interview* sont significativement différents (valeur de $p = ,0231$) ; le locuteur *Interview M* présente en effet globalement des durées plus importantes que la locutrice *Interview F*.

Globalement, l'Attaque est plus longue que les autres composantes syllabiques. Pour le style *Interview*, cette tendance est significative pour les deux locuteurs, mais uniquement par rapport à la Coda. Pour le style *Journal*, cette tendance est significative pour les deux locuteurs en ce qui concerne le Noyau, mais seule la locutrice *Journal F* présente une Attaque significativement plus longue que la Coda. Pour les deux locuteurs du style *Lecture* en revanche, la durée (et plus particulièrement le raccourcissement) est réparti de manière égale sur les différentes composantes syllabiques.

En résumé, les styles *Journal* et *Interview* présentent un allongement plus important de l'Attaque par rapport aux autres composantes syllabiques, alors que le raccourcissement est réparti de la même manière sur les éléments syllabiques pour le style *Lecture*.

5.3.5.4. Conclusion partielle de l'étude des allongements différentiels des ANC par catégories accentuelles

L'analyse des stratégies individuelles et inter-stylistiques par catégorie accentuelle nous a permis de mettre en évidence les catégories accentuelles les plus sensibles au facteur locuteurs.

Globalement, les catégories accentuelles rythmiques (AFS et AI) sont les plus robustes aux différents processus d'encodage, puisqu'elles présentent respectivement les mêmes stratégies d'allongement différentiel, quel que soit le locuteur ou le style de parole. Par contre, les catégories sémantico-pragmatiques (UI-NT, UI-T et EMP) sont extrêmement sensibles aux locuteurs et peu robustes.

Si on considère les stratégies *par style* de parole, il apparaît que le style *Lecture* se différencie significativement des styles *Interview* et *Journal*, sur la base des stratégies d'allongement global et différentiel des catégories sémantico-pragmatiques. La *Lecture* se particularise par un allongement syllabique beaucoup plus important des accents finaux d'unités intonatives terminales et non terminales (UI-NT et UI-T), et par un raccourcissement notable de l'accent initial emphatique (EMP).

- Pour la catégorie UI-NT cependant, les trois styles de parole conservent une hiérarchie des composantes syllabiques similaires, le Noyau étant significativement plus allongé que les autres éléments syllabiques.
- Pour la catégorie UI-T, la Coda joue un rôle important dans la distinction inter-stylistique : sur une échelle d'allongement de la Coda, les styles *Lecture* et *Interview* se trouvent aux extrémités opposées, le style *Journal* étant intermédiaire, mais plus proche du style *Lecture* (cf. Figure 34, p.219).

Il semblerait donc que la catégorie UI-T soit plus sensible aux divers processus d'encodage que la catégorie UI-NT, plus précisément au niveau infrasyllabique. Cependant, cette dernière manifeste une variabilité *intra-stylistique* très importante, notamment pour les styles *Interview* et *Journal*, où les locutrices allongent plus le Noyau que les locuteurs dans leur style respectif, et pour le style *Lecture* où le locuteur allonge plus la coda que la locutrice. Les locuteurs du style *Lecture* sont également non cohérents pour la catégorie UI-T, où la locutrice allonge plus la Coda que le locuteur. Il semble donc que, pour le style *Lecture*, les différences de stratégies d'allongement des UI-NT et UI-T observées au niveau syllabique global, s'expliquent essentiellement par des stratégies opposées d'allongement de la Coda. Nous sommes cependant consciente que ce résultat doit être tempéré par le faible nombre d'occurrences des éléments de Coda en français.

Pour la catégorie UI-T s'exprime donc une distinction inter-stylistique : *Lecture* favorise l'allongement de la Coda, alors que l'allongement concerne plutôt le couple Noyau et Coda pour *Journal* ; enfin, l'allongement est distribué de manière homogène pour *Interview*.

Dans la catégorie UI-NT s'expriment plutôt des différences intra-stylistiques, qui sont estompées au niveau stylistique global, où l'on observe la même tendance d'allongement supérieur de l'élément Noyau.

- Pour la catégorie EMP enfin, *Journal* et *Interview* ont un allongement global similaire (allongement substantiel) alors que *Lecture* se caractérise par un raccourcissement de EMP. Les trois styles ont cependant la même hiérarchie interne, l'Attaque étant plus allongée que le Noyau (non significatif cependant pour *Lecture*).

Le Tableau 24 résume les stratégies d'allongement globales et différentielles par style de parole, pour chaque catégorie accentuelle. Les limitations rencontrées lors de l'analyse détaillée par locuteur sont indiquées par des numéros qui renvoient à des notes au bas du tableau.

Tableau 24 : Tableau récapitulatif des stratégies d'allongement global et d'allongement différentiel par style de parole, en fonction des catégories accentuelles (S = significatif ; N.S. = non significatif) : seuls les résultats statistiquement significatifs sont indiqués dans les tendances d'allongement global et différentiel. (Lec. = Lecture ; Int. = Interview ; Jour. = Journal).

Catégories Accentuelles	Facteur Styles	Facteur ANC	Interaction ANC * Styles	Tendances d'allongement global	Tendances d'allongement différentiel
AFS	N.S.	S.	N.S.	Int. = Jour. = Lec.	$N > C$ ¹
UI-NT	S.	S.	N.S. ²	Lec. > (Int. = Jour.)	$A < N > C$ ³
UI-T	S.	N.S.	N.S. ⁴	Lec. > Int.	⁵
AI	N.S.	S.	N.S.	Int. = Jour. = Lec.	$A > (N = C)$
EMP	S.	S.	N.S.	Lec. < (Int. = Jour.)	$A > (N = C)$ ⁶
INA	N.S.	S.	N.S.	Int. > Lec.	⁷

¹ le Noyau est également plus long que l'Attaque pour le locuteur Journal M
² l'interaction 'ANC * locuteurs' est significative ; des tests post hocs montrent que chaque couple de locuteurs à l'intérieur d'un même style est significativement différent.
³ le locuteur Lecture M ne présente pas d'allongement du Noyau significativement supérieur à la Coda, cette dernière étant beaucoup plus allongée que les Codas des autres locuteurs.
⁴ seuls les locuteurs du style Lecture se distinguent significativement l'un de l'autre. Les locuteurs des styles Interview et Journal sont comparables dans leur style respectif.
⁵ chaque style a sa propre stratégie : Lecture allonge significativement plus la Coda [$C > (A = N)$] ; Journal allonge significativement l'ensemble Noyau + Coda [$(N = C) > A$] ; Interview n'allonge pas significativement une composante syllabique par rapport aux autres.
⁶ ce résultat n'est pas significatif pour le style Lecture, ni pour le locuteur Interview M qui allonge considérablement les autres composantes syllabiques également. Pour ces trois locuteurs néanmoins, la hiérarchie est respectée (Attaque plus longue que Noyau et Coda).
⁷ pour le style Interview, l'Attaque est supérieure à la Coda uniquement. Pour Journal, l'Attaque est significativement plus allongée que le Noyau, alors que pour Lecture, l'allongement (ou plutôt le raccourcissement) est réparti de manière égale sur les constituants syllabiques

6. DISCUSSION ET CONCLUSION GENERALE SUR LES TENDANCES D'ALLONGEMENT DIFFERENTIEL EN FONCTION DES CONTRAINTES PROSODIQUES ET DES STYLES DE PAROLE

Les résultats développés dans les paragraphes précédents nous ont permis de dégager un bon nombre de tendances dans la gestion des phénomènes de durées, et de répondre aux trois questions initiales de notre plan expérimental.

Distinction fonctionnelle des catégories accentuelles :

Tout d'abord, d'un point de vue méthodologique, la distinction fonctionnelle que nous avons postulée entre les catégories accentuelles semble pertinente, ce qui confirme par là même l'existence d'un système accentuel en français bien plus complexe que ce que les descriptions traditionnelles le laissaient supposer.

Au niveau syllabique global en effet, la durée permet de distinguer significativement entre les deux macro-classes accentuelles fonctionnelles : les syllabes accentuées qui relèvent des catégories sémantico-pragmatiques (accent emphatique initial EMP ; accent final d'unité intonative terminale UI-T, et non terminale UI-NT) sont significativement plus allongées que les syllabes accentuées relevant des catégories accentuelles rythmiques (accents initial et final de mot ou de syntagme, AI et AFS). Nous trouvons cependant deux limitations à ce constat :

- Pour le style *Lecture*, les deux types d'accents initiaux (AI rythmique et EMP sémantico-pragmatique) ne sont pas significativement distingués sur la base de leur durée. Cela pourrait expliquer pourquoi ces deux types d'accents ont souvent été confondus dans la littérature, puisque la plupart des descriptions et des modèles s'appuient sur la parole de laboratoire, et plus généralement sur des données contrôlées. Le besoin de modèles adaptatifs se fait donc sentir, qui prendraient en considération la réalité de la parole 'en situation', i.e. de la parole dans diverses situations de communication incluant différents degrés de types de discours spontanés.
- Pour le style *Interview*, les accents finaux d'unités intonatives terminales (UI-T) ne se distinguent pas significativement des accents finaux de syntagmes (AFS), aussi bien au niveau syllabique global qu'au niveau de la hiérarchie des constituants syllabiques. L'UI-T a tendance à être raccourcie en parole spontanée.

En conclusion, nos résultats ont validé la distinction que nous avons effectuée a priori entre les divers types d'accents en français. Nous rappelons quelles sont les bases théoriques de cette distinction : les mots lexicaux seraient marqués par un accent initial et un accent final à un niveau sous-jacent de la représentation phonologique. Ces accents sont actualisés au niveau de surface en fonction de règles d'eurythmie et de groupement, ainsi que de règles pragmatiques qui rendent compte de l'assignation des accents nucléaires et emphatiques (Di Cristo, sous presse). Nos résultats confirment la tendance à un marquage moins fort des accents initiaux et finaux rythmiques, dont le domaine est le mot ou le groupe de mot, comparativement aux accents initiaux et finaux de niveau supérieur, dont le domaine est l'énoncé.

Nous avons également conforté un résultat majeur présenté dans une communication antérieure (Astésano et al., 1995b), concernant l'existence d'un accent initial rythmique en français. Ce type d'accent a longtemps été ignoré, ou confondu avec l'accent 'didactique' ou emphatique. De plus, les descriptions qui en ont été faites insistaient sur le caractère purement mélodique de cet accent (Rossi, 1985b ; Fónagy, 1980, notamment). Or, nos résultats suggèrent que cet accent initial rythmique est également caractérisé par un allongement, porté essentiellement sur l'élément d'Attaque, qui le distingue significativement des syllabes inaccentuées.

Tendances d'allongement différentiel en français : " est-ce que l'allongement se répartit de manière différente sur les constituants syllabiques en fonction des catégories accentuelles? "

Nous avons mis en évidence des stratégies d'allongement différentiel pour le français, comparables à celles trouvées dans la littérature pour l'anglais, le suédois et le hollandais notamment (Campbell, 1992 et suivantes ; Fant et al., 1991a et b ; Hofhuis, 1993). La réponse à notre interrogation 'l'allongement se répartit-il différemment sur les constituants syllabiques en fonction des catégories accentuelles ?' (cf. p. 182) est donc affirmative. Ainsi, les catégories que nous avons pré-définies se distinguent sur le plan distributionnel par une tendance à l'allongement des Attaques pour les accents initiaux (AI et EMP), comparable à la tendance observée pour les accents lexicaux de l'anglais ('stress'), et par une tendance à l'allongement des éléments constitutifs de la Rime (Noyau et Coda) pour les accents finaux, également comparable aux résultats trouvés dans les autres langues pour les accents situés à des frontières de constituants prosodiques, particulièrement à des frontières d'unités intonatives. Cependant, Fant et al. (1991a et 1991b) remarquent que les accents mineurs en français (les accents finaux de syntagme, AFS) sont caractérisés par un allongement de l'Attaque. Nous n'avons pas trouvé les mêmes tendances dans notre corpus, l'élément Noyau étant toujours plus allongé, même si cette tendance est beaucoup moins marquée que pour les accents finaux à des frontières d'unités intonatives.

Campbell remarque également que l'allongement final est un cas particulier en anglais. Les segments dans les syllabes finales de phrases sont allongées différemment : les Attaques sont significativement plus courtes que les Noyaux, mais, alors que la Coda est légèrement plus longue en moyenne, elle n'est pas significativement plus longue que le Noyau. Nos résultats pour le français montrent que ce constat est à relativiser en fonction du type de frontière (mineure, majeure terminale ou non terminale) et en fonction du style de parole.

Pour ce qui est des accents initiaux, Séguinot (1976) avait déjà remarqué pour les accents emphatiques en français, cette tendance à l'allongement des éléments d'Attaque. Selon lui, la tenue de la consonne (ou du groupe consonantique) est responsable de l'accroissement de durée sous l'emphase. Nous avons remarqué quant à nous que l'allongement de durée frappe également les autres constituants syllabiques, même si l'Attaque est plus fortement marquée.

Enfin, Vaissière (1983) postule que l'allongement des éléments finaux semble avoir une fonction essentiellement démarcative (de frontière à droite) alors que l'allongement des consonnes initiales de mot semble avoir une fonction de marqueur à gauche des frontières (plus précisément des frontières lexicales). Elle précise que ce dernier type d'allongement

est plus largement utilisé pour marquer l'emphase. Or, il apparaît, pour le français, que l'allongement relatif des éléments initiaux de mot n'est pas uniquement réservé au marquage de l'emphase, mais s'étend d'une manière générale aux accents initiaux (rythmiques et emphatiques). Nous pensons qu'il s'agit avant tout d'un phénomène de démarcation à gauche des unités lexicales, combiné tantôt à une fonction rythmique, tantôt à une fonction sémantico-pragmatique (emphase), les deux fonctions n'étant pas mutuellement exclusives : la fonction sémantico-pragmatique peut en effet prendre appui sur un accent rythmique sous-jacent, et représenter en ce cas la version 'Hyper' de l'accent initial sous-jacent, comme le propose Di Cristo (sous presse). Nos résultats d'allongement différentiel pour ces deux accents initiaux vont d'ailleurs dans ce sens : les deux accents ont une 'figure' globale d'allongement infra-syllabique similaire (allongement de l'Attaque), cette tendance étant quasiment deux fois plus importante pour EMP.

En conclusion, l'hypothèse forte de Campbell & Isard (1991), selon laquelle l'allongement est distribué de manière homogène sur les segments syllabiques en fonction de leur élasticité propre, ne s'applique pas à nos données. Par contre, nos résultats indiquent que l'hypothèse faible est plus appropriée pour rendre compte des stratégies d'allongement dans l'énoncé. En effet, l'allongement s'applique sur les segments en fonction de leur place dans la syllabe (Attaque, Noyau ou Coda) et de la place de la syllabe dans l'énoncé (finale ou non finale).

Cependant, nos résultats indiquent également qu'à l'intérieur d'une même classe distributionnelle, l'allongement différentiel rend compte également de distinctions *fonctionnelles*, comme nous le montrons ci-dessous.

Plusieurs niveaux de constituance syllabique rendent compte de plusieurs types de distinctions accentuelles :

Les différents niveaux de constituance que nous avons analysés (niveau syllabique, niveaux infrasyllabiques en A/Rime et en ANC) entretiennent bien entendu des relations d'interdépendance. Ainsi, l'allongement des constituants syllabiques résulte en l'allongement syllabique global. Cependant, une certaine indépendance de ces différents niveaux existe. Deux catégories accentuelles peuvent avoir la même durée syllabique, mais des hiérarchies d'allongement des constituants syllabiques différentes. Nous avons vu que, à deux exceptions près, les catégories accentuelles sémantico-pragmatiques ne se distinguent pas sur la base de leurs durées syllabiques, et les catégories rythmiques ont également des allongements globaux similaires. Le niveau syllabique rend donc compte de la distinction fonctionnelle entre les catégories accentuelles. Le niveau des A/Rimes permet quant à lui une distinction entre les catégories accentuelles sur la base de leur distribution (initiale vs. finale), les accents finaux ayant des Rimes plus allongées que les Attaques, les accents initiaux présentant la tendance inverse.

Le niveau des ANC présente l'intérêt de permettre la distinction des catégories accentuelles à la fois sur le plan de leur distribution et de leur fonction. Comme au niveau des A/Rimes, les catégories sont distinguées sur le plan distributionnel, malgré une variabilité intra-stylistique plus importante au niveau des ANC. Sur le plan fonctionnel, les catégories sémantico-pragmatiques ont des tendances d'allongement plus marquées que les accents rythmiques. Enfin, le niveau des ANC permet également de distinguer les catégories à l'intérieur d'une même macro-classe accentuelle. Ainsi, les accents finaux

d'unités intonatives se distinguent globalement par la gestion différente des éléments de Coda. Nous rappelons qu'il faut néanmoins être prudent en ce qui concerne cette dernière tendance, eu égard au nombre d'occurrences limité de ce type de constituant syllabique (structure syllabique majoritairement CV en français).

Se pose alors la question de savoir quel niveau de constituance syllabique est le plus représentatif des stratégies d'allongement en fonction du contexte prosodique. Mais nous avons vu que les stratégies d'allongement sont également fonction des styles de parole, voire des locuteurs.

Variabilité interindividuelle et variabilité stylistique : " les styles de parole ont-ils des stratégies d'allongement spécifiques en fonction des catégories accentuelles? "

Notre plan expérimental a été conçu pour répondre à la question de l'influence des styles de parole sur les stratégies d'allongement différentiel. Au vu de nos résultats, il apparaît que la réponse est positive : les styles de parole présentent bien des stratégies d'allongement spécifiques et ce aux deux niveaux de constituance (syllabique et infra-syllabique).

Une originalité de notre travail est en effet de mettre en évidence une autre dimension de variabilité, qui n'a pas été prise en compte dans les travaux précédents sur les stratégies d'allongement du suédois, de l'anglais ou du hollandais. Nous pensons en effet que la variabilité induite par les divers styles de parole sur les phénomènes de durées devrait être prise en considération dans les tentatives d'implémentation des paramètres de durée dans les systèmes de synthèse, par exemple. De plus, la comparaison de plusieurs styles de parole permet de mettre en évidence les stratégies transversales à divers locuteurs et types de discours, permettant de distinguer entre les invariants représentatifs et constitutifs du système noyau de la langue et la variabilité imputable aux styles de parole. On peut, à partir de cette distinction, choisir d'implémenter sur une même base invariable, les caractéristiques prosodiques représentatives de tel ou tel type de discours.

Nos résultats ont montré que les trois styles de parole sont distingués par leurs stratégies d'allongement syllabique et infrasyllabique. Au niveau des constituants infrasyllabiques, le style *Lecture* est caractérisé globalement par un allongement plus important de la Coda, et marque de manière optimale les accents finaux au niveau syllabique. A l'opposé, le style *Interview* favorise l'allongement des éléments d'Attaque et marque plus distinctement les accents initiaux. Comment expliquer cette opposition radicale dans les stratégies d'allongement? Nous pensons que cette distinction stylistique est directement liée à la différence de processus d'encodage, i.e. à la tâche de production. Guaitella (1990) suggère que le style *Lecture*, de par la possibilité d'anticipation visuelle des éléments de parole, permet la programmation des événements prosodiques en fonction du contour intonatif de fin de phrase ou d'énoncé (principe de sélection à droite), alors que la parole spontanée, par essence non pré-programmée, structure ces événements prosodiques par rapport au début de l'énoncé (principe de sélection à gauche) (cf. § «Principales caractéristiques phonétiques et prosodiques distinguant les styles de parole », p. 117). Par analogie avec les phases de tension articulaire participant de la caractérisation rythmique des langues (Wenk, 1983), on pourrait penser que *Lecture* et *Interview* sont caractérisés par des 'phases de tension prosodique' inverses, le premier ayant une tension prosodique maximale en fin de syntagme (anticipation visuelle), le second en début de syntagme (adéquation aux phases de

conceptualisation). Ces différences de tension prosodique auraient pour conséquence directe une réduction articulatoire et acoustique des consonnes en position de Coda en parole spontanée (Duez, 1998).

Le style *Journal*, en revanche, occupe une position intermédiaire aux deux autres styles. Alors qu'au niveau syllabique il ne se distingue pas du style *Interview* (*Lecture* étant globalement distingué des deux autres styles), une analyse plus détaillée au niveau des constituants ANC révèle que le style *Journal* combine les deux stratégies, en marquant clairement à la fois les proéminences initiales et les proéminences finales.

La Figure 37 présente les stratégies globales d'allongement différentiel (toutes catégories accentuelles confondues) par style de parole, et révèle cette tendance.

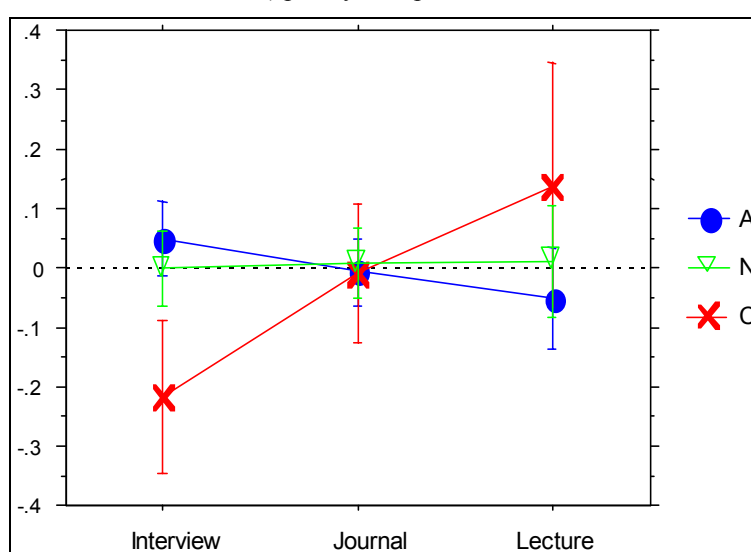


Figure 37: Allongement différentiel global (toutes catégories accentuelles confondues) pour les 3 styles de parole. (A = Attaque ; N = Noyau ; C = Coda).

Notre hypothèse initiale est donc également vérifiée : nous avons en effet choisi le style *Journal* comme élément de comparaison entre la parole spontanée et la parole lue, car il emprunte à la fois aux caractéristiques d'encodage du style *Lecture* et du style *Interview* (cf. § « Intérêt méthodologique de la gradation des types de discours sur un axe spontané/non spontané », p. 139). Les locuteurs du style *Journal* présentent des stratégies globales opposées sur le plan de l'allongement général (toutes catégories accentuelles confondues). Il serait fort intéressant d'élargir notre corpus afin de vérifier quelle tendance ce style de parole privilégie. Il se pourrait également que la distinction a priori entre les trois styles de parole doive être révisée a posteriori par une classification en deux styles de parole principaux : parole spontanée vs. parole lue.

Niveaux de constituance syllabique et variabilité : " Quels niveaux de constituance (syllabique ou infrasyllabique) est particulièrement sensible aux variations stylistiques? "

La réponse à notre interrogation 'Quel niveau de constituance est particulièrement sensible aux variations stylistiques (et individuelles) ?' n'est pas simple. Les deux niveaux de constituance (syllabique et infrasyllabique) sont concernés par la variabilité induite par les styles de parole : les catégories sémantico-pragmatiques sont particulièrement sensibles et peu robustes aux stratégies des locuteurs et des styles de parole, comparativement aux catégories accentuelles rythmiques, et ce aussi bien au niveau syllabique qu'infrasyllabique.

Ces différences de robustesse des macro-classes accentuelles sont directement liées aux fonctions des types d'accent. Les accents sémantico-pragmatiques correspondent à un choix de la part du locuteur à mettre en valeur certains mots, dans le cadre d'une focalisation large (accents nucléaires d'unités intonatives marquant le niveau de l'énoncé) ou d'une focalisation étroite (accents emphatiques marquant le niveau lexical). Il semble que ces choix pragmatiques se concrétisent au niveau acoustico-phonétique par des stratégies personnelles de marquage prosodique. On verra également que cette variabilité sur le plan de la gestion des durées se retrouve également en partie sur le plan de la gestion de la f0 (cf. § « Analyse de Variance », p. 246).

Cependant, une certaine compensation s'opère entre les différents niveaux de constituance, par laquelle une variabilité induite par les styles à un niveau est compensée par des stratégies communes aux styles de parole à un autre niveau. Ainsi, si les styles de parole se distinguent sur le plan de l'allongement global des catégories accentuelles, ces différences sont estompées au niveau des stratégies d'allongement différentiel, car les styles de parole présentent des organisations hiérarchiques entre constituants syllabiques similaires à l'intérieur des catégories accentuelles. L'exemple le plus frappant concerne l'accent emphatique EMP du style *Lecture*, considérablement raccourci par rapport aux autres styles de parole, mais qui garde néanmoins la caractéristique de supériorité hiérarchique de l'Attaque sur les autres éléments. Enfin, les éventuelles différences de stratégies individuelles observées au niveau des constituants syllabiques ANC sont parfois compensées au niveau de constituance supérieur (A/Rime), notamment dans le cas des accents finaux d'unités intonatives. Les questions demeurent cependant ouvertes concernant ce phénomène apparent de compensation : quelle est la motivation de cette compensation ? Est-elle motivée linguistiquement ? Est-elle le résultat de la recherche inconsciente d'une adéquation à un système noyau ? La variabilité à un niveau de constituance n'est-elle que la conséquence de contraintes de bas niveaux ? Si c'est le cas, la compensation au niveau de constituance syllabique supérieur serait bien le reflet d'une recherche d'adéquation aux contraintes phonologiques du système noyau. Or, nos résultats évoquent la tendance contraire, i.e. la compensation par une stratégie commune au niveau inférieur de constituance. La question demeure donc ouverte.

En ce qui concerne la variabilité intra-stylistique, les stratégies de compensation ne sont pas les mêmes. Le style *Journal* est plus homogène dans la gestion des allongements au niveau syllabique, alors qu'à l'inverse, le style *Interview* compense les différences interindividuelles au niveau des syllabes par une stratégie commune au niveau des A/Rimes. Le style *Lecture* est quant à lui plus cohérent à tous les niveaux de constituance, sauf pour les accents finaux d'unités intonatives.

Pour autant, aucun niveau de constituance ne peut être privilégié dans notre quête de la recherche d'invariant. Il apparaît notamment que le niveau des A/Rimes exprime davantage l'invariance intra-stylistique, alors que le niveau des ANC exprime plutôt un certain degré d'invariance inter-locuteurs et inter-stylistique. Les arguments en faveur du niveau intermédiaire de la Rime concernent la cohérence des stratégies d'allongement de ces constituants immédiats (Noyau et Coda). Campbell (1992) remarque notamment que les éléments de Coda semblent être plus homogènes à travers les locuteurs que les éléments d'Attaque, et cela l'incite à considérer la Rime comme un constituant plus adéquat pour rendre compte des stratégies d'allongement. Or, nos résultats sur le français montrent que la part de variabilité inter- et intra-stylistique en fonction des catégories accentuelles, mais également sur le plan global de l'allongement différentiel, est à imputer principalement à la Coda. De fait, au niveau de la Rime, où le comportement erratique de la Coda influe sur l'ensemble Noyau + Coda, la variabilité interindividuelle est plus importante qu'au niveau des ANC, dans lequel le Noyau a un comportement très similaire à travers les locuteurs. Cette tendance se retrouve dans les valeurs de p des Anova générales, où l'interaction entre les effets 'locuteurs * ANC * Catégories accentuelles' a une valeur de $p = ,1779$, alors que l'interaction entre 'locuteurs * A/Rimes * Catégories accentuelles' a une valeur de $p = ,0008$.

Dans une perspective d'implémentation à la synthèse vocale, il serait donc préférable d'utiliser le niveau des ANC, sachant que l'allongement du Noyau par catégorie accentuelle est très homogène à travers les locuteurs. Les poids relatifs des Attaques et Codas dépendrait globalement du style de parole à synthétiser, la parole spontanée marquant davantage les Attaques et la parole lue, les Codas. Une plus grande part de variabilité devrait également être prévue dans le comportement des éléments de Coda à travers les catégories accentuelles.

Il reste cependant à démontrer que ces tendances d'allongement différentiel, qu'elles soient le reflet du système noyau (invariants) ou le reflet d'une stratégie stylistique (variabilité), sont pertinentes sur le plan perceptif.

Pour conclure, nos résultats confirment la pertinence du modèle de Campbell, qui postule l'ajustement des durées segmentales dans un cadre syllabique. En effet, nous avons montré que l'allongement différentiel s'applique également au français, différenciant essentiellement entre les prééminences initiales et finales. Cependant, nous avons également montré que le niveau syllabique global rend compte de manière optimale de la distinction fonctionnelle des prééminences du système accentuel du français. Nous avons par là même confirmé la pertinence de la syllabe en tant qu'interface entre l'information segmentale et l'organisation temporelle (timing) de la parole. Il semblerait donc que la syllabe, ainsi que l'ajustement temporel des unités infrasyllabiques en fonction du contexte prosodique, soient les unités pertinentes de modélisation du timing de la parole. L'unité de programmation rythmique proposée par Barbosa & Bailly (1994), le GIPC, ne rendrait pas compte de la diversité des phénomènes accentuels en français, notamment en ce qui concerne les réalisations des accents initiaux rythmiques et emphatiques (AI et EMP), pour lesquels l'allongement se porte sur le début de la syllabe. Ce type d'unité (GIPC) est sans doute mieux adapté à la parole lue pour laquelle, nous l'avons vu, ce type d'accent ne présente pas d'allongement substantiel (Hirst et al., 1998).

Enfin, nous formulons une réserve sur le postulat de Campbell (sous presse), selon lequel les syllabes, plus robustes que les segments au changement de locuteurs, constitueraient une unité de 'normalisation' inter-locuteurs. L'analyse des stratégies d'allongement de divers types de discours nous a en effet conduite à considérer la nécessité de l'étude combinée des niveaux syllabique et infrasyllabique, afin de pouvoir pleinement rendre compte de la complexité du système accentuel en français, ainsi que des phénomènes d'invariance et de variabilité de l'organisation accentuelle du français (Astésano, 1998a).

7. RESUME

Nous analysons l'effet des contraintes prosodiques et stylistiques sur la durée des segments dans le cadre de la syllabe. Pour cela, nous utilisons le modèle de Campbell (1992 et suivantes) qui permet de rendre compte de l'allongement différentiel des segments en fonction de leur élasticité propre dans un cadre syllabique. Nous analysons l'allongement induit par les phénomènes accentuels à différents niveaux de constituance : syllabique, infrasyllabique en Attaque + Rime, et infrasyllabique en Attaque + Noyau + Coda. Nous émettons en effet l'hypothèse qu'il existe une certaine variabilité induite par les processus d'encodage différents, mais qu'une part d'invariabilité doit exister quel que soit le style de parole, comme reflet du système noyau du rythme du français. La question est de savoir s'il existe un niveau privilégié qui rend compte de ce système noyau, et, si variabilité il y a, à quel niveau elle se situe.

Nos résultats mettent en évidence plusieurs tendances :

- la distinction *fonctionnelle* des catégories accentuelles est pertinente. On observe en effet au niveau syllabique global un allongement plus important des catégories sémantico-pragmatiques par rapport aux catégories accentuelles rythmiques.
- les niveaux d'analyse infrasyllabiques permettent de mettre en évidence la distinction *distributionnelle* des catégories accentuelles, les accents initiaux présentant une Attaque plus allongée que les autres constituants syllabiques, et les accents finaux présentant un allongement plus conséquent des constituants Noyau et Coda (Rime). Ces résultats vont dans le sens des observations faites sur d'autres langues (Fant et al., 1991 a et b ; Campbell, 1992 et suivantes, etc.).
- les résultats confortent notamment l'existence d'un accent initial rythmique (AI) distinct de l'accent initial emphatique (EMP) (similitude dans la 'figure' globale de l'allongement infrasyllabique, mais différence notable des durées syllabiques).
- les styles de parole présentent des stratégies d'allongement différentiel spécifiques, à tous les niveaux d'observation. La *Lecture* marque plus l'allongement des accents finaux d'Unités Intonatives, alors que le style *Interview* marque comparativement plus les accents initiaux. Le style *Journal* distingue de manière optimale les proéminences rythmiques des proéminences sémantico-pragmatiques, à la fois au niveau syllabique global et au niveau de l'allongement différentiel. Globalement, le style *Lecture* favorise l'allongement des Codas, alors que le style *Interview* privilégie l'allongement des Attaques. Le style *Journal* est intermédiaire. Les différentes stratégies stylistiques sur le plan de la durée doivent selon nous être prises en compte dans les systèmes de synthèse de la parole.

- *CHAPITRE 6* -

Configurations Tonales

1. INTRODUCTION

Cette étude se situe dans la suite logique d'une caractérisation acoustico-phonétique des catégories accentuelles fonctionnelles du modèle de Di Cristo & Hirst (1997) et Di Cristo (sous presse). Cependant, à la différence de l'analyse des durées segmentales, qui visait à mettre en évidence la part potentielle d'invariabilité et la part de variabilité imputable aux différents styles de parole, nous avons centré notre analyse des configurations tonales sur la recherche de l'invariance. Notre corpus multistyle se justifie ici dans la recherche d'invariants tonals, susceptibles de rendre compte du système noyau du rythme du français, transversal à tout style de parole.

Nous nous situons ainsi dans la perspective selon laquelle les paramètres locaux de f_0 (variations locales de f_0) refléteraient l'information structurale et sémantico-pragmatique du message indépendamment du style de parole, alors que les paramètres globaux (étendue de la f_0 , notamment) seraient révélateurs des différences de processus d'encodage (Hirschberg, sous presse).

Il est connu que la f_0 joue un rôle prépondérant dans la perception des prééminences (Terken, 1993). Beckman (1993) postule cependant que si les paramètres tonals permettent la délimitation des groupements prosodiques en français, il n'y a pas de réelles contraintes relationnelles entre les mouvements de f_0 en français et les syllabes prééminentes. Ce postulat va, une fois de plus, à l'encontre de la reconnaissance d'un système accentuel en français. S'il est vrai que le français est caractérisé par une accentuation assez 'floue' (Fónagy, 1980), nous nous proposons de montrer que les catégories accentuelles définies dans le cadre du modèle de Di Cristo & Hirst (1997) peuvent être distinguées sur la base des paramètres tonals.

Dans la lignée des travaux de Terken (1993) et t'Hart et al. (1990), nous fondons notre analyse sur les facteurs dynamiques de la f_0 , eu égard au caractère continu et dynamique des variations de f_0 . Nous choisissons par conséquent de prendre en compte les propriétés temporelles de la f_0 telles que les pentes, l'empan syllabique et le timing des configurations, afin de rendre compte précisément de la spécificité de l'organisation tonale en français. Cette analyse est complémentaire aux études précédentes de l'organisation tonale du français en termes de tons discrets (Hirst et al., sous presse).

Notre plan expérimental est conçu pour répondre à deux objectifs :

- nous proposons d'une part un examen des caractéristiques phonétiques des catégories accentuelles prédéfinies en termes de configurations tonales locales ;
- nous voulons déterminer d'autre part dans quelle mesure la combinaison des différentes propriétés tonales retenues permet de prédire l'appartenance d'une prééminence particulière à une catégorie accentuelle donnée.

Avant de passer à l'exposé de notre méthodologie et à l'analyse de nos résultats, il nous semble nécessaire de faire le point sur une notion centrale, relative à l'organisation tonale des prééminences : le débat sur l'analyse en configurations vs. en niveaux tonals. Cette mise au point nous permettra de situer plus précisément l'enjeu et le cadre de notre étude.

1.1. Le débat sur l'analyse en configurations vs. l'analyse en niveaux tonals

Ce débat a donné lieu à un certain nombre de publications et de controverses (Bolinger, 1951 ; Ladd, 1996, notamment). Nous nous proposons de présenter brièvement les concepts que sous-tendent l'analyse en niveaux tonals (discrets) et l'analyse en configurations (dynamiques). Pour un compte-rendu détaillé des controverses liées à la phonologie intonative, nous renvoyons le lecteur à l'article de fond de Ladd (sous presse).

Les tenants de l'analyse de l'intonation en termes de séquences de segments tonals se situent dans le courant de la phonologie métrique auto-segmentale. La f_0 serait la manifestation d'une structure surimposée et indépendante du texte, dans laquelle les éléments tonals sont associés avec les éléments d'un texte de manière à refléter les relations de proéminence. Un pic de f_0 n'est donc pas considéré comme une propriété phonétique de la syllabe proéminente, mais comme un élément de la structure phonologique de l'énoncé, au même niveau que la syllabe proéminente elle-même. Dans cette perspective, il semble donc évident de rendre compte de l'association entre les pics de f_0 et la syllabe, et plus précisément de son alignement temporel avec la syllabe. L'alignement est défini comme la propriété phonétique du timing relatif des événements dans le contour de f_0 avec les événements dans la chaîne segmentale ; l'association par contre, est la propriété structurelle abstraite de cette relation (Ladd, 1992).

Bruce (1977) montre que ce qui est important dans la perception et la distinction de l'accent en suédois est l'alignement précis d'un point cible dans le temps. Le timing et la pente de la chute suivante ou la montée précédente importent peu.

Or, aux dires mêmes des tenants de la phonologie autosegmentale, la distance entre les points cibles n'est pas constante dans un même contexte prosodique et varie considérablement à travers les locuteurs (Arvaniti et al., 1998). L'alignement et le positionnement des points cibles seraient de surcroît contraints par le contenu segmental (Bruce, 1980). Dans ces conditions, on peut se demander si les informations de pentes, intégrant les distances temporelles et fréquentielles entre les points cibles, ne seraient pas plus pertinentes que la seule information de niveaux de tons pour la caractérisation de l'organisation tonale. Outre ce constat, les adversaires de l'analyse en points cibles fondent leur argumentation sur la base des caractéristiques physiques continues de la f_0 .

Dès 1951 en effet, Bolinger propose une critique de la notion de niveaux de hauteur proposée par Trager & Smith (1951). Ces auteurs postulent l'existence de 4 niveaux de tons relatifs, sans définir explicitement la notion de *relativité*. La principale critique de Bolinger tient à ce que deux contours intonatifs peuvent être décrits par une même séquence de tons (séquences allophoniques) alors qu'en termes de configurations et de sens, les deux séquences sont différentes. La notion même de configuration, défendue par Bolinger, s'inscrit dans le cadre de la reconnaissance des formes prônée par la théorie de la Gestalt. Selon Bolinger, une configuration peut se trouver accrue ou diminuée dans le registre du locuteur, elle sera toujours perçue comme la même configuration. Sa 'taille' et sa 'position' dans le registre ne l'affectent pas matériellement car la reconnaissance se fait sur sa forme, sur l'analogie entre les patrons. Bien que le registre (ou étendue tonale) semble avoir une fonction linguistique (Rietveld & Gussenhoven, 1985 ; 1997), il ne jouerait qu'un rôle secondaire par rapport aux patrons intonatifs. Bolinger propose une analogie avec la musique, où les relations entre 'tons' peuvent être modifiées (amplitude, hauteur relative

etc.) sans que la reconnaissance du patron soit affectée.

En rapport avec la notion de registre également, la phonologie autosegmentale prévoit de calculer la distance entre les maxima de f_0 et une ligne de base virtuelle pour estimer la magnitude des variations de f_0 et pour rendre compte des relations entre niveaux tonaux relatifs (Pierrehumbert, sous presse, notamment). Si l'auditeur semble traiter l'information de f_0 sur la base des pics et des creux de f_0 dans l'énoncé, l'utilisation de la ligne de base virtuelle dans ce processus serait remise en cause en parole 'naturelle', pour laquelle elle n'a pas de manifestation concrète (Terken, 1997). Dans cette perspective, la notion même de niveaux est remise en cause, et nous incite à rejoindre Bolinger (1951) selon lequel la reconnaissance des mouvements de f_0 (et par suite leur pertinence linguistique) se ferait plus en termes de 'patrons', qu'en termes de niveaux de hauteur discrets.

Pour les tenants d'une approche perceptive des patrons de f_0 en effet, les propriétés physiques dynamiques de la parole sont à la base de l'interprétation linguistique (t'Hart et al., 1990). Ils s'opposent en cela aux tenants de la phonologie autosegmentale pour lesquels les mouvements locaux de f_0 ne sont pas des primitives de l'analyse linguistique : ils ne seraient que des transitions ou des interpolations entre les points cibles tonaux (Arvaniti et al., 1998 ; Hirst et al., sous presse). Pour t'Hart et al. (1990), l'analyse phonétique de l'intonation en niveaux est une sur-simplification. Si cette tentative de réduction phonologique des données est intéressante, son application au niveau phonétique va à l'encontre des faits phonétiques de changements de hauteur en production et en perception. En effet, la f_0 est une propriété dynamique et continue, et l'intégration perceptive se fait sur la base de ces caractéristiques dynamiques. Il est donc plus logique de la décrire en termes de contours que de points discrets. Les mouvements de montée et descente sont des mouvements de base, qui doivent être décrits en fonction de propriétés temporelles, comme la direction, le timing et l'empan syllabique du mouvement tonal. Les configurations décrites en ces termes représentent les réalisations a priori en nombre illimité des contours de base en nombre fini.

Les deux approches partagent néanmoins le point de vue selon lequel l'auditeur ne serait sensible qu'aux changements de f_0 produits intentionnellement par le locuteur. Ce présupposé valide d'ailleurs notre démarche auditive et notre classification en catégories accentuelles fonctionnelles. Cependant, il reste à démontrer si la pertinence linguistique de ces variations locales de f_0 repose sur une représentation discrète ou continue de la f_0 . En choisissant de centrer notre étude sur les propriétés dynamiques de la f_0 , nous n'avons pas la prétention de répondre à cette question. Nous prenons néanmoins le parti de présupposer de la pertinence des paramètres dynamiques de f_0 , à l'instar de Bolinger (1951) et dans la lignée de la théorie de la Gestalt.

2. CARACTERISATION DES CONFIGURATIONS TONALES DES PROEMINENCES EN FRANÇAIS ⁶

Nous nous proposons d'analyser les propriétés tonales des catégories accentuelles fonctionnelles définies dans le cadre du modèle de Di Cristo & Hirst (1997), et Di Cristo (sous presse). Nous avons précédemment exposé notre volonté d'inscrire notre recherche dans la lignée de l'école Hollandaise, pour laquelle la notion de mouvements tonals est centrale. Ce parti-pris correspond également au postulat des psycholinguistes, selon lesquels le temps est structurateur et intégrateur des phénomènes accentuels (Fraisse, 1956 ; 1967 ; 1974).

Notre étude est donc fondée sur l'analyse des paramètres dynamiques de la f_0 , et plus précisément sur les configurations tonales des proéminences.

2.1. Procédure expérimentale

Pour cette étude, la fréquence fondamentale a été modélisée à l'aide d'une fonction spline quadratique utilisant un algorithme de modélisation (Momel), et a été corrigée manuellement. Cette correction concerne les erreurs flagrantes de modélisation (points cibles manquants ou redondants) et a été effectuée dans environ 8% des cas (pour une discussion sur les erreurs de l'algorithme Momel, voir Astésano et al., 1997a). La f_0 modélisée est représentée par une séquence de points cibles <ms ; Hz> correspondant aux variations locales pertinentes (Hirst & Espesser, 1993). Précisons que l'algorithme de modélisation de la f_0 Momel est ici détourné de sa fonction originelle. En effet, nous ne l'utilisons pas comme base d'une analyse phonologique de surface en termes de niveaux de tons, mais simplement comme une aide à l'automatisation des relevés des configurations tonales.

Les configurations tonales sont définies par une série de 3 points cibles. Le point cible le plus proche de l'attaque vocalique de la syllabe proéminente constitue le centre de la configuration tonale, dont les frontières sont délimitées par les points cibles gauche et droit adjacents au point cible central. Les configurations tonales dont les frontières droites ou gauches s'étendaient au delà d'une pause silencieuse ont été éliminées.

Sur la base de ces configurations, les 8 paramètres suivants ont été extraits (cf. Figure 38) : l'empan syllabique des configurations tonales, les distances entre les 3 points cibles en termes de ms et de valeurs ERB (Hermes & Van Gestel, 1991), les pentes droites et gauches (ERB/ms), et la distance du point cible central à l'attaque vocalique de la syllabe proéminente.

⁶ Cette étude a été réalisée avec la collaboration de Pascale Nicolas, Robert Espesser et Emmanuel Flachaire. Ce travail a donné lieu à deux communications, l'une lors du Workshop sur l'intonation à Athènes en Septembre 1997 (Astésano et al., 1997) et l'autre lors des Journées d'Etude sur la Parole à Martigny en Juin 1998 (Astésano et al., 1998).

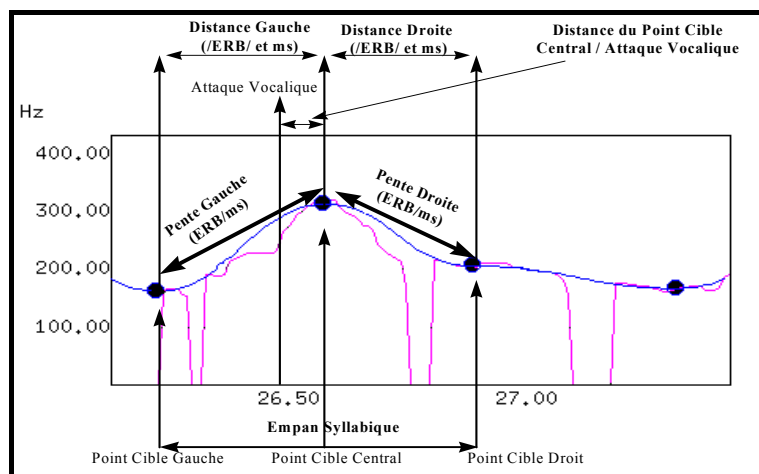


Figure 38 : Modélisation de la f_0 d'une proéminence perçue. 3 points cibles définissent la configuration tonale dont sont extraits les 8 paramètres.

Les paramètres ont été calculés comme suit ⁷:

- **Écart Gauche** = |valeur ERB du PC Central - valeur ERB du PC Gauche |
- **Écart Droit** = |valeur ERB du PC Central - valeur ERB du PC Droit |
- **Distance Gauche** = valeur ms du PC Central - valeur ms du PC Gauche
- **Distance Droite** = valeur ms du PC Droit - valeur ms du PC Central
- **Pente Gauche** = (valeur ERB du PC Central - valeur ERB du PC Gauche) / Distance Gauche
- **Pente Droite** = (valeur ERB du PC Droit - valeur ERB du PC Central) / Distance Droite
- **Empan Syllabique** = étendue, en nombre de syllabes, de la configuration tonale
- **Distance du Point Cible Central par rapport à l'Attaque Vocalique (Noyau)** = calculée en ms.

Nous avons choisi les paramètres d'écart tonal (Écart Gauche et Écart Droit) entre les points cibles pour rendre compte de l'étendue tonale de la proéminence. Les paramètres de distance entre les points cibles nous renseignent sur l'étendue temporelle de la configuration. Les paramètres de pentes droite et gauche intègrent les informations d'étendue tonale et temporelle, et indiquent la direction de pente. C'est plus particulièrement en fonction de ces deux derniers paramètres que l'on décrira les

⁷ les abréviations PC, ms et ERB renvoient respectivement aux notions de Point Cible, millisecondes et valeur de f_0 normalisée selon l'échelle psychoacoustique d' « Equivalent Rectangular Bandwidth » (Hermes & Van Gestel, 1991).

configurations tonales des proéminences. Le paramètre d'Empan Syllabique nous renseigne sur l'étendue de la configuration tonale en termes de syllabes. Enfin, la distance du Point Cible Central par rapport à l'attaque vocalique rend compte du phénomène d'alignement des événements tonals par rapport au texte, en fonction des types de proéminences (Bruce, 1980 ; Kohler, 1987 ; Heuft & Portele, 1996 ; Portele & Heuft, 1997 ; Ladd, sous presse).

Les écarts tonals droit et gauche sont exprimés en valeur absolue pour deux raisons essentielles : l'indication du *poids* de l'écart tonal nous semblait plus pertinente que le sens (+ ou -) de cet écart, sens dont nous disposons par ailleurs dans les paramètres de pentes. D'autre part, l'indication du signe des variations tonales pour les paramètres d'écarts et de pentes auraient été redondants et auraient induit une trop grande corrélation entre les paramètres d'écarts et de pentes, corrélation qui aurait pu fausser l'interprétation des résultats.

2.1.1. Méthodologie statistique

Notre but est de savoir si les 8 paramètres tonals (mesures quantitatives) permettent la distinction des catégories accentuelles définies dans le cadre du modèle de Di Cristo & Hirst. Le traitement statistique d'une analyse multi-paramétrique peut se faire d'un point de vue descriptif à l'aide d'analyses factorielles (Analyse Discriminante ou Analyse en Composantes Principales), ou d'un point de vue quantitatif avec une Analyse de Régression des variables dépendantes qualitatives. C'est cette dernière méthode que nous avons choisi d'appliquer à nos données, à l'aide d'un modèle de régression permettant l'analyse de variables qualitatives (modèle *Logit Multinomial*).

Le modèle *Logit Multinomial* permet d'obtenir des probabilités d'appartenance à un groupe (ici, catégories accentuelles) en fonction des mesures explicatives (paramètres tonals). Cette méthode est toutefois coûteuse lorsque le nombre de variables est important. Afin de réduire le nombre de variables étudiées, la sélection des paramètres tonals pertinents à intégrer dans le modèle *Logit* s'est faite à l'aide d'une *Analyse de Variance*. L'Anova est également à la base de la description phonétique des configurations tonales.

Notre méthodologie comprend donc deux étapes principales : d'une part, dans notre quête des invariants tonals, l'Anova va nous permettre de sélectionner les paramètres indépendants du locuteur qui participent à la distinction des catégories accentuelles. Seuls ces paramètres pertinents seront intégrés dans le modèle de régression ; d'autre part, l'analyse de régression (modèle *Logit*) va nous permettre de quantifier le poids de chacun des paramètres sélectionnés dans la distinction des catégories accentuelles. A l'issue de ce procédé, nous pensons donc pouvoir proposer le ou les paramètres tonals les plus pertinents dans la caractérisation tonale des proéminences en français.

2.1.1.1. Analyse de Variance (Anova) ⁸

Un pré-traitement des variables a été nécessaire avant l'étape d'analyse. Afin de minimiser le 'bruit' ou la variabilité non pertinente, nous normalisons les valeurs de f0. Nous utilisons la même méthode de normalisation que pour les durées segmentales

⁸ Les tests statistiques d'Anova ont été réalisés avec le logiciel StatView pour Windows, version 4.57.

(Transformée Z), par locuteur. Cette méthode nous permet de comparer les variables indépendamment du locuteur ou du style de parole (fluctuation à long terme de la f_0 et registre du locuteur). Par ailleurs, les valeurs de f_0 ont été transformées en ERB (Hermes & Van Gestel, 1991) nous permettant ainsi d'interpréter les variations de f_0 sur une échelle perceptive.

Les 8 paramètres tonals ont été choisis comme variables dépendantes, les facteurs 'locuteurs' et 'catégories accentuelles' étant les variables indépendantes.

2.1.1.2. **Modèle de régression Logit**

Ce type de modèle de régression manipule des variables dépendantes qualitatives (discrètes), alors que les modèles de régression traditionnels manipulent des variables dépendantes quantitatives (continues). Ce modèle étant à choix discret, le R^2 standard est remplacé par un *pseudo- R^2* .

Dans le modèle *Logit*, la quantification de l'effet des variables explicatives est donnée par le *pseudo- R^2* : une valeur proche de 0 signifie que les variables explicatives ne permettent pas de distinguer les catégories, alors qu'elles sont clairement distinguées si la statistique est proche de 1.

Le *pseudo- R^2* mesure donc l'importance des variables explicatives non constantes dans le modèle. Le degré ou le poids de l'effet des facteurs (paramètres tonals) sur la distinction des variables dépendantes qualitatives (catégories accentuelles) est donné par des statistiques de T-Student⁹. Nous avons choisi le R^2 de Nagelkerke (Nagelkerke, 1991) qui pallie la minoration systématique des R^2 de type Mac Fadden (Mac Fadden, 1974).

Nous allons donc tester l'hypothèse selon laquelle les catégories accentuelles vont être distinguées sur la base des paramètres tonals sélectionnés (cf. Figure 39).

⁹ Les T-Student sont une mesure de la significativité des paramètres lorsque, pour 100 observations, ils ont une valeur $> 2,62$ (niveau de significativité : 1%), ou une valeur $> 1,98$ (niveau de significativité : 5%).

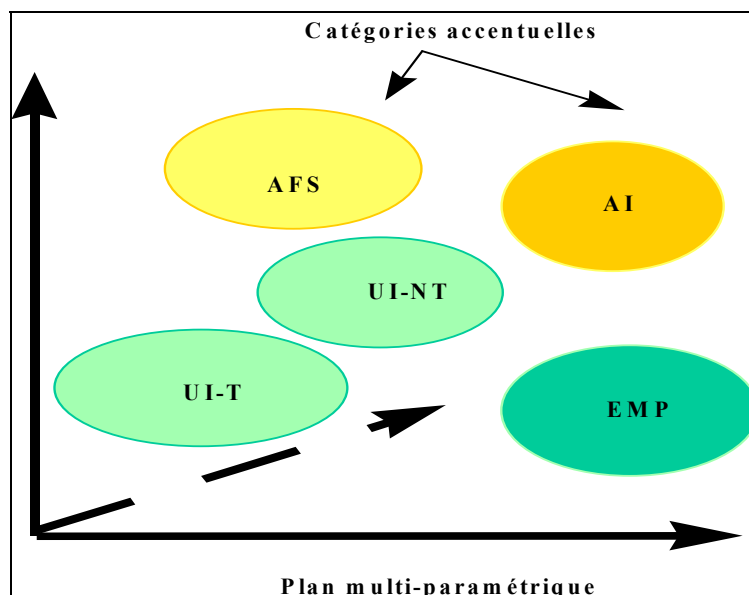


Figure 39 : Hypothèse : les catégories accentuelles peuvent être distinguées sur la base des paramètres tonals sélectionnés (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique).

Le modèle de régression *Logit* permet de quantifier l'effet des paramètres tonals dans la distinction de l'ensemble des catégories (modèle *Logit Multinomial*). Pour mesurer l'importance des mesures explicatives dans la distinction de deux catégories accentuelles seulement (opposition binaire), on procède à une analyse de régression dans le cadre d'un modèle *Logit Binaire*, qui utilise la même statistique (*Pseudo-R²*).

2.1.2. Justification de l'extraction automatique des paramètres et discussion sur la variabilité observée

Nous avons pris le parti d'automatiser l'extraction des configurations tonales pour deux raisons principales :

- un relevé manuel aurait été trop fastidieux et peu pratique pour l'extraction des mesures.
- nous désirions également éviter une trop grande subjectivité dans le relevé des configurations tonales. A cette fin, l'automatisation des relevés permet d'appliquer une méthodologie rigoureuse et identique, dans toutes les situations.

En sortie de l'extraction automatique des configurations tonales, nous recueillons 4 types de configurations tonales : en pic, en creux, montante et descendante. Le pourcentage de type de configuration par catégorie accentuelle est donné dans le Tableau 25 :

Tableau 25 : Pourcentages de configurations tonales par type de configuration et par catégorie accentuelle. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique).

Catégories accentuelles	Configurations tonales en Pics	Configurations tonales montantes	Configurations tonales descendantes et en creux
AFS	76%	5%	19%
UI-NT	50%	15%	35%
UI-T	22%	1%	77%
AI	66%	12%	22%
EMP	89%	7%	4%

On peut remarquer que, pour les proéminences initiales et finales non terminales, plus de 70% des configurations tonales sont en pics, et environ 10% sont montantes. La proéminence de type emphatique initiale (EMP) présente le plus grand pourcentage de configurations en pics. Nous remarquons néanmoins que la proéminence de type finale d'unité intonative non terminale (UI-NT) présente un taux assez faible de configuration en pics, et globalement une assez grande disparité de types de configurations.

Cette variabilité dans les types de configurations tonales peut avoir plusieurs raisons. L'algorithme de modélisation automatique de la f0 n'est peut être pas assez précis temporellement, et la fluctuation temporelle peut être une source de 'bruit'. D'autre part, l'automatisation du relevé du point cible central (le plus proche du noyau) peut induire un choix forcé qui n'aurait pas été celui d'un expérimentateur humain ; en effet, à quelques millisecondes près, un point cible peut être choisi comme central de configuration, alors qu'il apparaîtrait plus logique de le considérer plutôt comme point cible gauche ou droit. C'est ainsi que l'on se retrouve avec 4 types de configurations tonales pour chaque type de proéminence.

Nous avons choisi de ne pas prendre en considération les 4 types de configurations tonales séparément, mais plutôt de les moyenniser, et ce pour plusieurs raisons :

- il existe dans l'ensemble une cohérence non négligeable de types de configurations par catégorie accentuelle ;
- d'un point de vue statistique, l'éclatement en 4 types de configurations aurait rendu nos données peu exploitables, certains types de configurations étant très minoritaires et le nombre d'individus par conséquent trop restreint ;
- dans le même ordre d'idée, choisir uniquement les configurations en pics (pour les proéminences non terminales) et les configurations en creux pour la proéminence terminale d'unité intonative, aurait diminué d'autant le nombre d'occurrences de nos proéminences, et rendu délicat une analyse statistique.
- enfin, les 4 types de configurations peuvent être considérés comme des alloconfigurations d'une même proéminence, comme les manifestations de surface de contraintes rythmiques (taille des groupes rythmiques, équilibrage eurythmique, contexte accentuel, présence ou absence d'un accent initial dans le groupe rythmique, débit local, etc.) et de contraintes linguistiques et sémantico-pragmatiques. En ce sens, il est concevable que la variabilité observée ne soit pas uniquement due à la méthode

automatique d'extraction des données, mais soit motivée linguistiquement. On peut en effet imaginer que les proéminences aient des représentations bimodales ou trimodales (Bolinger, 1985). Cette variabilité linguistique doit être prise en compte dans le cadre d'une analyse syntagmatique (contexte accentuel, pragmatique et discursif). Or, notre travail est centré essentiellement sur l'analyse paradigmatique des proéminences. Rien n'exclurait cependant à l'avenir de procéder à une telle analyse qui, de notre point de vue, ne saurait être alors que qualitative et descriptive.

Une possibilité d'éviter la subjectivité des relevés ainsi que la variabilité des types de configurations aurait pu être de prendre les valeurs de f_0 systématiquement sur la syllabe proéminente (syllabe 0) et les syllabes immédiatement adjacentes (syllabes -1 et +1). Or, cette méthodologie aurait interdit toute prise en considération de l'empan syllabique et de l'étendue temporelle des configurations tonales, information pourtant essentielle dans la distinction perceptive des mouvements tonals (Rossi, 1971).

2.2. Résultats

2.2.1. *Analyse de Variance*

A la lumière des premiers résultats d'Anova, il est apparu que la proéminence de type UI-NT avait un comportement instable d'un locuteur à l'autre. On pourrait rapprocher cela des observations de Vaissière (1983) selon laquelle, dans une tâche de segmentation automatique de la parole en unités prosodiques, on ne peut pas se fier à un type d'information phonétique particulier. En effet, il semblerait que les locuteurs utilisent de manière variable la f_0 et/ou la durée pour marquer les frontières d'unités intonatives. La fonction sémantico-pragmatique de la catégorie UI-NT la rend sans doute particulièrement sensible aux stratégies individuelles de marquage prosodique. Ces résultats vont dans le sens de ce que nous avons constaté pour les durées segmentales, où cette catégorie accentuelle apparaissait être particulièrement sensible aux variations individuelles (cf. § « Niveau de constituance syllabique : Conclusions partielles », p. 192).

Cependant, la variabilité de cette catégorie accentuelle n'est pas à imputer uniquement à sa fonction sémantico-pragmatique. En effet, on aurait pu penser que les catégories accentuelles relevant d'un choix du locuteur (accents emphatiques et accents nucléaires de fin d'énoncé ou de groupes de sens) seraient plus variables que les catégories accentuelles relevant du code linguistique (accents rythmiques initiaux et finaux). Or, les catégories sémantico-pragmatiques UI-T (accents nucléaires d'énoncé) et EMP (accents initiaux emphatiques) présentent une variabilité parfois moindre que celle des catégories rythmiques, aussi bien pour les paramètres tonals que pour les paramètres de durées (pour les analyses détaillées, cf. Astésano, 1999). La variabilité de la catégorie UI-NT pour ce qui est des configurations tonales est donc sans doute à imputer en grande partie au contexte segmental. Nous avons vu dans le Chapitre 5 sur les durées segmentales que cette catégorie accentuelle présente généralement l'allongement syllabique le plus important, ainsi qu'un allongement du noyau vocalique proportionnellement plus important que les autres constituants syllabiques. Sur un tel allongement, plusieurs points cibles sont parfois nécessaires pour rendre compte de la configuration tonale de ce type de proéminence (jusqu'à 3 points cibles sur une même syllabe). Ainsi, le point cible le plus proche du noyau vocalique peut éventuellement se trouver dans le creux de la configuration, i.e. au début de

la syllabe. Un relevé manuel des points cibles constituant la configuration tonale n'aurait sans doute pas choisi le point cible le plus proche du noyau comme point cible central de la configuration, mais aurait choisi logiquement, le point cible le plus haut sur la syllabe (cf. paragraphe précédent sur la justification de l'extraction automatique). Ceci explique sans doute pourquoi ce type de proéminence présente une telle variabilité de configurations tonales.

A la lumière de ces résultats, nous avons pris le parti d'éliminer cette proéminence de notre analyse, afin de réduire une part de variabilité trop importante. Cette variabilité aurait pu nuire à l'estimation globale du modèle de régression et aurait réduit la fiabilité des résultats pour les autres catégories accentuelles.

2.2.1.1. Choix des paramètres à intégrer dans le modèle de régression *Logit*

Les variables pertinentes à retenir pour le modèle *Logit* sont les paramètres indépendants du locuteur (invariants tonals) permettant de distinguer les catégories accentuelles. L'interaction entre les deux variables indépendantes (locuteurs et catégories accentuelles) devait de surcroît être non significative pour augmenter la fiabilité des résultats. En effet, si l'interaction entre les variables indépendantes est non significative, cela indique que les locuteurs ont des stratégies tonales identiques pour chaque catégorie accentuelle. Les détails des Anova et des tests post hoc sont présentés dans Astésano (1999).

Le Tableau 26 indique les paramètres tonals non pertinents pour l'analyse de régression.

Tableau 26 : Paramètres non pertinents pour la description des proéminences (Interaction = interaction entre les facteurs 'catégories accentuelles' et 'locuteurs' ; P-val. = valeur de P ; F-val. = valeur de F).

	Écart Droit (ERB)		Distance Gauche (ms)		Distance Droite (ms)		Distance PC Central/Noyau	
	P-val	F-val	P-val	F-val	P-val	F-val	P-val	F-val
Catégories Accentuelles	,0007	5,751	,2605	1,340	<,0001	26,61	,0101	3,812
Locuteurs	,1786	1,530	,2386	1,358	<,0001	12,95	,0751	2,015
Interaction	,0006	2,691	,1716	1,343	<,0001	7,580	,0217	1,892

Ces 4 paramètres n'ont pas été retenus, pour des raisons différentes : l'Écart Droit présente une interaction des variables indépendantes significative, indiquant que les locuteurs n'ont pas les mêmes stratégies à travers les catégories accentuelles. La Distance Gauche ne permet pas de distinguer les catégories accentuelles. La Distance Droite, en revanche, permet la distinction des catégories mais indique que les locuteurs n'ont pas des stratégies identiques. Le dernier paramètre (Distance du Point Cible central au Noyau) présente une interaction significative, indiquant une variabilité à travers les locuteurs en fonction des catégories accentuelles. Cette variabilité nous interdit donc d'interpréter les résultats de ce phénomène tonal. Notons cependant que les tendances observées semblent néanmoins aller en partie dans le sens des résultats de Kohler (1987), Heuft & Portele (1996) et Portele & Heuft (1997) selon lesquels les proéminences les plus importantes sur le plan perceptif ont tendance à avoir un pic tardif. Ceci semble être le cas notamment pour la proéminence de type emphatique (EMP).

Au vu des résultats de l'Anova, nous conservons donc pour le modèle *Logit* les 4 paramètres suivants : Pentes Droite et Gauche, Écart Gauche (ERB) et Empan Syllabique (cf. Tableau 27).

Tableau 27 : Paramètres pertinents pour la description des proéminences, à inclure dans le modèle *Logit*. (Interaction = interaction entre les facteurs 'catégories accentuelles' et 'locuteurs' ; P-val. = valeur de P ; F-val. = valeur de F).

	Pente Gauche (ERB/ms)		Pente Droite (ERB/ms)		Écart Gauche (ERB)		Empan Syllabique	
	P-val	F-val	P-val	F-val	P-val	F-val	P-val	F-val
Catégories Accentuelles	<,0001	16,67	,0050	4,329	<,0001	11,75	,0175	3,405
Locuteurs	,5529	,796	,4547	,940	,1615	1,589	,5149	,850
Interaction	,0553	1,660	,0432	1,723	,0937	1,519	,8578	,622

Les tests post hoc indiquent que le paramètre de Pente Gauche distingue les catégories deux à deux, sauf AI vs. AFS. La Pente Droite distingue significativement les catégories AFS vs. AI, ainsi que les catégories AI vs. EMP, et UI-T vs. EMP. Le paramètre d'Écart Gauche distingue toutes les catégories deux à deux, et le paramètre d'Empan Syllabique distingue la plupart des catégories, sauf UI-T vs. AFS, et UI-T vs. EMP. Pour ce dernier paramètre, il faut noter que les différences (écarts moyens) entre les catégories sont peu importantes. Nous verrons cependant que ce paramètre semble avoir un effet non négligeable dans la distinction de certaines catégories.

Il est intéressant de noter que les paramètres comprenant uniquement des durées brutes (timing) se sont avérés non pertinents. Apparemment, il existerait une 'normalisation' interlocuteurs en termes de syllabes (empan syllabique) mais pas en termes de durée brute, ce qui semble confirmer le rôle de la syllabe comme unité de planification.

2.2.2. Modèle *Logit Multinomial*

Nous prenons en compte dans le modèle 4 groupes d'accents distincts [AFS, UI-T, AI et EMP] et 4 variables explicatives [Pente Gauche, Pente Droite, Écart Gauche et Empan Syllabique].

Le corpus comprenait à l'origine 790 proéminences perçues. Après avoir éliminé les configurations interrompues par une pause et les proéminences de type UI-NT, le corpus est maintenant constitué de 543 données, dont la distribution est la suivante : 372 AFS ; 36 UI-T ; 81 AI ; 53 EMP.

On peut constater que les données ne sont pas distribuées de manière homogène à travers les catégories accentuelles, ce qui peut rendre délicate la tentative de modélisation.

Tableau 28 : Résultats des pseudo-R² (AFS = accent final de syntagme ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique).

R ² général	R ² AFS/UI-T	R ² AFS/AI	R ² AFS/EMP	R ² UI-T/AI	R ² UI-T/EMP	R ² AI/EMP
0.27	0.32	0.13	0.17	0.52	0.73	0.22

Le Tableau 28 présente les résultats du pseudo-R² du modèle *Logit Multinomial* (R²

général) et ceux des modèles *Logit binomiaux* (catégories accentuelles deux à deux). La catégorie UI-T, étant donné son contour globalement descendant, se distingue clairement dans les oppositions de catégories deux à deux (modèles binomiaux). Le modèle général explique 27% de la variabilité¹⁰ ; les paramètres tonals semblent donc avoir un effet explicatif non négligeable dans la prédiction d'appartenance d'une proéminence à une catégorie accentuelle (cf. Figure 40).

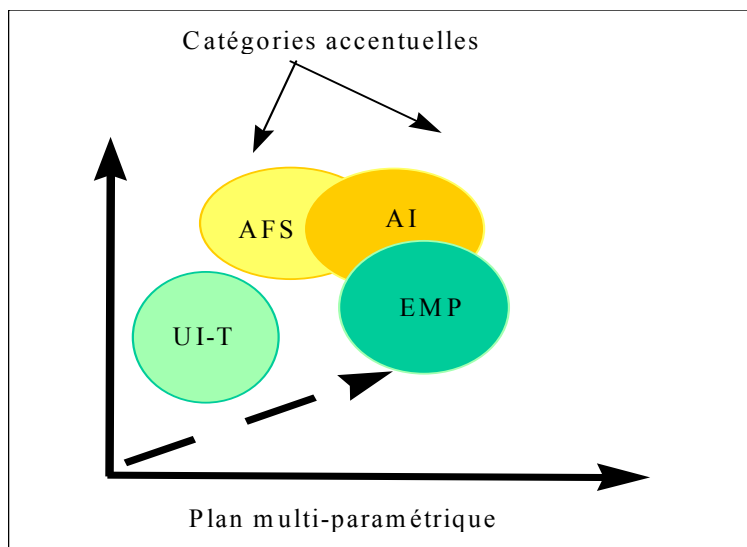


Figure 40 : Résultats : Le modèle général explique 27% de la variabilité : les paramètres tonals semblent donc avoir un effet explicatif non négligeable dans la prédiction d'appartenance d'une proéminence à une catégorie accentuelle. (AFS = accent final de syntagme ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique).

Le détail des tests binaires (modèle *Logit Binomial*) et du modèle général (modèle *Logit Multinomial*), ainsi que les valeurs des T-Student sont présentés dans Astésano (1999).

2.3. Discussion

Individuellement, les 4 paramètres permettent de distinguer clairement les catégories accentuelles (tests post-hocs de l'Anova). Pour le modèle *Logit* général, les statistiques de Student montrent pour les catégories EMP et UI-T notamment, que le rejet de l'hypothèse de non-influence est le mieux testé par la pente gauche ($t_{\text{Pente Gauche}} = 7,5$ et $t_{\text{Pente Gauche}} = -6,2$ respectivement). En combinaison dans les modèles *Logits Binomiaux* néanmoins, les paramètres pertinents diffèrent suivant les catégories opposées.

Considérant seulement les comparaisons phonologiques pertinentes, deux résultats majeurs émergent de ces analyses : la confirmation de la fonction rythmique des proéminences initiales et finales de mots (AI et AF), ainsi que la distinction - relative - entre les proéminences initiales rythmiques et emphatiques (AI et EMP).

¹⁰ Et non 56% comme annoncé par erreur dans Astésano et al., 1998.

- **Fonction rythmique des AI et AFS**

Di Cristo (sous presse) fait état de la ‘visée eurythmique’ des prééminences initiales de mots et finales de mots prosodiques. La tendance à éviter la contiguïté accentuelle ou à favoriser l’équilibre rythmique inter-accentuel en termes de syllabes, s’actualise à un niveau de surface par l’occurrence alternée de prééminences initiales et finales de mots à l’intérieur de l’Unité Intonative (cf. § « Le modèle actuel », p. 95). Les propriétés rythmiques de AI et AFS se manifestent de plus au niveau tonal par une tendance à former un ‘arc accentuel’ (Fónagy, 1980) (pour une discussion, cf. § « L’accentuation initiale : évolution du système accentuel en français », p. 80).

Nos résultats confirment ces caractéristiques à la fois du point de vue de l’empan syllabique et des configurations tonales générales de ces catégories accentuelles. Ces deux types de prééminences ont en effet les Empans Syllabiques les plus étendus, ce qui signifie qu’aucune variation de f_0 majeure ne se produit dans leur voisinage. Il faut noter cependant que les tests post hoc d’Anova distinguent légèrement ces deux types d’accents rythmiques, l’AI ayant un Empan Syllabique plus important que l’AFS. Cette tendance confirme donc l’hypothèse d’eurythmie, notamment pour l’AI. La Pente Gauche ne distingue pas AFS et AI (Tests post-hocs Anova : valeur de $p = 0,4722$; Logit Binaire : T-student = 1,38). En revanche, la Pente Droite les distingue significativement (valeur de $p = 0,001$; T-student = 3,91). La configuration tonale de AFS est symétrique, alors que celle de AI est asymétrique (Pente Droite en moyenne deux fois moins abrupte que Pente Gauche).

La Figure 41 présente les configurations tonales (Pentes Gauches et Pentes Droites) des prééminences de type AFS et AI :

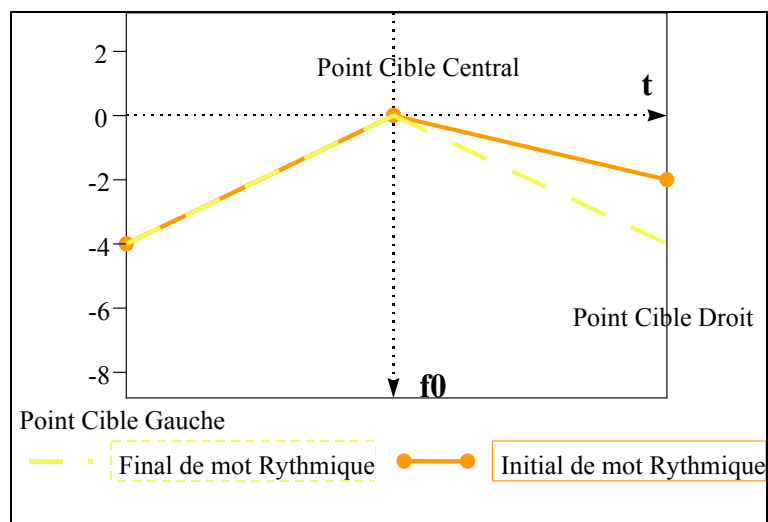


Figure 41 : Configurations tonales des prééminences de type AFS (Final de mot Rythmique) et AI (Initial de mot Rythmique). Les configurations sont déterminées par les Points Cibles Central, Gauche et Droit. L’axe des ordonnées est exprimé en $ERB * 1000$. L’axe du temps (t) est purement indicatif.

La configuration asymétrique de AI pourrait refléter le rôle de ce type de prééminence dans la démarcation à gauche des unités lexicales, et illustrerait la réalité phonétique de

l'arc accentuel. Ces résultats nous amènent également à une comparaison avec les résultats de Jun & Fougeron (1995). Dans une certaine mesure, et malgré nos approches méthodologiques différentes, nos résultats semblent corroborer la description de ce que ces auteurs appellent 'le Syntagme Accentuel' du français ('Accentual Phrase'). Pour ces auteurs en effet, le Syntagme Accentuel est représenté, au niveau sous-jacent, par la combinaison de tons ((LHi)LH) : le premier Hi correspond au ton Haut associé au début du mot ou du groupe de mot, et est optionnel ; le deuxième H correspond au ton Haut associé à la syllabe finale du mot ou du groupe. Dans cette étude, elles décrivent les mouvements ascendants et descendants (Pentes Gauche et Droite) de ces tons H. Ces deux tons ont les mêmes Pentes Gauches, alors que leurs Pentes Droites diffèrent : le timing (et par conséquent la direction) de la Pente Droite du ton H final est constant, alors que celui de la Pente Droite du ton H initial dépend du nombre de syllabes le séparant du ton H final. Par conséquent, la Pente Droite du ton Haut Initial est asymétrique, et moins 'raide' que celle du ton H final.

La question reste de savoir si ces similarités confirment ou non l'existence, à un niveau phonologique profond, de ce Syntagme Accentuel borné à droite par un accent final et à gauche par un accent initial. Pour une discussion plus détaillée, nous renvoyons le lecteur à Hirst & Di Cristo (1996), ainsi qu'à notre propre discussion, p. 93 (§ « L'unité tonale est-elle adaptée à la description prosodique du français ? »).

Enfin, le modèle *Logit binaire* opposant ces deux catégories accentuelles présente un R² faible (R²_{AFS/AI} = 0,13), ce qui signifie que, même si la Pente Droite est un paramètre pertinent dans leur distinction, la *combinaison* des 4 paramètres tonals ne permet pas de les distinguer nettement. Leur similitude fonctionnelle se traduit en effet au plan phonétique par un comportement tonal quasi identique.

- ***Proéminences initiales rythmiques (AI) et emphatiques (EMP) :***

Les 4 paramètres tonals distinguent significativement les deux types de proéminences (tests post-hocs de l'Anova), confirmant par là même leurs rôles distincts sur le plan fonctionnel (pour une discussion sur leur assimilation fréquente dans la littérature, cf. § «

Confusion fonctionnelle des accents initiaux », p. 83). En effet, la configuration tonale de AI n'est pas très saillante et s'étend sur plusieurs syllabes, alors que celle de EMP est très saillante et a un empan syllabique réduit. Elles ont toutes deux la même configuration générale (asymétrique), celle de EMP étant en moyenne 2 fois plus grande que celle de AI.

La figure présente les configurations tonales (Pentes Gauches et Pentes Droites) des proéminences de type AI et EMP :

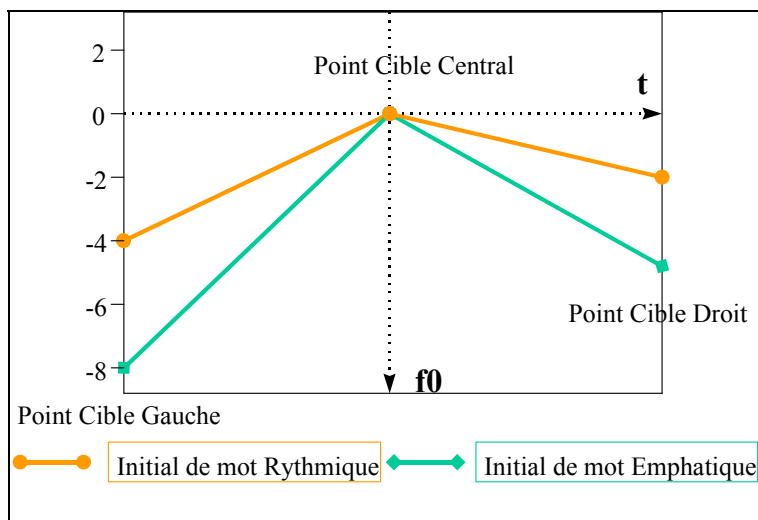


Figure 42 : Configurations tonales des proéminences de type AI (Initial de mot Rythmique) et EMP (Initial de mot Emphatique). Les configurations sont déterminées par les Points Cibles Central, Gauche et Droit. L'axe des ordonnées est exprimé en $ERB * 1000$. L'axe du temps (t) est purement indicatif.

Le modèle *Logit binaire* révèle cependant que seuls l'Empan Syllabique (T-student = -2,82) et la Pente Droite (T-student = -2,34) sont pertinents dans la distinction des catégories. Assez étrangement, en effet, le paramètre d'Écart Gauche (étendue de la variation de f_0) ne semble pas pertinent dans la distinction de ces deux types d'accents (T-student = -0,31). Pourtant, les tests post hocs d'Anova pour le paramètre de Écart Gauche indiquent qu'ils sont significativement distingués par ce paramètre. On peut en déduire que le poids de ce paramètre par rapport aux autres paramètres dans la distinction des proéminences de type AI et EMP n'est pas suffisant. L'on voit ici l'intérêt d'une approche combinée (Anova et modèle de régression multi-paramétrique) pour estimer la pertinence des paramètres tonals dans la caractérisation des proéminences.

Globalement, les paramètres expliquent une part non négligeable de la variabilité du modèle binomial ($R^2_{AI/EMP} = 0,22$) ; cette valeur néanmoins assez faible est sans doute due à la similitude dans l'aspect général de leurs configurations tonales. A ce sujet, nous avons vu que Di Cristo (sous presse) décrit l'accent emphatique initial comme étant l'expression 'Hyper' de l'accent initial rythmique, hypothèse qui semble ici confirmée par nos résultats. Il apparaît donc qu'un changement graduel dans les corrélats phonétiques induit un changement dans la catégorisation fonctionnelle de ces proéminences.

Nous avons particulièrement insisté, dans la partie méthodologique, sur nos difficultés à

distinguer ces deux types d'accents initiaux sur le plan perceptif. Une étude en cours (Jankowski et al., 1999) vise à établir dans quelle mesure la perception de ces accents initiaux est catégorielle ou continue, et quels sont les indices acoustiques les plus pertinents dans cette distinction. Des études sur la perception des proéminences montrent en effet que les auditeurs sont à même de distinguer les proéminences à la fois sur un plan paradigmatique (catégories) et syntagmatique (relations hiérarchiques entre proéminences). La perception de la hauteur est donc à la fois catégorielle et graduelle (Terken, 1993). Il nous semble cependant que la perception des proéminences est catégorielle eu égard à la distribution des accents (finaux vs. initiaux), mais paraît être continue pour les accents de même distribution (finaux rythmiques vs. finaux sémantico-pragmatiques, et initiaux rythmiques vs. initiaux sémantico-pragmatiques).

Malgré la similitude physique des proéminences de type AI et EMP, aussi bien au niveau de la distribution de l'allongement sur les éléments syllabiques qu'au niveau des configurations tonales, nos résultats indiquent que l'on peut distinguer significativement les accents initiaux sur la base de leur durée et des paramètres de f_0 . Si ces résultats valident en quelque sorte la catégorisation fonctionnelle des proéminences du modèle de Di Cristo & Hirst, une étude perceptive serait à l'avenir souhaitable pour compléter ce type de résultats.

Outre ces deux résultats importants, notre étude a également permis de confirmer, sur la base des paramètres tonaux dynamiques, les distinctions fonctionnelles et distributionnelles entre les différents types de proéminences.

- ***Distinction fonctionnelle des proéminences Rythmiques vs. Sémantico-pragmatiques :***

Les proéminences Rythmiques ont un Empan Syllabique plus important et présentent par conséquent des configurations tonales peu saillantes sur le plan physique. En revanche, les proéminences Sémantico-pragmatiques que nous avons pu retenir dans cette étude n'ont pas un comportement identique. En effet, les proéminences de type emphatique (EMP) sont caractérisées par des configurations tonales saillantes (grande variation de f_0 sur peu de syllabes), alors que les proéminences de type UI-T sont particulièrement peu marquées sur la base de leurs caractéristiques tonales (cf. Figure 43). Par contre, les paramètres temporels comme l'Empan Syllabique faible, couplés aux caractéristiques de durées syllabiques (cf. Chapitre 5 - Durées Segmentales -) suggèrent que le caractère saillant de ce type de proéminence relève plus du domaine temporel que fréquentiel. Ceci va dans le sens des résultats d'études antérieures qui montrent la difficulté à percevoir les accents finaux d'unités intonatives terminales en français, de par leurs caractéristiques physiques essentiellement temporelles, moins prégnantes sur le plan perceptif (cf. § « Perception des phénomènes accentuels », p. 65). Bien que nous n'ayons pu inclure dans le modèle descriptif les proéminences de type UI-NT (finale d'unité intonative non terminale), à cause de leur trop grande variabilité, nous aimerions faire remarquer que ce type de proéminence présente des configurations tonales en moyenne très saillantes $[-9,5 ; 9,5]$ sur un faible Empan Syllabique, à l'instar des catégories sémantico-pragmatiques EMP. Leurs configurations sont également symétriques, comme celles des autres types de proéminences finales.

Il apparaît donc que, parmi les proéminences sémantico-pragmatiques, l'UI-T (finale

d'unité intonative terminale) se démarque sur le plan tonal, non seulement par son mouvement tonal descendant, mais également par le caractère peu saillant de ses Pentas Gauche et Droite.

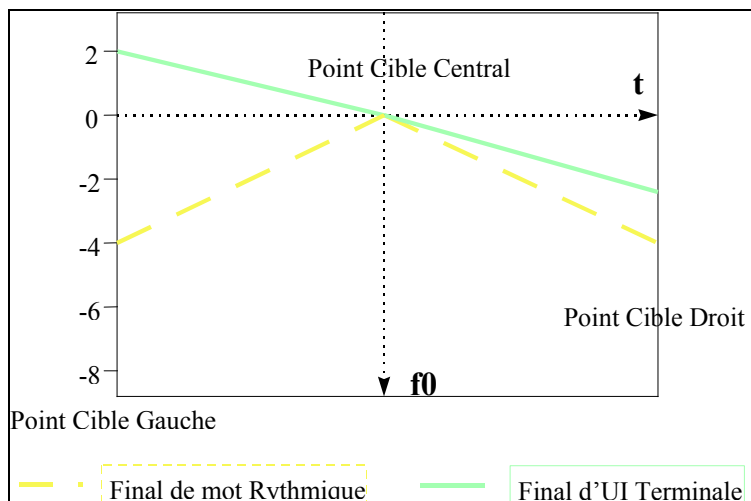


Figure 43 : Configurations tonales des proéminences de type AFS (Final de mot Rythmique) et UI-T (Final d'Unité Intonative Terminale). Les configurations sont déterminées par les Points Cibles Central, Gauche et Droit. L'axe des ordonnées est exprimé en ERB * 1000. L'axe du temps (t) est purement indicatif.

Distinction distributionnelle des proéminences Initiales vs. Finales :

Les configurations tonales des proéminences Finales sont *symétriques*, alors que celles des proéminences Initiales sont *asymétriques*. Pour ces dernières, nous pensons que cette asymétrie reflète le rôle de ces proéminences dans la démarcation à gauche des unités lexicales et leur participation à la formation d'arcs accentuels, intégrateurs d'unités rythmiques (cf. Chapitre sur l'organisation temporelle des proéminences, p. 259).

Les principales caractéristiques tonales des proéminences accentuelles sont résumées dans le Tableau 29.

Tableau 29 : Résumé des caractéristiques tonales des catégories accentuelles (AFS = accent final de syntagme ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique).

Catégories Accentuelles	Hauteur (f0)	Configurations Saillantes	Symétrie des Configurations	Empan Syllabique
AFS	+	-	symétrique	+
UI-T	--	--	symétrique	-
AI	++	-	asymétrique	++
EMP	+++	++	asymétrique	--

3. CONCLUSION

La combinaison des deux approches statistiques (analyse de variance et modèle de régression) nous a permis d'avoir une vision plus précise et plus complète des spécificités des configurations tonales du français. En effet, le modèle de régression *Logit*, en prenant en compte la *combinaison* des paramètres tonals et en *quantifiant* leurs rôles relatifs dans la distinction des catégories accentuelles, nous a permis de rendre compte précisément des similitudes et différences phonétiques, corollaires des rôles fonctionnels des catégories accentuelles.

En effet, si l'analyse de variance révèle qu'un paramètre tonal, pris individuellement, permet de distinguer deux catégories accentuelles, l'analyse de régression multi-paramétrique permet en revanche de relativiser le poids de ce paramètre comparativement aux autres. Le cas du paramètre d'Écart Gauche est à ce titre éloquent : ce paramètre est le seul à pouvoir distinguer très fortement toutes les catégories deux à deux (cf. détails des Anova dans Astésano, 1999). Cependant, en combinaison avec les autres paramètres tonals, il apparaît que, même s'il a un effet non négligeable dans les distinctions, l'Écart Gauche est moins pertinent que les paramètres de pentes notamment.

Il semblerait donc que la prise en compte des paramètres dynamiques de f_0 soit plus pertinente, dans la distinction et la caractérisation des proéminences accentuelles, que le seul paramètre de hauteur de f_0 . On ne peut cependant pas tirer de conclusions au sujet d'une éventuelle programmation des configurations en termes de pentes ou d'écart tonal (niveaux de tons). Notre étude n'a d'ailleurs pas été conçue pour entretenir une controverse à ce sujet. Par contre, nous pensons que la prise en compte des facteurs dynamiques est d'une part plus proche de la réalité perceptive (Théorie de la Gestalt, Bolinger, 1951) et permet d'autre part de savoir quels sont les éléments les plus pertinents dans la distinction des catégories accentuelles. Les tenants d'une approche phonologique autosegmentale ignorent généralement les aspects dynamiques de f_0 . Or, on peut supposer que ce genre d'information est central eu égard à l'intégration perceptive des variations tonales (Rossi, 1971, notamment).

Comme nous l'avons vu, nos résultats indiquent que les variations tonales importantes (pour les accents emphatiques et les accents finaux d'unités intonatives terminales et non terminales en l'occurrence) interviennent sur des empanns syllabiques moins larges que les variations tonales réduites (typiques des catégories accentuelles rythmiques initiales AI et finales AFS). Les durées syllabiques des premiers types de catégories accentuelles (sémantico-pragmatiques) étant largement supérieures à celles des catégories accentuelles rythmiques¹¹, il y a donc ralentissement du débit local, de la 'cadence' en terme musical. Ces résultats sont cohérents avec les exigences perceptives des auditeurs, selon lesquelles les événements tonals pertinents (figures signifiantes) doivent se détacher du 'fond' de manière optimale en maximisant la fission perceptive (McAdams, 1997).

¹¹ cf. § « Niveau de constituance syllabique : Conclusions partielles », p. 192

« Le degré de séparation [perceptive] dépend de la différence entre les fréquences (Δf) et de la cadence des événements, les cadences plus lentes exigeant un plus grand Δf avant de se scinder en deux flux. Réciproquement, un plus petit Δf exige une cadence plus rapide pour la fission perceptive. » (McAdams, 1997 : 8).

Il serait intéressant de procéder à une étude complémentaire afin de valider, sur le plan perceptif, les tendances dégagées par notre analyse acoustico-phonétique des configurations tonales des proéminences.

Pour conclure, la capacité du modèle *Logit* général à expliquer une part non négligeable de la variabilité sur la seule base des paramètres tonals nous invite à poursuivre dans cette approche. Une étude sur l'intégration des paramètres temporels (durée syllabique) aurait été souhaitable, mais la variabilité de durée inter-stylistique et inter-locuteur notamment pour les catégories sémantico-pragmatiques, nous interdit malheureusement d'appliquer la même méthodologie aux durées¹².

Enfin, l'invariance observée au niveau paradigmatique à travers les locuteurs (sauf pour la catégorie UI-NT), nous permet de conclure que les stratégies locales d'actualisation des proéminences sont identiques quels que soient les locuteurs ou les styles de parole. La distinction stylistique que l'on pourrait opérer sur la base des paramètres objectifs de f_0 doit se situer dans l'organisation temporelle des proéminences, i.e. la gestion rythmique des paramètres de f_0 . Une grande part de la variabilité des configurations tonales peut en effet s'expliquer non seulement par le contexte segmental, non négligeable, mais également par le contexte accentuel, le débit local et global, etc. Cette gestion rythmique devrait s'actualiser en termes de succession des événements dans le temps, la fréquence des proéminences et la proportion des catégories accentuelles et par conséquent des types de constituants métriques.

C'est ce que nous nous proposons d'analyser maintenant dans le chapitre consacré à l'organisation temporelle de la parole.

¹² cf. § « Discussion et conclusion générale sur les tendances d'allongement différentiel en fonction des contraintes prosodiques et des styles de parole », p. 225

4. RESUME

Pour cette analyse, notre corpus multistyle se justifie dans la recherche d'invariants tonals susceptibles de rendre compte du système noyau du rythme. Nous pensons en effet que les paramètres locaux de f0 reflètent l'information structurale et sémantico-pragmatique du message indépendamment des styles de parole, alors que les paramètres globaux sont plutôt révélateurs des différences de processus d'encodage (Hirschberg, sous presse).

Notre analyse est fondée sur les paramètres dynamiques de f0 (pentes, empan syllabique des configurations tonales etc.). Nous cherchons à savoir si les configurations tonales participent de la distinction des catégories accentuelles, et, si c'est le cas, quels paramètres de f0 sont les plus pertinents dans la distinction. A cette fin, nous procédons à une analyse de régression permettant de quantifier le poids des paramètres tonals dans la distinction des catégories accentuelles.

Deux résultats majeurs émergent de cette étude :

- d'une part, la fonction rythmique des proéminences finales et initiales de mot (AI et AFS) est confirmée (empan syllabique étendu). Ces proéminences présentent des Pentes Gauches similaires, et sont essentiellement distinguées par leurs Pentes Droites.

- d'autre part, la distinction fonctionnelle entre les proéminences initiales emphatiques (EMP) et rythmiques (AI) est validée (la configuration tonale de EMP est en moyenne deux fois plus saillante que celle de AI). Par contre, leur similitude distributionnelle est reflétée dans la forme générale de leurs configurations tonales (asymétrique).

De manière générale, l'analyse des configurations tonales des proéminences a permis de mettre en évidence la distinction fonctionnelle des catégories accentuelles, les accents rythmiques présentant des configurations tonales moins saillantes que les accents sémantico-pragmatiques. De même, la distinction distributionnelle des proéminences est validée, les accents finaux étant caractérisés par des configurations tonales symétriques, et les accents initiaux par des configurations asymétriques.

La forme asymétrique des configurations tonales des catégories accentuelles initiales reflète selon nous le rôle de ces proéminences dans la démarcation à gauche des unités lexicales et leur participation à la formation d'arcs accentuels, intégrateurs rythmiques des unités lexicales et sémantiques.

- CHAPITRE 7 -

***L'Organisation Temporelle
de la Parole***

I. INTRODUCTION

Nous avons vu, dans les chapitres consacrés à l'organisation paradigmatique des proéminences, que les phénomènes prosodiques locaux (durées segmentales et configurations tonales des proéminences) présentent dans une certaine mesure des caractéristiques communes aux différents styles de parole. Hirschberg (sous presse) constate que la variabilité imputable aux différents processus d'encodage se manifeste essentiellement dans les phénomènes prosodiques globaux. Nous nous proposons dans ce chapitre d'examiner les paramètres temporels globaux, et plus particulièrement l'organisation temporelle des proéminences, en relation également avec des phénomènes rythmiques tels les pauses et le débit. C'est en effet à travers l'organisation temporelle que nous pensons plus particulièrement pouvoir mettre en évidence la variabilité stylistique.

Lors de l'exposé de notre cadre théorique et méthodologique, nous avons constaté que le français possède une double identité accentuelle et rythmique (cf. § « Introduction : le mythe de l'isochronie et de l'isosyllabité », p. 32 ; et § « L'accentuation initiale : évolution du système accentuel en français », p. 80). D'une part, rythme syllabique et rythme accentuel semblent coexister (Wenk & Wioland, 1982 ; Fant et al., 1991a et b, notamment). D'autre part, il existe une tendance de plus en plus répandue à marquer le début des mots ou des groupes de mots par un accent initial ayant une 'parenté' physique avec les accents de l'anglais par exemple, réputée langue à isochronie accentuelle. Cette tendance peut contribuer dans une certaine mesure à la perception globale d'une récurrence régulière subjective des accents. Il y aurait donc aussi coexistence de l'accentuation finale traditionnelle, marquée par l'allongement, avec une accentuation initiale (mélodique) (Fónagy & Fónagy, 1976 ; Fónagy, 1980 ; Hirst & Di Cristo, 1984 ; Padeloup, 1990 ; Vaissière, 1991 ; Di Cristo & Hirst, 1997 ; Di Cristo, sous presse).

Cette 'dualité' dans les formes accentuelles du français, ainsi que la tendance à réaliser des groupes rythmiques réguliers malgré le poids syllabique important, nous conduit à considérer le français comme une langue où succession syllabique et alternance accentuelle (Allen, 1980) contribuent à sa spécificité rythmique.

Pour rendre compte de cette spécificité rythmique et accentuelle du français, le besoin d'un modèle descriptif adapté au français se fait sentir. En effet, on a souvent pu lire que les langues n'ayant pas les mêmes caractéristiques prosodiques que les langues de 'prestige' (ou plus connues) sont des langues 'arythmiques'. Ce fut longtemps le cas notamment du français. Des recherches indépendantes des courants de descriptions traditionnels ('anglo-centrés') s'imposent pour des langues comme le français tout particulièrement, langue que l'on oppose systématiquement à l'anglais sur le plan prosodique et phonétique.

*« All languages have rhythm, but more **independent research** needs to be done to discover exactly what the rhythmic principles are in languages which do not show a tendency towards 'stress-timing' » (Dauer, 1987 : 449).*

C'est pourquoi, pour la description du rythme du français, notre choix s'est porté sur le modèle de Di Cristo & Hirst. Ce modèle de description de l'accentuation en français présente l'avantage de se situer dans un cadre de description métrique qui fait appel aux connaissances universelles sur le rythme linguistique. Cependant, il prend en compte les

spécificités prosodiques du français et, dans sa version actuelle, permet de rendre compte des phénomènes accentuels propres aux situations de communication spontanées (cf. § « Le modèle actuel », p. 95).

Nous centrons notre analyse sur l'observation des notions de base, fondements même du rythme : la 'périodicité' et la 'structure'. La périodicité, selon Fraisse (1956, 1974), est la récurrence d'éléments discrets qui se distinguent de leur environnement. En l'occurrence, il s'agit ici des proéminences. La 'structure' a trait aux éléments constitutifs de ce qui est répété périodiquement. A ces notions de 'périodicité' et 'structure' sont souvent rattachées les notions de 'dissimilation' et 'assimilation' (cf. « Loi d'assimilation et de dissimilation », p. 29). Nous verrons que les éléments périodiques sont créateurs de structure, et par conséquent, les lois de dissimilation et d'assimilation constituent les deux facettes d'une seule réalité.

Si la rythmicité d'une langue peut varier suivant le style de parole, nous ne pensons pas que la parole lue et la parole spontanée soient régies par des principes d'organisation différents. En effet, selon Guaïtella (1991), la différence de processus d'encodage se refléterait dans l'inclination pour la parole lue à favoriser une organisation métrique, i.e. à la régularisation des intervalles interaccentuels ('assimilation'), et pour la parole spontanée à favoriser une organisation rythmique, i.e. à la production d'événements contingents ('dissimilation'). Dans le même ordre d'idée, Fónagy (1967) propose de distinguer la poésie et la parole spontanée, en insistant sur le caractère prévisible et programmé du premier type de discours qui lui confère un aspect redondant.

« Le rythme du vers métrique est par définition plus redondant que le rythme d'une conversation. Dans le vers métrique, le retour des syllabes marquées par l'accent ou par la longueur, comme le retour des pauses est réglé d'avance, est préétabli. Rien de tel dans la parole où la langue ne limite que le champ de variation de distribution des syllabes accentuées. » (Fónagy, 1967 : 316).

Nous sommes consciente des différences fondamentales entre la parole lue et la parole spontanée. Les processus cognitifs impliqués dans la parole spontanée sont différents de ceux de la parole lue. Cette différence majeure se reflète dans l'organisation prosodique globale des deux types de discours, et dans la production d'événements propres à la parole spontanée, comme une organisation pausale particulière (hésitations vocales, distribution des pauses), la production d'événements d'ordre énonciatif (réalisation des éléments d'emphase, très marqués prosodiquement, alors qu'ils sont moins marqués en parole lue, notamment au niveau de la durée).

Nous cherchons, par l'utilisation d'un cadre méthodologique identique appliqué à différents styles de parole, à mettre l'accent sur les similitudes prosodiques et à mettre en valeur, le cas échéant, les particularités rythmiques et accentuelles des différents styles de parole. Par là même, nous ne nous focalisons pas sur les différences posées 'a priori' entre les styles de parole. Nous pensons que l'utilisation d'un cadre méthodologique identique, rigoureux, permet justement de mieux mettre en valeur les différences stylistiques. Ces différences seront tangibles, voire quantifiables, et en tout cas comparables terme à terme, ce que ne permet pas une démarche postulant 'a priori' des différences radicales entre les styles de parole.

Nous émettons l'hypothèse que les deux tendances accentuelles et rythmiques coexistent dans les styles (assimilation et dissimilation), et contribuent également à les situer sur une 'échelle de métricité'. Ce qui distingue les styles de parole semble être le *degré* d'assimilation ou de dissimilation. Nous émettons l'hypothèse que ces deux tendances se retrouvent à différents niveaux de l'organisation prosodique :

- au niveau accentuel, la dissimilation est représentée par les proéminences initiales. En effet, les accents initiaux sont différents des accents finaux 'traditionnels' tant sur le plan physique que sur le plan distributionnel. En ce sens, on peut supposer qu'ils introduisent dans le rythme de la langue une rupture sonore. En revanche, l'assimilation est représentée par les proéminences finales, qui participent à la formation de constituants hiérarchiques.

- au niveau des groupements hiérarchiques, la dissimilation pourrait être représentée par la tendance à une plus grande variabilité structurelle des unités prosodiques (coefficient de variation), et l'assimilation par la tendance à moins de variabilité des constituants rythmiques.

Les différences de processus d'encodage se traduiraient également par la prédilection pour un *niveau* d'actualisation du rythme : les syllabes seraient plus isochrones en parole lue, les Groupes de Souffle plus réguliers en parole spontanée (Vaissière, 1991). A la suite de Couper-Kuhlen (1993), nous postulons que la rythmicité se situe à plusieurs niveaux de la structure prosodique. Cependant, nous émettons l'hypothèse que les différents niveaux d'actualisation du rythme ne sont pas sollicités de la même manière suivant le style de parole. Nous pensons que la façon dont les différents niveaux de constituance prosodique sont affectés par la rythmicité contribue à 'classer' les différents styles de parole sur une 'échelle de métricité'.

A travers l'étude de trois styles de parole (*Interview, Journal, Lecture*), nous nous proposons donc de mettre en évidence à la fois les schèmes rythmiques spécifiques du français, et la variabilité rythmique induite par les processus d'encodage de la lecture et de la parole spontanée.

2. PLAN EXPERIMENTAL

A l'instar de Jassem et al. (1984), nous désirons répondre à plusieurs questions concernant le principe de rythmicité.

- Quelles sont les mesures appropriées que l'on doit appliquer au signal pour tester l'hypothèse de la rythmicité ? Existe-t-il par exemple des unités de référence dans la parole, en fonction desquelles les principes rythmiques s'organisent ?

- Quelles unités sont affectées par le phénomène d'isochronie : les phones, les syllabes, les syllabes accentuées, les Pieds Métriques, ou encore les niveaux métriques supérieurs (Couper-Kuhlen, 1993) ?

- La rythmicité est-elle vraie de manière identique pour tous les styles de parole, ou bien est-ce une variable dépendante étant par exemple plus évidente en lecture de poésie, moins en lecture de prose et en parole lente et contrôlée, et peut-être encore moins évidente en discours rapide et non contrôlé ? Il s'agit d'une question centrale

dans notre étude : notre hypothèse est qu'il y a bien une variabilité induite par les styles de parole ; cependant, cette variabilité n'est pas le reflet de systèmes différents, mais l'expression d'alloformes d'un système noyau.

Afin de répondre à ces questions, nous proposons de rendre compte de la rythmicité du français à travers 4 niveaux d'analyse :

1. l'analyse de la variabilité des niveaux de constituance ;
2. l'analyse des rapports de rythmicité entre les événements prosodiques et des unités de référence de l'organisation rythmique (syllabes inaccentuées, pauses et Intervalles InterAccentuels moyens) ;
3. l'analyse de l'organisation temporelle des Mots Prosodiques en fonction de la *nature* des accents les constituant ;
4. et enfin, l'analyse de la gestion stylistique des phénomènes de pause.

Avant d'exposer nos résultats, nous présentons notre cadre méthodologique et nous définissons les unités d'observation de la rythmicité.

2.1. Cadre méthodologique et procédure expérimentale

Notre cadre théorique comprend les niveaux métriques de la syllabe, du Pied Métrique, et du Mot Prosodique. Cependant, nous choisissons dans cette étude d'examiner la rythmicité du français essentiellement à travers l'analyse de deux niveaux de constituance métrique : le Mot Prosodique (MP) et la syllabe.

- Le niveau du MP nous permet de rendre compte de la notion de groupement et d'alternance accentuelle,
- et le niveau de la syllabe, de la notion de succession et de régularité syllabique, qui serait caractéristique du rythme du français.

Nous tenons compte également du Groupe de Souffle (GS), délimité par deux pauses respiratoires, en tant qu'unité de performance.

A partir de nos 10 minutes de corpus constitué de trois styles de parole, un expert a localisé les proéminences perceptivement et les degrés de frontières. Sur un total de plus de 2600 syllabes, le corpus comprend 790 proéminences, donc autant de Pieds Métriques (PM), 630 Mots Prosodiques (MP) et 156 Groupes de Souffle (GS).

Nous avons exprimé la variance aux niveaux de la syllabe inaccentuée (INA), du MP et du GS, afin de rendre compte de la régularité des événements à un même niveau de constituance ; cette variance est exprimée par le rapport de l'écart type à la durée moyenne des constituants (coefficient de variation) et est exprimée en pourcentage.

Également, nous avons voulu établir les rapports de rythmicité qui pourraient exister entre des unités de référence (le 'fond') et des unités variables (la 'figure') :

- au niveau syllabique, nous avons mis en rapport la durée de la syllabe inaccentuée avec la durée des syllabes accentuées ;

- au niveau des constituants, nous avons comparé la durée de l'intervalle interaccentuel moyen à celle de l'entité pause + syllabe prépausale (Fant et al. 1991a et b ; Fant, 1991).

L'analyse de la nature et de la durée des pauses, enfin, nous permet de rendre compte des relations rythmiques que les pauses entretiennent avec les autres phénomènes prosodiques, comme les syllabes accentuées et le débit de parole.

2.1.1. Définition des unités d'observation

2.1.1.1. Niveau du Mot Prosodique

Le Mot Prosodique, tel que nous l'avons défini, correspond au 'Mot Phonologique' tel que défini par Jassem et al. (1984) (Unité Rythmique Totale), et Crompton (1980). Selon eux, un Mot Phonologique peut comporter deux groupes accentuels, par le jeu des accents secondaires lexicaux. Ce niveau métrique permet de rendre compte des phénomènes d'anacrouse (syllabes inaccentuées précédant une syllabe accentuée au début d'un énoncé en anglais). Les auteurs se démarquent ainsi de l'approche d'Abercrombie (1964) notamment, qui ne prend en compte que les Pieds Métriques au sens strict et considère les syllabes précédant les syllabes accentuées en anglais comme extramétriques.

Le Mot Prosodique tel que nous le concevons présente une meilleure adéquation avec les unités lexicales et les groupes syntaxiques. De plus, en français, les Mots Prosodiques définissent le domaine d'application des règles de liaison, et en ce sens constituent une unité phonologique cohérente. Ils sont également l'unité d'actualisation du schéma bipolaire de l'accentuation du français (accentuation initiale et finale d'un mot ou d'un groupe de mot), et par conséquent le lieu de réalisation des règles d'équilibrage rythmique (Di Cristo, sous presse). A ce sujet, Di Cristo distingue essentiellement 2 cas de schémas accentuels bipolaires :

- les mots lexicaux de plus de deux syllabes sont doublement accentués (accent initial sur la première syllabe ; accent final sur la dernière syllabe). Ces schémas seraient plus caractéristiques de certains phonostyles (style journalistique notamment) ;
- les groupes de mots ou syntagmes sont accentués à l'initiale et à la finale du groupe. Cette bipolarisation accentuelle au niveau du syntagme a deux motivations : elle permet d'une part d'éviter les phénomènes de collisions accentuelles, et d'autre part de mettre en valeur un 'motif' rythmique non motivé par l'évitement d'une collision accentuelle. Dans le cas d'une bipolarisation au niveau du syntagme, on trouve généralement la manifestation du phénomène de désaccentuation, qui entraîne l'effacement d'un accent final. C'est à ce niveau de bipolarisation que Di Cristo voit la manifestation de « l'arc accentuel » tel qu'il a été défini par Fónagy (1980).

Dans la suite de notre exposé, nous ferons parfois référence aux Mots Prosodiques 'bipolaires' par le terme d' « Arcs Accentuels », qu'ils se situent au niveau du mot lexical ou au niveau du syntagme.

En nous inspirant du modèle de Di Cristo & Hirst (1997) (développé dans Di Cristo, sous presse), nous distinguons deux types de MP :

– les MP Simples constitués d'un seul Pied Métrique (PM) borné à droite par un accent final situé à une frontière prosodique mineure (accent final de syntagme, AFS), ou à une frontière prosodique majeure (Unité Intonative) non terminale (UINT) ou terminale (UIT) ; c'est ainsi que certains MP seront congruents avec le niveau métrique supérieur, le niveau de l'Unité Intonative ;

– les MP Complexes enfin, constitués de deux PM, le premier PM étant borné à droite par un accent initial rythmique (AI) ou emphatique (EMP) et le deuxième PM borné à droite par un accent final (de frontière mineure ou majeure). Nous considérons ce type de MP Complexes comme la manifestation de l'arc accentuel (Fónagy, 1980).

Le modèle accentuel que nous utilisons (Di Cristo & Hirst, 1997 ; Di Cristo, sous presse) nous permet par conséquent de mettre en évidence l'influence de la nature de l'accent (final de syntagme ou d'unité intonative ; initial rythmique ou emphatique) sur l'organisation temporelle des Mots Prosodiques.

La Figure 44 présente la grille métrique pour la partie d'énoncé « un catalogue de promesses pour les femmes » prononcée par la locutrice *Interview F*. Les niveaux du Pied Métrique, du Mot Prosodique et de l'Unité Intonative sont représentés. La figure présente le cas des MP Complexes, constitués d'un accent initial et d'un accent final (2 PM), ainsi que le cas d'un Mot Prosodique Simple congruent avec l'unité métrique immédiatement supérieure : l'Unité Intonative.

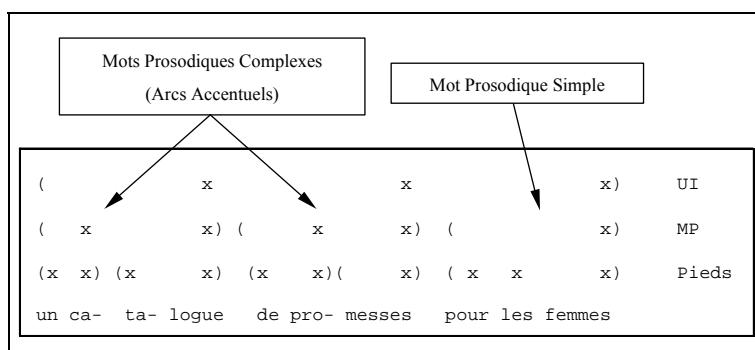


Figure 44 : Représentation des niveaux de constitution métrique d'une partie d'énoncé produit par la locutrice *Interview F*, selon la représentation en grille métrique parenthésée de Halle & Vergnaud (1987). Les Mots Prosodiques complexes représentent, dans cet exemple, une bipolarisation accentuelle au niveau du mot lexical.

Nous préférons la représentation en double parenthésage proposée par Halle & Vergnaud (1987) car elle permet une meilleure représentation des groupements et des structures, notamment au niveau du MP.

Sans remettre foncièrement en cause la notion de Pied Métrique (PM), nous privilégions le niveau du MP en tant qu'unité d'intégration rythmique qui favorise la formation « d'arcs accentuels ». A l'instar de Fónagy (1980), nous pensons que l'arc accentuel, en favorisant le groupement par le marquage initial (gauche) d'une première unité lexicale et le marquage final (droit) d'une deuxième unité lexicale, est un intégrateur rythmique des unités sémantiques. L'arc accentuel peut également délimiter une seule unité lexicale à droite et à gauche comme c'est le cas dans l'exemple ci-dessus ; dans ce cas, l'accent initial a

également une fonction de démarcation lexicale. Nous pensons que c'est à l'intérieur des MP que s'opèrent les règles phonologiques d'équilibrage rythmique, participant de la 'visée eurythmique', par l'alternance et l'agencement des syllabes accentuées et inaccentuées.

2.1.1.2. Niveau de la syllabe : syllabe inaccentuée, syllabe de référence

Outre ce constituant rythmique, central dans notre étude, nous intégrons également à notre recherche le constituant de plus bas niveau dans la hiérarchie métrique, la syllabe, et plus précisément à l'intérieur de la syllabe, la syllabe inaccentuée (INA). Cette dernière est en effet décrite par les psycholinguistes comme la syllabe 'étalon', dont la durée constitue une référence dans la structure rythmique de la phrase (Fraisie, 1974 ; Guañella, 1988). Également, elle est supposée être de durée relativement stable en français, comparativement aux langues à rythmicité accentuelle.

Les syllabes inaccentuées représenteraient donc une unité de référence, jouant le rôle de 'fond' sur lequel, d'après le principe de Gestalt, se détacherait la 'figure' représentée par les proéminences. Cependant, il nous paraît intéressant de faire remarquer qu'une organisation temporelle particulière des syllabes atones est à l'œuvre, qui favoriserait le détachement perceptif optimal des proéminences selon leur distribution (initiale ou finale). Une analyse syntagmatique des syllabes atones encadrant la réalisation des accents initiaux et finaux semble en effet indiquer une stratégie apparemment différente de mise en relief des deux types d'accents. Jankowski (en cours) analyse la durée des syllabes atones situées immédiatement avant et immédiatement après une syllabe accentuée (syllabes pré- et post-accentuelles), ainsi que la durée des syllabes inaccentuées ne précédant ni ne suivant immédiatement une syllabe accentuée (syllabes atones, à proprement parler). Il note une dissymétrie de durée des syllabes 'pré' et 'post', dont la tendance est inverse suivant qu'elles sont dans le voisinage d'un accent initial ou final. Autour de l'accent initial, l'auteur observe une augmentation de la durée de la syllabe pré-accentuelle, et un raccourcissement de la syllabe post-accentuelle (plus courte que les syllabes atones). L'inverse se produit pour les accents finaux, pour lesquels la syllabe pré-accentuelle est notablement raccourcie, alors que la syllabe post-accentuelle est légèrement allongée. Jankowski en déduit que les stratégies de mise en valeur (figure sur fond) sont spécifiques au type d'accent, sur la base de leur distribution initiale ou finale. Il semblerait au vu de ces résultats qu'une proéminence démarcative à droite (finale) prenne pour référence et contraigne la syllabe précédente, alors qu'une proéminence démarcative à gauche (initiale) prend pour référence et contraint la syllabe suivante.

2.1.1.3. Pauses : unités rythmiques de référence

Selon Fant et al. (1991a et b), un lecteur soucieux de produire un rythme régulier tend à produire des intervalles interaccentuels moyens (non suivis de pause) de durée proportionnelle (multiple entier) à la durée moyenne de la syllabe prépausale + celle de la pause. Le rythme accentuel est donc mis en valeur par les pauses.

Pour notre étude, nous choisissons les intervalles interaccentuels (I.I.A.) délimités par des accents finaux de syntagmes (AFS) comme intervalles de référence. L'unité de référence rythmique que nous avons choisie est le Mot Prosodique Simple, mais syllabe accentuée (finale) exclue : il s'agit donc de l'Intervalle Inter-Accentuel à proprement parler. On se fonde ainsi sur une réalité psychoacoustique (Gestalt, psychologie de la forme) selon

laquelle la reconnaissance de patrons de syllabes accentuées récurrentes sur un fond de syllabes inaccentuées rend compte du fait que des intervalles très différents soient perçus comme égaux (Faure et al., 1980).

En ce qui concerne la durée moyenne de la syllabe prépausale, nous choisissons celle des syllabes accentuées finales d'unité intonative non terminale (UI-NT), syllabes qui sont suivies de pauses à plus de 80% des cas dans notre corpus.

Enfin, nous avons retenu les Pauses Respiratoires comme unité pausale de base.

Notre but est de vérifier dans quelle mesure cette relation d'équilibre rythmique est vérifiée à travers les différents styles de parole. Nous avons donc mis en relation la durée moyenne (en ms) des Intervalles Inter-Accentuels (IIA) formés par des AFS avec la durée moyenne (en ms) de l'entité rythmique allongement prépausal UI-NT + Pauses Respiratoires.

3. RESULTATS

3.1. Niveaux de rythmicité

Selon Couper-Kuhlen (1993), l'établissement de patrons isochrones, et donc la rythmicité, peut intervenir à n'importe quel niveau de la hiérarchie prosodique. Ce constat contraste avec la croyance traditionnelle selon laquelle on trouverait l'isochronie, si elle existe, uniquement au niveau des syllabes accentuées dans les langues à rythmicité accentuelle comme l'anglais, et non pas au niveau de la syllabe comme se serait le cas par exemple pour les langues à rythmicité syllabique. L'auteur va également plus loin que la vision traditionnelle de l'isochronie en ce sens qu'elle établit l'existence de l'isochronie à des niveaux supérieurs aux syllabes accentuées, i.e. entre les nuclei (ou accents de hauteur majeurs) dans les séquences de syntagmes intonatifs. Enfin, selon Couper-Kuhlen, plusieurs niveaux de rythmicité peuvent se superposer, probablement pour augmenter la saillance globale du rythme.

Nous nous proposons dans le paragraphe suivant de tester l'hypothèse des niveaux multiples de rythmicité.

3.1.1. *Constituance et variabilité rythmique*

Une manière d'évaluer la stabilité temporelle des constituants prosodiques consiste à mesurer le pourcentage de dispersion autour de la moyenne (en durée brute) de chaque constituant (coefficient de variation).

- Au niveau syllabique, la stabilité de durée des syllabes inaccentuées nous renseigne sur la régularité syllabique.
- Au niveau du Mot Prosodique (MP), qu'il soit simple (un PM) ou complexe (deux PM), le coefficient de variation nous renseigne sur la régularité accentuelle, autrement dit sur la tendance à la récurrence des accents à intervalles réguliers. Les MP simples finaux comprennent l'ensemble des unités rythmiques délimitées par des accents finaux, qu'ils soient situés à des frontières prosodiques mineures (AFS) ou majeures (UINT et UIT). Les MP complexes sont constitués d'un PM défini par un

accent initial rythmique ou emphatique (AI et EMP) et d'un PM défini par l'un des trois accents finaux.

Le Tableau 30 donne les coefficients de variation correspondant aux différents niveaux de constituance pour les trois styles de parole.

Tableau 30 : Coefficient de variation (rapport de l'écart type sur la moyenne de durée brute, exprimé en pourcentage) et niveaux de constituance en fonction du style de parole. (INA = syllabes inaccentuées ; -/Finaux = Mots Prosodiques Simples délimités par des accents finaux de syntagme (AFS), ou d'unité intonative terminale (UI-T) ou non terminale (UI-T) ; AI/Finaux = Mots Prosodiques Complexes contenant un PM défini par un Accent Initial rythmique (AI) ou emphatique (EMP)).

Styles de parole	Syllabes INA	Mots Prosodiques simples et complexes		Groupes de Souffle
		-/Finaux	AI/Finaux	
Interview	46 %	36.5 %	26.5%	47.5 %
Lecture	33 %	27 %	22 %	43 %
Journal	33.4 %	35 %	30 %	53 %

Les résultats présentés dans ce tableau sont repris sous formes d'histogramme dans la Figure 45, p. 270.

Il apparaît que la marge de variabilité est en moyenne plus faible pour le niveau des MP qu'aux niveaux des syllabes inaccentuées ou des Groupes de Souffle, ce qui va dans le sens d'une certaine isochronie accentuelle en français (Fant et al., 1991a et b). A l'intérieur des MP, les MP complexes sont nettement plus réguliers que les MP simples, surtout à l'intérieur du style *Interview*. L'unité rythmique de référence serait donc le MP complexe, ce qui confère à l'arc accentuel un rôle majeur de cohésion rythmique des groupes de mots.

Enfin, le niveau du Groupe de Souffle est le plus variable et ne semble pas constituer, même pour la parole spontanée, une unité rythmique privilégiée.

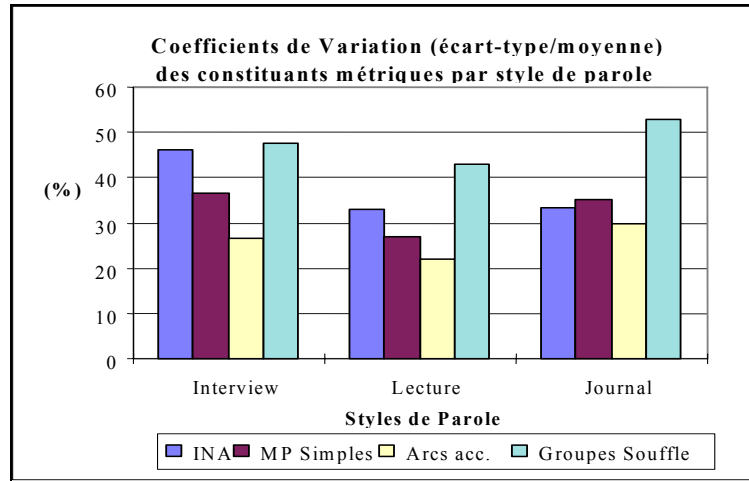


Figure 45 : Coefficients de variations (écart-types / moyennes) des constituants prosodiques par style de parole (INA = syllabes inaccentuées ; MP Simples = Mots Prosodiques contenant un accent final ; Arcs acc. = Mots Prosodiques contenant un accent final et un accent initial rythmique (AI) ou emphatique (EMP) ; Groupes de souffles = unités de performance comprises entre deux pauses respiratoires.)

Il est intéressant de noter que le style *Lecture* présente une variabilité moins importante que les styles *Journal* et *Interview* respectivement, et ce à tous les niveaux de la hiérarchie prosodique. Cette constatation nous amène à conclure que ce qui distingue a priori la parole lue de la parole spontanée, n'est pas tant le choix d'un niveau différent d'actualisation du rythme (Vaissière, 1991), mais bien la tendance pour la parole lue à exprimer une certaine régularité à tous les niveaux de la hiérarchie prosodique. Fónagy & Fónagy (1976) remarquent une plus grande variabilité dans la taille des groupes rythmiques en parole spontanée. La lecture est caractérisée, selon ces auteurs, par une propension à produire des paires accentuelles (i.e. deux groupes rythmiques identiques qui se suivent) ou des groupes rythmiques adjacents dans des proportions simples ou doubles, ce qui explique la moindre variabilité des groupes rythmiques pour ce type de discours. Ces tendances sont reflétées également dans nos propres résultats.

La différence de processus d'encodage se manifeste donc par le *degré* de régularité (étendue de la variance) et par la *superposition* des niveaux de rhythmicité.

A ce sujet, on constate que le style *Journal* se situe entre les styles *Interview* et *Lecture* sur l'échelle de métricit .

3.2. Unit s de r f rence de l'organisation rythmique

3.2.1. Syllabes inaccentu es, syllabes de r f rence

On pourrait supposer que le r le de la syllabe inaccentu e, en tant que syllabe de r f rence, pourrait  tre moindre en parole spontan e, eu  gard   la variabilit  importante observ e   ce niveau pour ce style de parole (cf. § 3.1.1 pr c dent), ou bien le contraste entre syllabes inaccentu es et accentu es pourrait  tre plus ou moins fort suivant le style de parole. Or, lorsqu'on fait le rapport entre la dur e moyenne des syllabes accentu es toutes

catégories accentuelles confondues et celle des syllabes inaccentuées par locuteur, on obtient le même résultat (environ 1.7) pour tous les styles de parole. La syllabe inaccentuée représente donc bien une valeur de référence pour la structuration rythmique des énoncés, y compris en parole spontanée, et le rapport qu'elle entretient avec les syllabes accentuées ne participe pas de la distinction rythmique entre les styles de parole. Ce chiffre (1.7) correspond exactement au résultat que Fant et al. (1991a et b) obtiennent pour des phrases lues en français.

Le Tableau 31 p. 272 donne les rapports de durées entre syllabes accentuées (toutes catégories accentuelles confondues et par catégorie accentuelles) et les syllabes inaccentuées (pour une présentation détaillée, cf. Astésano, 1999).

Si l'on compare les rapports de durée entre les syllabes inaccentuées de référence et les syllabes accentuées par catégorie accentuelle, on s'aperçoit que le style *Lecture* accentue le contraste entre les syllabes inaccentuées et les syllabes accentuées finales, particulièrement aux frontières prosodiques majeures (Unités Intonatives), comparativement aux autres styles de parole. La 'figure' des accents finaux majeurs se détache donc particulièrement sur le 'fond' des syllabes inaccentuées. Les accents finaux, générateurs de constituance, favorisent la formation de groupements. En ce sens, nous rejoignons Guaitella (1991) selon laquelle la parole lue favorise une organisation 'métrique' plus traditionnelle basée sur le groupement.

Le rapport des syllabes inaccentuées aux accents initiaux rythmiques est comparable chez l'ensemble des locuteurs. Il s'agit du rapport le plus faible. Cependant, pour le style *Lecture*, le rapport entre syllabe inaccentuée et accent initial emphatique est dans une proportion comparable à celle caractérisant l'accent initial rythmique. Cette tendance rappelle d'ailleurs les résultats que nous avons mis en évidence pour les durées syllabiques (cf § « Niveau de constituance syllabique : Conclusions partielles », p. 192). Par contre, le rapport entre accent initial emphatique et syllabe inaccentuée est plus important pour les styles *Journal* et *Interview*, notamment pour les locuteurs masculins de ces deux styles, pour lesquels le contraste avec l'accent emphatique est plus fort que le contraste avec les accents à des frontières prosodiques majeures.

Si l'on constate par conséquent une coexistence en français moderne de l'organisation temporelle traditionnelle (accentuation finale), et du rythme accentuel initial, on note néanmoins que la prédilection pour l'un ou l'autre des rythmes dépend du style de parole. Fónagy et Fónagy (1976) avaient déjà montré que, alors qu'en parole conversationnelle et en lecture d'histoire, les syllabes finales sont perçues comme plus proéminentes, en style journalistique les syllabes initiales sont perçues comme plus proéminentes dans 74 % des cas. L'utilisation de plus en plus fréquente de l'accent emphatique au début des mots par les journalistes et les politiciens serait même de moins en moins perçue dans sa fonction emphatique, mais comme un phonostyle professionnel.

Cette remarque semble s'étendre dans une certaine mesure également au style *Interview*.

Tableau 31 : Rapport de durée (ms) entre les syllabes accentuées et inaccentuées par catégorie accentuelle, en fonction des locuteurs et des styles de parole. (INA = syllabes inaccentuées ; AFS = accent final de syntagme ; UI-NT = accent final d'Unité Intonative non terminale ; UI-T = accent final d'unité intonative terminale ; EMP = accent initial emphatique ; AI = accent initial rythmique ; ACC = syllabes accentuées, toutes catégories confondues).

Locuteurs	AFS/INA	UI-NT/INA	UI-T/INA	EMP/INA	AI/INA	Total ACC/INA
Interview M	1.56	1.72	1.72	2.1	1.33	1.64
Interview F	1.59	2.2	2.1	1.54	1.2	1.71
INTERVIEW	1.57	1.95	1.9	1.8	1.25	1.65
Lecture M	1.59	2.5	2.1	1.1	1.2	1.7
Lecture F	1.7	2.4	2.7	1.2	1.3	1.87
LECTURE	1.65	2.45	2.4	1.15	1.25	1.79
Journal M	1.6	1.9	1.9	2	1.25	1.7
Journal F	1.56	2.2	1.85	1.7	1.2	1.65
JOURNAL	1.58	2.1	1.88	1.85	1.23	1.68

Les rapports de durée entre syllabes accentuées et syllabes inaccentuées indiquent que le seuil minimum de différenciation (Just Noticeable Difference) de 20% proposé par Klatt (1976) et Rossi (1972) est largement dépassé. On remarquera néanmoins que l'accent initial rythmique (AI) se situe à la limite du seuil (en moyenne à travers les locuteurs, 25%), ce qui va dans le sens des descriptions traditionnelles de ce type d'accent, comme accent essentiellement mélodique.

3.2.2. *Pauses, unités rythmiques de référence*

Comme on l'a vu dans le paragraphe précédent, les locuteurs (et par conséquent les styles de parole) ne se distinguent pas sur la base du rapport des syllabes inaccentuées avec les prééminences, toutes catégories accentuelles confondues. Ils se distinguent par contre par la prédilection pour des 'figures' de prééminences finales ou initiales.

Il semble en fait que la distinction majeure induite par les différents processus d'encodage réside dans la gestion des groupements et des pauses.

Les résultats montrent que tous les locuteurs, à l'exception du locuteur masculin du style *Interview*, adoptent des stratégies d'équilibre rythmique nettes.

L'analyse des résultats concernant le style *Journal* ainsi que la locutrice du style *Interview* montre qu'il existe une relation de durée proportionnelle entre les unités rythmiques, la 'pause + syllabe prépausale' étant deux fois plus longue que l'intervalle interaccentuel de référence ; pour les locuteurs du style *Lecture* cette relation varie du simple au triple, les pauses ayant une durée bien supérieure à celle des autres styles (environ 600 ms contre 300 ms) (pour les résultats détaillés, cf. Astésano, 1999).

Selon Fant et al. (1991a et b), la tendance à produire des unités rythmiques dans un rapport de proportionnalité est un indice de rythmicité accentuelle en français, l'isochronie accentuelle étant mise en valeur par les phénomènes de pause.

Cette tendance n'est pas uniquement l'apanage de la lecture, mais également de la parole professionnelle et en partie de la parole spontanée.

Tableau 32 : Durées des Pauses Respiratoires par locuteur (Seulement Pause pour les locuteurs du style Lecture) ; durées des Intervalles Interaccentuels définis par des AFS ; rapports entre la durée des IIA AFS et la durée de l'entité rythmique 'Pauses Respiratoires + syllabes accentuées pré-pausales'. (IIA AFS = Intervalle InterAccentuel formé par un Accent Final de Syntagme ; Durée syllabique UI = durée moyenne des syllabes finales d'Unité Intonative Non Terminale ; PR = Pauses Respiratoires)

Locuteurs	Durée Pauses Respiratoires	Durée Pauses	Durée syllabique UI	Durée entité 'Pause + UI'	Durée IIA AFS	Rapport IIA AFS/ UI + PR
Interview M	285	-	254	539	402	0.7
Interview F	323	-	299	622	330	0.5
INTERVIEW	304	-	276	580	366	0.6
Lecture M	-	544	325	869	329	0.38
Lecture F	-	756	343	1099	352	0.32
LECTURE	-	650	334	984	340	0.35
Journal M	344	-	302	646	316	0.5
Journal F	370	-	323	693	357	0.5
JOURNAL	357	-	312	669	336	0.5

3.3. Organisation temporelle des Mots Prosodiques

3.3.1. Taille des Mots Prosodiques et variabilité stylistique

Selon Fónagy (1980), Wenk & Wioland (1982), Fletcher (1991) et Delais (1994) notamment, les groupes rythmiques (ou Mots prosodiques) en français seraient relativement courts (entre trois et quatre syllabes).

Nos résultats indiquent effectivement une prédilection pour les groupes de trois et quatre syllabes. Nous notons cependant une différence dans la distribution des Mots Prosodiques selon le style de parole, manifestée essentiellement dans la proportion des groupes longs (5 syllabes et plus) (pour les résultats détaillés, cf. Astésano, 1999).

Ainsi, nous constatons une distribution plus large, notamment vers la droite, pour le style *Interview*, ce qui signifie que les locuteurs dans cette situation de communication utilisent certes une majorité de groupes tri- et quadrisyllabiques (respectivement 21,4 et 25,2% des cas), mais également une proportion non négligeable de groupes longs (de 5 à 10 syllabes). Les groupes de 5 à 7 syllabes inclus représentent 31,8%, et les groupes de 8 syllabes et plus représentent 6,7%.

* Le total des groupes rythmiques de plus de 5 syllabes est donc de **38,5%** pour le style *Interview* (parole spontanée).

Le style *Lecture* par contre présente une distribution plus gaussienne, avec une forte proportion de groupe quadrisyllabiques (31%) et une moindre proportion de trisyllabiques (20%). La taille des groupes rythmiques en *Lecture* ne dépasse pas 7 syllabes, l'ensemble des groupes de 5 à 7 syllabes ne représentant pas plus de 23,9%.

* Le style *Lecture* présente seulement **24,7%** de groupes de 5 syllabes et plus (Les groupes de 8 syllabes et plus ne représentant que 0,8%).

Le style *Journal* se trouve une fois de plus entre les deux styles *Interview* et *Lecture* sur l'échelle de métricité. Comme le style *Interview*, la distribution fréquentielle des groupes

rythmiques est assez étendue, allant jusqu'à 10 syllabes. Cependant, les groupes de 5 à 7 syllabes inclus ne représentent que 25,3%, et les groupes de 8 syllabes et plus 3,1% (moitié moins par rapport à *Interview*). Par contre, ce qui rapproche le style *Journal* du style *Lecture* est la proportion élevée de groupes bisyllabiques (respectivement 19,5% pour *Journal* et 16,5% pour *Lecture*, contre 8,4% seulement pour *Interview*).

* Le total des groupes rythmiques de plus de 5 syllabes est donc de **28,4%** pour le style *Journal*.

En conclusion, on peut dire que les 3 styles présentent la même tendance de prédilection pour les groupes tri- et quadrisyllabiques. Ces tendances correspondent aux préférences 'naturelles' pour les groupements de 4 éléments, suivis de très près par les groupements de 3 et 2 éléments (Handel, 1989). Le style *Interview* présente une distribution assez étalée vers la droite (longs groupes), et le style *Lecture* une distribution légèrement étalée vers la gauche (groupes courts). Le style *Journal* se situe entre les deux autres styles, présentant tout de même une particularité en ce qui concerne la proportion quasi égale de groupes bi-, tri- et quadrisyllabiques. Enfin, les trois styles ont une proportion comparable de monosyllabes (6,3% pour *Interview* ; 8,3% pour *Lecture* et 8,6% pour *Journal*), *Lecture* et *Journal* représentant plus ce type de groupes rythmiques que le style spontané. Ces résultats sont résumés dans le Tableau 33.

Il apparaît donc que la parole spontanée se caractérise par des groupes rythmiques plus longs que la parole lue ou interprétée. Nos résultats vont dans le sens de ceux de Fónagy (1980) et Fónagy & Fónagy (1976), dont les corpus d'analyse sont également issus de parole en situation. En ce sens, nos résultats et ceux de Fónagy & Fónagy (1976) vont à l'encontre des résultats de Lucci (1973 ; 1983) et Bruce (1995) notamment, dont on a vu qu'ils n'analysent pas de la parole en situation mais de la pseudo-parole lue et de la parole pseudo-spontanée (cf. § « Critique des méthodes traditionnelles de comparaison des types de discours », p. 119).

Tableau 33 : Pourcentages des groupes rythmiques en fonction du nombre de syllabes les constituant, par style de parole (Mono = Monosyllabiques ; Bi = Bisyllabiques ; Tri = Trisyllabiques ; Quadri = Quadrisyllabiques ; 5 à 7 = groupes de 5 à 7 syllabes inclus ; 8 à 10 = groupes de 8 à 10 syllabes inclus ; 5 à 10 = total des groupes de 5 à 10 syllabes).

Styles	Taille des Mots Prosodiques (nombre de syllabes)						
	Mono	Bi	Tri	Quadri	5 à 7	8 à 10	5 à 10
Interview	6,3	8,4	21,4	25,2	31,8	6,7	38,5
Journal	8,6	19,5	21,5	21,8	25,3	3,1	28,4
Lecture	8,3	16,5	19,8	30,6	23,9	0,8	24,7

3.3.2. Densité accentuelle

La densité accentuelle est directement liée à la taille des Mots Prosodiques. Si un style de parole présente une tendance à avoir des Mots Prosodiques en moyenne longs, ce même style de parole présentera une densité accentuelle faible.

La densité accentuelle est définie comme le pourcentage de syllabes accentuées par rapport au nombre total de syllabes.

Le Tableau 34 p. 275 indique les pourcentages de proéminences selon leur distribution

(initiale ou finale) et leur fonction (rythmique ou sémantico-pragmatique), par locuteur et par style de parole.

On remarque que les locuteurs du style *Interview* ont tendance à avoir une densité accentuelle moindre que les locuteurs des styles *Lecture* et *Journal* (en moyenne, respectivement 26.5%, 32% et 33.5%). Par conséquent, le style *Interview* possède des intervalles interaccentuels plus grands que les deux autres styles, un plus grand nombre de syllabes inaccentuées séparant deux accents.

Les styles se distinguent également par le pourcentage d'accents initiaux, caractéristiques du style journalistique professionnel (Fónagy & Fónagy, 1976). Ils représentent pour ce dernier 23.5% des prééminences en moyenne, contre 19.5% pour le style *Lecture* et 18.5% pour *Interview*.

Tableau 34 : Densité accentuelle (%) et pourcentages de syllabes accentuées par catégories accentuelles, et selon leurs fonctions (rythmiques vs. sémantico-pragmatiques) et leur distribution (initiales vs. finales). (AFS = accent final de syntagme ; AI = accent initial rythmique ; UI = accent final d'Unité Intonative terminale ou non terminale ; EMP = accent initial emphatique).

Locuteurs	Densité accentuelle	AFS	AI	Accents Rythmiques	UI	EMP	Accents Sémantico-pragmatiques	Accents Finaux	Accents Initiaux
Interview M	25	48.5	10	58.5	31.5	10	41.5	80	20
Interview F	28	51.5	10.5	62	31	7	38	83	17
INTERVIEW	26,5	50	10,2	60,2	31,2	8,5	39,7	81,5	18,5
Lecture M	33	48.5	13	61.5	28.5	10	38.5	77	23
Lecture F	31	55	7.5	62.5	28.5	9	37.5	84	16
LECTURE	32	51,7	10,2	62	28,5	9,5	38	80,5	19,5
Journal M	35	46.5	14.5	61	28.5	10.5	39	75	25
Journal F	32	51.5	16.5	68	26.5	5.5	32	78	22
JOURNAL	33,5	49	15	64,5	27	8	35,5	76,5	23,5

On observe que le type de discours a une influence sur la densité accentuelle. Or, s'il est vrai que, globalement, les types de discours ont une certaine cohérence du point de vue de la densité accentuelle globale, la répartition des syllabes accentuées par catégories accentuelles individuelles, ou regroupées selon leurs fonctions et leur distribution, ne présente aucune cohérence intrastylistique.

Il apparaît notamment que les locutrices de chaque style produisent plus d'accents finaux rythmiques (AFS) que les locuteurs ; cette remarque est particulièrement vraie pour le style *Lecture*, où l'on constate une stratégie rythmique différente pour le locuteur *Lecture M* (qui privilégie l'accentuation rythmique initiale) et pour la locutrice *Lecture F* (accentuation rythmique finale).

Les locuteurs masculins compensent dans l'ensemble par un nombre plus élevé d'accents initiaux (rythmiques et principalement emphatiques). Le pourcentage d'accents nucléaires est quasi identique pour les locuteurs et les locutrices, alors que l'accentuation initiale emphatique semble être une stratégie plus masculine. Ainsi, il apparaît que les locutrices réalisent en moyenne plus d'accents à caractères rythmiques (plus précisément finaux), alors que les locuteurs privilégient en moyenne les accents sémantico-pragmatiques

(plus particulièrement emphatiques initiaux).

3.3.3. *Taille des Mots Prosodiques en fonction des catégories accentuelles*

La catégorisation fonctionnelle des proéminences selon le modèle de Di Cristo & Hirst nous permet de mettre en évidence l'influence de la nature de l'accent sur la taille des Mots Prosodiques (MP).

Lehiste (1973) montre que la taille (en durée et nombre de syllabes) des intervalles interaccentuels augmente avec le degré de la frontière syntaxique et prosodique. Cet argument atténue la réalité objective de l'isochronie accentuelle en anglais, mais va dans le sens d'une isochronie subjective eu égard à la tendance naturelle au ralentissement du débit et de la 'cadence' en fin de groupes.

Nous observons également cette tendance en français pour les frontières prosodiques. Les MP délimités par des accents à des frontières prosodiques mineures (AFS) sont plus courts que les MP délimités par des accents de frontières prosodiques majeures non terminales (UI-NT) et, dans une moindre mesure, terminales (UI-T) (pour les résultats détaillés, cf. Astésano, 1999).

Cette distinction est significative entre les AFS et les UI-NT pour les styles *Interview* et *Journal*, mais non significative pour *Lecture* (valeur de $p = < .0001$ pour *Interview* et *Journal*, et valeur de $p = .0585$ pour *Lecture*). Pour ce dernier en effet, les trois types de MP ont des tailles équivalentes, ce qui contribue à renforcer la tendance à l'isochronie subjective. Ce résultat explique sans doute aussi pourquoi le style *Lecture* présente une régularité plus importante au niveau de constituance des Mots prosodiques (cf. § 3.1.1 « Constituance et variabilité rythmique », p. 268). Cependant, les frontières majeures sont marquées à 90% par des pauses en *Lecture*, alors qu'elles ne le sont que dans une moindre mesure pour les autres styles, leur durée étant également nettement moins importante (cf. § 3.4.1 « Nature et durée des pauses », p. 277). Il est donc probable que les pauses jouent un rôle compensatoire dans les tailles des groupes rythmiques, notamment au niveau des Mots Prosodiques de frontières majeures (Unités Intonatives). Elles rétablissent ainsi la tendance objective de groupements plus longs aux frontières majeures.

Une autre tendance se retrouve dans la taille des MP Simples vs. les MP Complexes. Ainsi, les MP Complexes délimités par un AI (accent initial rythmique) et un AFS (accent final rythmique) sont-ils significativement plus longs (en moyenne 900 ms, et 5 syllabes) que les MP simples délimités par un AFS (en moyenne 600 ms et 3.7 syllabes), et ce pour les styles *Interview* et *Journal* (valeurs de $p = < .0001$, pour les deux styles). Ce résultat renforce le rôle essentiellement eurythmique de l'AI (tendance à l'équilibrage rythmique interaccentuel, selon Fónagy, 1980). Nous centrons notre réflexion sur les MP délimités par des accents rythmiques (AFS et AI) car, de par leur fonction rythmique, ils relèvent du code linguistique, du système noyau de la langue. A ce sujet, il est intéressant de remarquer que l'intervalle interaccentuel moyen dans de nombreuses langues est d'environ 550 ms (Fant et al., 1991b). Cette durée semble correspondre au 'tempo naturel' de l'être humain (rythme biologique), qui se situe aux alentours de 600 ms selon Handel (1989). Cette valeur correspond à la durée moyenne des MP simples délimités par un accent final rythmique (AFS).

3.4. Gestion stylistique des phénomènes de pause

Les phénomènes d'interruption verbale et vocale, selon les termes de Guaïtella (1991), sont des indices particulièrement importants de l'activité cognitive des locuteurs dans diverses situations de communication (Goldman-Eisler, 1968 ; Grosjean, 1980 ; Guaïtella, 1991, notamment).

Nous nous proposons de mettre en valeur les stratégies communes et les spécificités des styles de parole eu égard aux phénomènes de pause essentiellement à travers l'étude de trois aspects :

- la nature et la durée des pauses ;
- le lien entre les phénomènes de pause et le débit de parole ;
- la corrélation entre les pauses et les phénomènes d'allongement prépausaux.

Une présentation détaillée des graphes, tableaux de moyennes et tests post hoc de Fisher des différents types de pauses, par locuteur est disponible dans Astésano (1999).

3.4.1. Nature et durée des pauses

Comme nous l'avons fait remarquer dans la partie méthodologique, nous n'avons pas pu distinguer pour le style *Lecture* les pauses vides des pauses respiratoires : nous n'avons donc repéré qu'un type de pauses silencieuses que nous qualifions de 'Pauses'. Le Tableau 35 présente le nombre et la proportion des différents types de pause par locuteur.

Tableau 35 : Types de pauses par locuteur (en occurrence et en pourcentage (%)). Les pourcentages indiquent le rapport entre les deux macro-classes 'Hésitations Vocales' et 'Pauses Silencieuses', ainsi que le rapport entre les 'Pauses Respiratoires' et les 'Pauses Vides' à l'intérieur de la macro-classe des 'Pauses Silencieuses'.

Locuteurs	Hésitations Vocales	Pauses Silencieuses	Pauses Respiratoires	Pauses Vides	Total Pauses
Interview M	1 (3%)	44 (97%)	37 (84%)	7 (16%)	45
Interview F	18 (46%)	21 (54%)	20 (95%)	1 (5%)	39
Lecture M	0	20 (100%)	-	-	20
Lecture F	0	18 (100%)	-	-	18
Journal M	0	54 (100%)	39 (72%)	15 (28%)	54
Journal F	0	39 (100%)	32 (82%)	7 (18%)	39

Les deux locuteurs de *Lecture* produisent environ le même nombre de pauses (20 pour *Lecture M* ; 18 pour *Lecture F*).

Les hésitations vocales sont l'apanage du style *Interview*. Les deux autres styles (*Lecture* et *Journal*) n'en produisent pas, de par la spécificité des processus d'encodage (lecture et lecture interprétée). A l'intérieur du style *Interview*, les deux locuteurs semblent avoir des stratégies de pause différentes : ainsi, *Interview M* ne produit-il quasiment que des pauses silencieuses (97%), alors que *Interview F* produit proportionnellement plus d'hésitations vocales (46%). Enfin, à l'intérieur de la macro-classe des pauses silencieuses, la locutrice *Interview F* produit proportionnellement plus de pauses respiratoires que le locuteur *Interview M*.

En ce qui concerne le style *Journal*, la locutrice *Journal F* produit également plus de pauses respiratoires que le locuteur *Journal M* proportionnellement. Les locutrices *Journal F* et *Interview F* produisent en proportion 10% de pauses respiratoires de plus que les locuteurs masculins.

Le Tableau 36 présente la durée des pauses par locuteur et par style de parole.

Tableau 36 : Durée des pauses par locuteur et par style de parole (en ms)

Locuteurs	Hésitations Vocales	Pauses Silencieuses	Pauses Respiratoires	Pauses Vides	Total Pauses
Interview M	196	230	285	174	218
Interview F	241	199	323	74	213
INTERVIEW	219	215	304	124	215
Lecture M	-	544	-	-	544
Lecture F	-	756	-	-	756
LECTURE	-	650	-	-	650
Journal M	-	242	344	139	242
Journal F	-	242	370	114	242
JOURNAL	-	242	357	126	242

Les locuteurs du style *Lecture* produisent les pauses silencieuses les plus longues. Puis viennent les locuteurs de *Journal*, et enfin de *Interview*.

Pour ces deux derniers styles, on voit que les pauses respiratoires sont plus longues que les pauses vides (valeur de $p = <,0001$). Pour ce qui est des pauses silencieuses, on remarque une grande cohérence à l'intérieur des styles *Journal* et *Interview*, et également entre ces deux styles (les 4 locuteurs ne se distinguent pas les uns des autres sur la base des valeurs moyennes des pauses, cf. Astésano, 1999) On remarque également la tendance pour les locutrices *Journal F* et *Interview F* à produire des pauses respiratoires plus longues (respectivement environ 11% plus longues) que celles de locuteurs *Journal M* et *Interview F*. Cette différence est compensée par une durée légèrement supérieure des pauses vides pour les locuteurs *Journal M* et *Interview M*.

Dans le style *Interview* enfin, si la locutrice *Interview F* produit des pauses respiratoires plus longues que le locuteur *Interview M*, ce dernier compense avec la durée des pauses vides, au point d'avoir des pauses silencieuses globalement plus longues que celles de *Interview F*. Cependant, dans la comparaison des deux macro-classes de pauses (Hésitations vocales et Pauses silencieuses), la locutrice *Interview F* compense par une durée moyenne des pauses sonores (hésitations vocales) bien supérieure à celle de *Interview M*. Globalement pour le style *Interview*, on s'aperçoit que les deux locuteurs ont des stratégies de pauses bien différentes, la locutrice partageant presque équitablement son temps de pause entre les pauses sonores et les pauses silencieuses, alors que le locuteur *Interview M* ne produit pratiquement que des pauses silencieuses. Cependant, la répartition en nombre et en durée s'équilibre globalement.

Il apparaît donc que la nature et la durée des pauses participent à la 'classification' des types de discours sur l'échelle de spontanéité. Ainsi, le style *Journal* partage-t-il la même stratégie globale de pause que le style *Interview* en termes de durées, mais il partage avec le style *Lecture* la caractéristique de ne pas posséder d'hésitations vocales, propres à la parole

spontanée. Par conséquent, le style *Journal* se trouve dans une position intermédiaire entre la parole spontanée et la parole lue, empruntant aux deux stratégies d'encodage.

Au regard de ces tendances, on peut se demander si les différentes stratégies pausales en termes de durée et de types de pause agissent sur le débit de parole, dont on sait qu'il participe de la distinction des styles de parole (Fónagy & Fónagy, 1976 ; Grosjean, 1980 ; Lucci, 1983).

3.4.2. *Gestion des pauses et débit de parole*

Grosjean (1980) définit les variables temporelles caractéristiques du timing de la parole. Après avoir défini un certain nombre de variables, il se propose de comparer le timing de différentes situations de communication (description de dessins et interview), afin de distinguer les stratégies spécifiques à un type de discours. Il s'agit pour lui d'obtenir des informations sur les processus cognitifs mis en jeu dans diverses tâches linguistiques. La facilité avec laquelle le locuteur va encoder ses énoncés dépend, entre autre, du nombre de contraintes imposées par la tâche linguistique. Ceci est à son tour reflété par des variables telles que le débit de parole, le nombre de pauses, le nombre d'hésitations etc. Mais à la différence de Goldman-Eisler (1968), qui ne trouvait pas de différence de temps de pause entre la description de dessin et la conversation et qui en déduisait donc que ces deux tâches linguistiques requièrent en moyenne le même effort cognitif, Grosjean montre que les deux tâches linguistiques diffèrent eu égard aux différentes variables temporelles qu'il prend en considération. Ainsi, il distingue :

- les *variables temporelles primaires*, comprenant le débit de parole (pauses incluses), le ratio phonation/temps, la vitesse d'articulation (débit hors pauses), la longueur des pauses silencieuses (ces dernières reflètent un grand nombre d'opérations difficiles à séparer et souvent concomitantes, telles que la planification, l'hésitation, la respiration, le marquage de frontières grammaticales, le marquage d'accent emphatique etc.) et enfin les longueurs des morceaux de parole (entre deux pauses),
- et les *variables temporelles secondaires*, dont la présence n'est selon lui pas obligatoire dans la parole et qui concernent tous les phénomènes d'hésitation (pauses remplies, allongements syllabiques, répétitions et faux départs).

Grosjean note que le débit de parole hors pause (vitesse d'articulation) est plus lent en lecture qu'en parole spontanée, les pauses étant également de durée moins importante en lecture. Par conséquent, ce qui distinguerait la parole spontanée de la lecture serait l'utilisation plus fréquente en parole spontanée de pauses plus longues, et ce afin de permettre la planification de l'énoncé, les pauses plus nombreuses et plus importantes étant le reflet de la complexité de l'activité cognitive. Ainsi, le débit de parole (pauses incluses) serait plus lent en parole spontanée qu'en parole lue, et ceci afin de faciliter l'encodage du message (Lucci, 1983).

Nos résultats indiquent une tendance inverse (cf. Tableau 37). En effet, si la vitesse d'articulation est légèrement plus importante en moyenne en parole spontanée (*Interview*) qu'en *Lecture*, les pauses ont peu d'influence sur le débit de parole. Le style *Journal* se situe entre les deux autres styles, en ce sens que la vitesse d'articulation est moins

importante (cf. à ce sujet les conséquences sur la durée vocalique, Figure 10 : Durées en millisecondes des voyelles et des consonnes par locuteur », p. 146), mais l'influence des pauses est également peu importante, à l'instar du style *Interview*.

Tableau 37 : Débit hors pauses (vitesse d'articulation), débit pauses incluses et influence des pauses sur la vitesse d'articulation (Rapport en pourcentage entre le débit hors pauses et le débit pauses incluses). Les résultats sont présentés par locuteur et par style de parole.

Locuteurs	Débit (hors pauses)	Débit (pauses incluses)	Rapport débit Hors Pauses/débit avec Pauses
Interview M	5.9	5.2	1.13 (13%)
Interview F	6.1	5.6	1.09 (9%)
INTERVIEW	6	5.4	1.11 (11%)
Lecture M	6.3	4.8	1.31 (31%)
Lecture F	5.6	4.2	1.33 (33%)
LECTURE	5.9	4.5	1.32 (32%)
Journal F	5.5	4.9	1.12 (12%)
Journal M	5.2	4.6	1.13 (13%)
JOURNAL	5.3	4.7	1.13 (13%)

Si l'on compare la vitesse d'articulation des locuteurs, on note tout d'abord une plus grande homogénéité dans les styles *Interview* et *Journal* que dans le style *Lecture*. On note également une légère tendance du style *Interview* à avoir une vitesse d'articulation plus importante que les styles de parole non spontanée ; la différence est notamment plus grande avec le style *Journal* professionnel.

Si l'on compare maintenant le débit des locuteurs pauses incluses, on s'aperçoit que l'influence des pauses sur le débit est plus important en *Lecture* que dans les autres styles. Ainsi, les pauses ralentissent le débit de 32% en moyenne en *Lecture*, mais seulement de 12% en moyenne pour *Journal*, et 11% pour *Interview* en moyenne.

La vitesse d'articulation plus lente en *Journal* est donc une variable plus importante qu'en *Lecture*, pour laquelle les pauses influencent grandement le débit de parole.

Le style *Lecture* se distingue donc des autres styles par la variable 'temps de pause', alors que l'*Interview* se distingue des deux autres styles par la variable 'vitesse d'articulation'. Le style *Journal* est intermédiaire.

Si les pauses en *Lecture* sont plus importantes en termes de durée et en termes de rapport rythmique aux Intervalles Inter-Accentuels moyens (cf. § « Pauses, unités rythmiques de référence », p. 272), on peut penser que les phénomènes de pauses entretiennent un rapport de corrélation plus étroit avec les autres événements prosodiques, comme les syllabes accentuées prépausales par exemple.

3.4.3. Corrélation entre les durées de pauses et les durées des syllabes accentuées précédentes

Afin d'observer le lien entre les pauses et les syllabes accentuées prépausales, dont on dit qu'elles participent du même phénomène rythmique sous-jacent (Duez, 1987, notamment), nous proposons d'analyser la corrélation entre la durée des pauses silencieuses

et la durée des syllabes prépausales, quelle qu'elles soient, par style de parole, puis par locuteur.

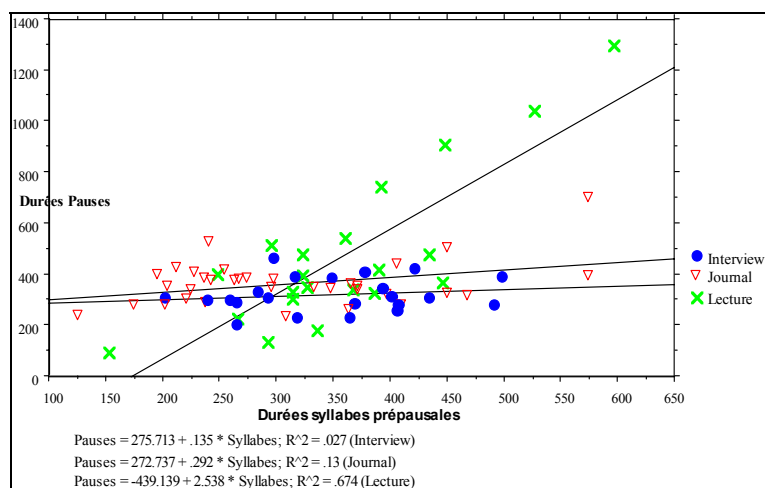


Figure 46 : Corrélation entre la durée brute (ms) des syllabes accentuées précédant les pauses, et la durée brute (ms) des pauses, par style de parole

Il apparaît que la corrélation de durée des syllabes prépausales et des pauses silencieuses est fonction du processus d'encodage. Ainsi, plus on tend vers la parole spontanée, moins la corrélation est importante.

Le style *Lecture* se démarque des deux autres styles par une forte corrélation ($R = 0.82$), suivi de *Journal* ($R = 0.36$) et *Interview* ($R = 0.16$) (cf. Figure 46)¹³.

Si l'on regarde les corrélations pour chaque locuteur (cf. Figure 47), on se rend compte que les locutrices, quel que soit le style auquel elles appartiennent, présentent une corrélation plus importante que les locuteurs. Ainsi, et en respectant la tendance générale selon laquelle la corrélation augmente en proportion inverse du caractère spontané du type de discours, la locutrice *Interview F* présente une corrélation plus forte que le locuteur *Interview M* ($R = 0.3$, contre $R = 0.07$), la locutrice *Journal F* une corrélation plus forte que le locuteur *Journal M* ($R = 0.565$, contre $R = 0.322$), et la locutrice *Lecture F* une corrélation plus forte que le locuteur *Lecture M* ($R = 0.879$, contre $R = 0.353$).

Cette tendance va de pair avec la tendance plus marquée chez les locutrices (*Journal* et *Interview* notamment) à privilégier les Pauses Respiratoires, en nombre et en durée, par rapport aux Pauses Vides, les Pauses Respiratoires représentant la grande majorité des phénomènes de pauses silencieuses. De même, la corrélation plus forte entre les pauses silencieuses et les syllabes prépausales chez les locutrices doit-elle être liée à la prédilection des locutrices pour les proéminences finales (cf. § « Densité accentuelle », p. 274).

¹³ L'indice de corrélation R correspond à la racine carrée du R^2 .

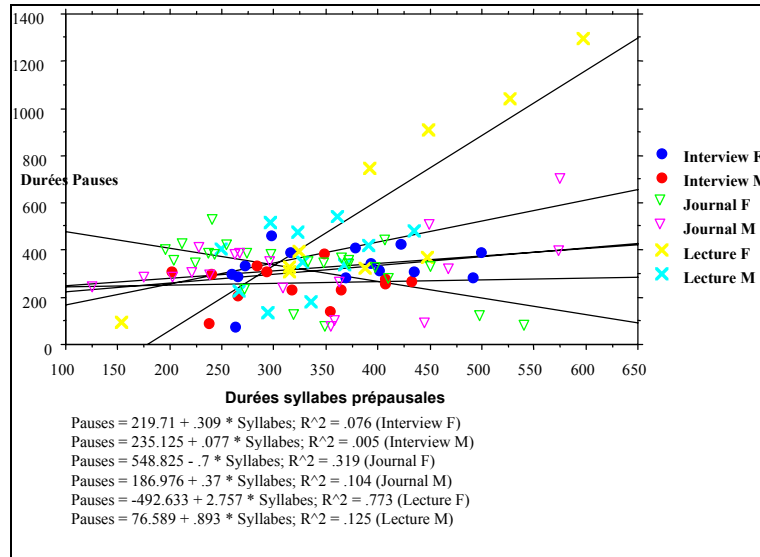


Figure 47 : Corrélation entre la durée brute (ms) des syllabes accentuées précédant les pauses, et la durée brute (ms) des pauses, par locuteur

On peut imputer ces résultats aux différentes stratégies de renforcement des groupes prosodiques par les pauses. Les styles et/ou les locuteurs qui privilégient les pauses respiratoires sur les autres types de pause, et qui tendent à marquer les pauses par une durée plus importante, présentent des corrélations plus importantes avec les syllabes accentuées précédentes. Ce serait peut être là le reflet (ou la conséquence) d'une programmation rythmique plus équilibrée.

En effet, en *Lecture*, la quasi totalité des frontières prosodiques majeures sont marquées par des pauses, ce qui participe au renforcement de frontière. Il y a donc quasi syncrétisme entre les Unités Intonatives (UI) et les Groupes de Souffle (*Lecture M* = 100% ; *Lecture F* = 83%).

En ce qui concerne le style *Journal*, plus des deux tiers des UI sont renforcées par des pauses (*Journal M* = 67% ; *Journal F* = 83%), alors que pour *Interview*, le renforcement des frontières prosodiques par des pauses concerne moins de 50% des cas (*Interview M* = 53% ; *Interview F* = 36%).

On assiste une fois de plus à une 'classification' similaire sur l'échelle de métricité, où le style *Journal* est intermédiaire entre le style *Interview* et le style *Lecture* sur la base des paramètres liés aux phénomènes de pause.

4. DISCUSSION ET CONCLUSION

Nos résultats nous permettent de mettre en évidence un certain nombre de tendances concernant l'organisation temporelle de la parole en français, et apportent des éléments de réponse aux questions posées dans notre plan expérimental.

- ***Peut-on mesurer l'isochronie ? Existe-t-il des unités rythmiques de référence dans la parole ?***

Les tendances que nous avons pu dégager nous permettent de répondre affirmativement à ces questions.

D'une part, les constituants métriques que nous avons retenus (Syllabes Inaccentuées, Mots Prosodiques Simples et Complexes) expriment le degré de régularité (isochronie) à chaque niveau de constituance, dont on a vu qu'il varie suivant les styles de parole.

D'autre part, l'équilibrage rythmique, par la relation immuable qu'entretiennent les syllabes accentuées et inaccentuées entre elles, et la relation proportionnelle qu'entretiennent les différents constituants rythmiques entre eux (Intervalle InterAccentuel moyen et entité 'Pause + syllabe prépausale), est une caractéristique commune aux trois styles de parole (cf. § « Unités de référence de l'organisation rythmique », p. 270).

Le rythme a donc une réalité phonétique indéniable, qu'il est possible de quantifier au moyen d'unités rythmiques communes aux différents types de discours.

- ***Quels niveaux de constituance sont affectés par la rythmicité ?***

Nous avons émis l'hypothèse, à la suite de Couper-Kuhlen (1993), que la rythmicité et la tendance à l'isochronie se retrouvent à tous les niveaux de constituance métrique. Cette hypothèse est vérifiée, mais doit être tempérée suivant les styles de parole.

Nous avons montré (cf. § « Constituance et variabilité rythmique », p. 268) que la tendance à la régularisation des intervalles interaccentuels participe également du rythme du français, et ceci en concurrence avec la tendance à la régularité syllabique : l'exemple le plus flagrant est le style *Lecture* où les deux aspects interagissent. Nous avons donc apporté la preuve expérimentale que les deux tendances coexistent à divers degrés et à différents niveaux dans les styles de parole : la *Lecture* présente une régularité plus importante à tous les niveaux de constituance, alors que le style *Interview*, représentatif de la parole spontanée, manifeste une certaine régularité essentiellement au niveau du MP Complexe.

Il est d'ailleurs intéressant de remarquer que, aussi bien pour le style *Lecture* que pour le style *Interview*, les Mots Prosodiques (Simples ou Complexes) sont plus réguliers que les Syllabes Inaccentuées. Par conséquent, on peut avancer que l'isochronie accentuelle constitue une rythmicité plus prégnante que l'isosyllabicit  en français.

Couper-Kuhlen (1993) note pour l'anglais que le niveau du syntagme (le Mot Prosodique) est le niveau d'isochronie privilégié en parole spontanée. Il semblerait que ce soit le cas également pour le français. Cependant, Couper-Kuhlen indique que la rythmicité est susceptible de se porter à des niveaux différents dans des séquences de parole adjacentes. Nous n'avons pas distingué, pour notre part, de *séquences* de parole, et n'avons

donc pu observer l'empan de l'isochronie à un niveau particulier de constituance. Il aurait fallu pour cela délimiter a priori une unité d'analyse discursive à l'intérieur de laquelle l'isochronie s'actualiserait à un niveau de rythmicité. Notre analyse de l'isochronie s'est faite globalement, sur un niveau paradigmatique d'observation. Cela explique sans doute une part de la variabilité observée à chaque niveau de constituance. Une analyse syntagmatique sur des empan de parole déterminés, aurait sans doute pu révéler une moindre variabilité.

Une autre remarque fondamentale que l'on peut faire à la lumière de nos résultats concerne le statut du Mot Prosodique Complexe (Arc accentuel). Pour les trois styles de parole, ce constituant présente la variabilité la moins importante par rapport aux autres constituants métriques, ce qui confère ainsi à l'Arc Accentuel un rôle rythmique majeur (intégration rythmique des unités sémantiques et lexicales). L'accent initial (rythmique ou emphatique), par la propension qu'il a à créer des arcs accentuels favorisant la constitution de groupements, contribue donc à l'émergence d'un niveau de constituance métrique indépendant : le Mot Prosodique Complexe.

- ***La rythmicité est-elle la même quel que soit le style de parole ?***

Nous avons en partie répondu à cette question dans les points développés ci-dessus. Nos résultats indiquent clairement que, si l'on peut dégager des tendances communes aux différents types de discours, la combinaison des paramètres rythmiques participe à classer les styles de parole sur une échelle de métricité.

Nous ne partageons pas le point de vue de Guaïtella (1991) selon lequel la parole lue favorise une organisation métrique (prédilection pour la loi d'assimilation) et la parole spontanée une organisation rythmique (prédilection pour la loi de dissimilation), car il sous-tend l'existence de deux systèmes différents. Pour contre, nous avons émis l'hypothèse que les styles de parole se distinguent par le degré d'assimilation et de dissimilation, manifesté au niveau des proéminences et au niveau des groupements.

Au niveau des proéminences, nous pensons que les accents initiaux (rythmiques et emphatiques) participent de la tendance à la dissimilation de par leur substance phonétique. Le style *Interview* et particulièrement le style *Journal* favorisent les proéminences initiales par rapport aux proéminences finales du point de vue de la densité accentuelle, mais également dans le rapport que les accents initiaux emphatiques entretiennent avec les syllabes inaccentuées, plus marqué que le rapport des syllabes accentuées finales aux syllabes inaccentuées. En ce sens, ces deux types de discours privilégient la tendance à la dissimilation. En revanche, le style *Lecture* est caractérisé par une forte tendance à l'assimilation. En effet, ce style privilégie les proéminences finales, tant du point de vue de la densité accentuelle que sur le plan de la mise en valeur des accents finaux sur le 'fond' constitué par les syllabes inaccentuées. Cette tendance est accentuée par un important renforcement des frontières prosodiques par les pauses, ainsi qu'une forte corrélation de durée entre les syllabes accentuées finales prépausales et les pauses.

Au niveau des groupements hiérarchiques, l'assimilation se caractérise pour le style *Lecture* par une plus grande régularité à tous les niveaux de constituance (forte métricité). Le style *Interview* se situe en revanche à l'autre bout de l'échelle de métricité, car il manifeste une plus grande variabilité à tous les niveaux de constituance (dissimilation). Cependant, il faut tempérer cette tendance eu égard à la régularité observée pour ce style de

parole au niveau des Mots Prosodiques Complexes. En ce sens, l'accentuation initiale, élément dissimilateur au niveau accentuel, contribue à constituer une figure rythmique cohésive (arc accentuel) et participe de ce fait à la tendance à l'assimilation. Le style *Journal*, enfin, est intermédiaire sur l'échelle de métricité.

Ces caractéristiques stylistiques sont cependant à pondérer par certaines particularités liées au sexe du locuteur. Nous avons vu en effet que les locutrices manifestent une prédilection pour l'accentuation finale, et ce quel que soit le style de parole. Cette tendance est renforcée par la forte corrélation de durée entre les pauses et les syllabes prépausales, nettement moindre proportionnellement pour les locuteurs masculins dans chaque style, ainsi que par les durées de pauses plus importantes chez les locutrices. Parmi les locutrices, la hiérarchie stylistique est cependant conservée, la locutrice *Lecture F* marquant plus fortement cette tendance que les locutrices *Journal F* et *Interview F* respectivement.

Il serait cependant hâtif de conclure que la variabilité rythmique observée s'explique davantage par l'appartenance à un groupe sexuel que par des caractéristiques stylistiques ou individuelles. Il serait intéressant d'étayer ces résultats par une étude portant sur des corpus plus conséquents.

En aucun cas toutefois, ces quelques particularités liées à l'appartenance sexuelle ne remettent en cause les tendances rythmiques globales observées par style de parole.

Le Tableau 38 présente un résumé des caractéristiques rythmiques des styles de parole.

Tableau 38 : Caractéristiques rythmiques des styles de parole. Le sigle φ indique les paramètres rythmiques pour lesquels il existe des particularités liées à l'appartenance sexuelle.

CARACTÉRISTIQUES RYTHMIQUES	LECTURE	JOURNAL	INTERVIEW
Régularité des niveaux de constituance	++	+	-
'Figure' accentuelle privilégiée φ	finale (assimilation)	initiale (dissimilation)	initiale (dissimilation)
Nature des pauses	silencieuses	silencieuses	silencieuses et sonores
Durée des pauses φ	+++	+	-
Influence des pauses / débit	++	-	-
Corrélation pauses / syllabes prépausales φ	+++	+	-
Densité accentuelle	+	++	-
Taille des groupes	-	+	++

Les différences entre les styles sont issues des différentes contraintes imposées par la situation de communication. D'après le tableau ci-dessus, nous constatons que le style *Journal* présente des tendances rythmiques intermédiaires à celles des types de discours spontané et lu. Ce style subit probablement les deux contraintes 'antagonistes' ce qui lui confère sa particularité prosodique. Le style *Journal* se situe par conséquent entre le style *Lecture* et le style *Interview* sur l'échelle de métricité. Il est intéressant de remarquer que l'échelle de métricité est congruente à l'échelle de spontanéité, ce qui tend à valider notre hypothèse selon laquelle le style journalistique emprunte à la fois aux stratégies

prosodiques de la parole lue et de la parole spontanée (cf. § « Style journalistique : hétérogénéité linguistique », p. 113).

En résumé, nous avons apporté la preuve de l'existence d'un système noyau en français. Notre approche méthodologique unifiée se justifiait en effet par la recherche des invariants rythmiques participant du système noyau du français, ainsi que par la mise en évidence de la part de variabilité imputable aux styles de parole.

La tendance à l'équilibrage rythmique, fonction des rapports qu'entretiennent les unités rythmiques les unes par rapport aux autres, et la mise en évidence d'une unité phonologique (Mot Prosodique Complexe ou Arc Accentuel) comme 'figure' de cohésion rythmique, sont en effet les éléments centraux d'un système noyau commun aux différents styles de parole. Nous avons montré que les différences stylistiques et la variabilité observée par rapport au système noyau ne sont qu'affaires de *degré* de rythmicité.

5. RESUME

Nous proposons de mettre en évidence par cette analyse la double identité accentuelle du français : en effet, rythme syllabique et rythme accentuel semblent coexister en français, ainsi que l'accentuation initiale et l'accentuation finale traditionnelle, ce qui confère ainsi au français sa spécificité rythmique.

Nous pensons que les principes d'assimilation et de dissimilation issus de la théorie de la Gestalt caractérisent tous les styles de parole ; par contre, le *degré* d'assimilation ou de dissimilation, au niveau accentuel (distribution accentuelle initiale ou finale) et au niveau des groupements hiérarchiques (variabilité structurelle des constituants rythmiques à différents niveaux de constituance métrique) est différent et contribue à situer les styles de parole sur une échelle de métricité.

Afin de rendre compte de la rythmicité du français, nous proposons d'analyser :

- la variabilité des niveaux de constituance métrique ;
- les rapports de rythmicité entre les événements prosodiques et des unités de référence de l'organisation rythmique (syllabes inaccentuées, pauses, et Intervalles Interaccentuels moyens) ;
- l'organisation temporelle des Mots Prosodiques en fonction des catégories accentuelles ;
- la gestion stylistique des phénomènes de pause.

Nos résultats indiquent que les styles de parole se distinguent essentiellement par la taille des Mots Prosodiques (plus longs en parole spontanée), et par la tendance pour les styles *Interview* et *Journal* à favoriser les 'figures' accentuelles initiales (dissimilation), alors que le style *Lecture* favorise les 'figures' accentuelles finales. Cette tendance est renforcée pour le style *Lecture* par la durée plus importante des pauses (imposant un débit de parole plus lent) ainsi que par l'importante corrélation de durée entre les pauses et les syllabes (finales) prépausales. Également, le style *Lecture* est caractérisé par un *degré* de rythmicité plus important que les autres styles (moins de variabilité à tous les niveaux de constituance). La combinaison des paramètres rythmiques permet donc de situer les styles sur une échelle de métricité, qui s'avère être congruente à l'échelle de spontanéité (*Lecture* - *Journal* - *Interview*).

Nos résultats indiquent également que le rythme a une réalité phonétique, quantifiable à l'aide d'unités rythmiques communes aux différents types de discours. Ainsi, nous mettons en évidence un système rythmique noyau en français, qui se manifeste par la tendance commune à l'équilibrage rythmique (rapport des syllabes inaccentuées aux syllabes accentuées ; rapport des Intervalles Interaccentuels aux entités 'Pause + syllabe prépausale'), et par l'émergence d'un niveau de constituance métrique (le Mot Prosodique Complexe ou « Arc accentuel ») constituant une figure de cohésion rythmique commune aux trois styles de parole.

Enfin, la moindre variabilité au niveau du Mot Prosodique pour les trois styles de parole nous invite à penser que l'isochronie accentuelle est plus prégnante en français que l'isochronie syllabique.

- *Conclusions et Perspectives* -

Ce travail avait pour objectif de mettre en évidence les indices prosodiques caractéristiques du rythme et de l'organisation accentuelle du français. Cette analyse était motivée par deux hypothèses :

- la validation d'un cadre fonctionnel de l'accentuation du français ;
- l'existence d'un système linguistique noyau (invariants) supposant une structure prosodique sous-jacente présente dans tout type de discours, mais dont la réalisation de surface serait contrainte par le processus d'encodage (variabilité). Cette hypothèse s'oppose à celle qui postule l'existence d'autant de systèmes prosodiques qu'il y a de styles de parole.

Nous avons choisi d'orienter notre recherche de façon à nous situer dans une approche théorique alliant à la fois certaines hypothèses de la mouvance comportementaliste (Bolinger, 1972 et suivantes ; Fraïsse, 1967 et suivantes ; Guaitella, 1991 et suivantes, notamment) et des concepts issus du courant post-généraliste (Lieberman & Prince, 1977 ; Ladd, 1996 ; Di Cristo & Hirst, 1997 ; Di Cristo, sous presse). Dans cette perspective, nous avons proposé l'analyse de plusieurs types de discours, comprenant de la parole spontanée, en adoptant un cadre théorique strict fondé sur une approche métrique de l'accentuation.

Nos analyses empiriques sur les caractéristiques tonales et temporelles du rythme nous ont permis d'obtenir un certain nombre de résultats, qui convergent vers la validation de nos hypothèses de départ. L'ensemble des résultats confirme en effet la pertinence du modèle accentuel dont nous nous inspirons. Ce dernier postule une organisation accentuelle en français plus complexe que celle proposée communément dans les descriptions traditionnelles. Ainsi, les distinctions fonctionnelles et distributionnelles des catégories accentuelles sont vérifiées. De même, certaines 'figures' rythmiques communes aux différents types de discours sont mises en évidence par nos résultats.

Nous proposons de discuter dans cette conclusion des points suivants :

- la confirmation des distinctions fonctionnelles et distributionnelles des catégories accentuelles ;
- la mise en évidence de 'figures' rythmiques communes aux différents types de discours ;
- la variabilité stylistique ;
- les perspectives de recherche à l'issue de ce travail.

◆ *Confirmation des distinctions fonctionnelles et distributionnelles des catégories accentuelles*

Sur un plan fonctionnel, la distinction entre les accents rythmiques et les accents sémantico-pragmatiques a été vérifiée aussi bien en ce qui concerne l'analyse des paramètres temporels que celle des paramètres tonals. Globalement, les catégories accentuelles sémantico-pragmatiques (accent initial emphatique ; accent final d'unité intonative terminale et non terminale) présentent des allongements syllabiques plus importants et des configurations tonales plus saillantes que les catégories accentuelles

rythmiques (accents rythmiques initiaux et finaux).

Nos résultats ont également permis de révéler les similitudes entre les catégories accentuelles d'une même macro-classe fonctionnelle. Par exemple, l'analyse des configurations tonales a confirmé le rôle rythmique des accents initiaux et finaux de mot (AI et AFS). Ces prééminences sont en effet caractérisées par des configurations tonales s'étalant sur un large empan syllabique. Les durées syllabiques de ces deux types d'accents sont de surcroît similaires, la différence étant marquée au niveau infrasyllabique, ce qui justifie par là même notre analyse en constituants syllabiques. Enfin, l'accent rythmique initial (AI) est réalisé lorsque les groupes rythmiques (Mots Prosodiques) sont trop longs (plus de 4 syllabes), confirmant ainsi le rôle de cet accent dans l'équilibre rythmique des groupements (règles d'eurythmie).

Les catégories accentuelles sont également distinguées sur la base de leur distribution. Sur le plan tonal, les prééminences initiales et finales se distinguent par la forme générale de leurs configurations tonales (asymétriques et symétriques, respectivement). Nous interprétons l'asymétrie des configurations tonales des prééminences initiales comme l'indication de leur rôle de démarcation des unités lexicales, ainsi que de leur participation à la formation des arcs accentuels. Sur le plan de la durée des constituants infrasyllabiques, les prééminences initiales et finales se distinguent par la 'configuration' de l'allongement différentiel des segments (Attaque allongée et Rime allongée, respectivement).

Le cas des prééminences initiales rythmiques et emphatiques mérite une attention particulière. Bien qu'elles présentent une parenté certaine sur le plan des 'figures' temporelles et tonales, nos résultats suggèrent qu'elles sont significativement distinguées sur la base de la dynamique de leurs variations. C'est ainsi que la prééminence de type emphatique est significativement plus allongée que la prééminence de type rythmique. D'autre part, sa configuration tonale est en moyenne deux fois plus saillante. Nos résultats confirment par là même l'hypothèse formulée par Di Cristo (sous presse), selon laquelle la prééminence initiale emphatique est l'expression 'Hyper' de l'accent initial rythmique.

Ces tendances confirment empiriquement le bien-fondé du principe de subordination accentuelle tel que le définit Di Cristo (sous presse) :

« (...) Nous supposons également que cette organisation métrique est soumise à un autre principe, le Principe de Subordination Accentuelle qui attribue à l'élément terminal désigné d'un constituant prosodique un niveau d'accentuation correspondant au rang qu'il occupe dans la représentation phonologique. Ce second principe est le moteur de la construction de la grille métrique (...) » (Di Cristo, sous presse: 13).

◆ **Mise en évidence de 'figures' rythmiques communes aux différents types de discours**

L'analyse de différents styles de parole confirme non seulement la robustesse de notre cadre d'analyse accentuel, mais permet également d'identifier des structures rythmiques communes, qui reflètent un système noyau sous-jacent. Pour les trois styles de parole que nous avons retenus, nous mettons en évidence des rapports de durée entiers entre les structures (ou 'figures') rythmiques représentées par les Intervalles Interaccentuels moyens et l'entité 'Pause + syllabe prépausale'. A ce sujet, il est intéressant de noter que la durée de

l'Intervalle Interaccentuel moyen (550 ms) - observée également par d'autres auteurs (Fant et al., 1991 a et b ; Fletcher, 1991) - correspond à la durée naturelle de prédilection du rythme biologique (Handel, 1989). En ce sens, on peut penser que les contraintes d'équilibrage rythmique révélées par nos résultats se combinent avec des contraintes biologiques universelles indépendantes des contraintes rythmiques liées au code linguistique.

Enfin, il ressort également de nos analyses que la syllabe inaccentuée semble constituer un 'fond' sur lequel les syllabes accentuées se détachent ('figures'), dans une proportion similaire à travers les styles de parole (rapport de 1.7, identique aux résultats de Fant et al., 1991a).

Ces principes rythmiques fondamentaux communs aux différents types de discours sont créateurs de structure et sont garants de l'impression de périodicité, fondements mêmes de la théorie de la Gestalt.

De manière plus générale, nous nous sommes efforcé dans ce travail à mettre en valeur les relations entre proéminence et constituance en français, notamment à travers l'émergence du Mot Prosodique Complexe en tant qu'unité rythmique de référence dans l'organisation prosodique du français. Dans le prolongement des travaux de Fónagy (1980) et à l'instar de Di Cristo (sous presse), nous avons en effet montré que l'Arc Accentuel est une 'figure' accentuelle prépondérante en français, caractérisant à la fois les styles de parole lue et spontanée. Nous avons donc non seulement montré que l'accentuation initiale fait partie intégrante du système accentuel du français, mais qu'elle est de surcroît centrale dans une description des propriétés métriques de cette langue.

Outre la confirmation de l'existence d'un système noyau, nos résultats indiquent que la variabilité stylistique observée ne concerne en réalité que les *dimensions* des paramètres du rythme, et non le *code* linguistique lui-même.

◆ *Variabilité stylistique*

Ainsi, la variabilité stylistique ne concerne que le *degré* de rythmicité exprimé en fonction des unités rythmiques de référence, permettant de situer les styles de parole sur une échelle de métricité en relation étroite avec l'échelle de spontanéité. La variabilité stylistique observée au niveau des durées segmentales ne peut d'ailleurs s'expliquer qu'en relation avec les dimensions de variation sur l'échelle de métricité.

Nous avons émis l'hypothèse que les trois styles de parole retenus dans le cadre de cette recherche peuvent être situés sur une échelle de spontanéité (*Lecture - Journal - Interview*). Le style *Journal* serait ainsi intermédiaire car, de par l'hétérogénéité linguistique due à la spécificité de ce processus d'encodage, il emprunterait à la fois les caractéristiques accentuelles et rythmiques de la parole lue et de la parole spontanée. Il est intéressant de souligner que la position intermédiaire qu'occupe le style *Journal* est de nature à conforter à la fois notre interprétation des caractéristiques rythmiques de chaque style de parole, mais également l'hypothèse de l'existence d'un système prosodique noyau.

Nos résultats nous permettent de dégager les tendances suivantes :

- Le style *Lecture* privilégie le marquage de l'accentuation finale (assimilation), aussi bien sur le plan paradigmatique (durées syllabiques) que sur le plan syntagmatique ('figure' accentuelle finale, renforcée par une importante corrélation entre les durées des pauses et les syllabes finales prépausales).
- Le style *Interview* privilégie le marquage de l'accentuation initiale (dissimilation) sur le plan paradigmatique (durées syllabiques) et syntagmatique ('figure' accentuelle initiale). Il faut ajouter à cela une très faible corrélation des durées de pauses et des syllabes finales prépausales, ce qui diminue d'autant la prégnance de l'accentuation finale dans ce style de parole.
- Le style *Journal*, enfin, indique un marquage équilibré des accents initiaux et des accents finaux, aussi bien sur le plan des durées syllabiques que de l'organisation temporelle. D'une part, ce style favorise la 'figure' accentuelle initiale, à l'instar de la parole spontanée. D'autre part, il produit une proportion non négligeable de groupes rythmiques (ou Mots Prosodiques) longs (parole spontanée), tout en privilégiant les groupes mono et bisyllabiques (parole lue). Enfin, il présente une corrélation non négligeable entre les durées des pauses et les syllabes prépausales (intermédiaire entre parole lue et parole spontanée). En d'autres termes, le style *Journal* présente de manière optimale les caractéristiques accentuelles duales du français, telles que nous les avons décrites (coexistence de l'accentuation initiale et de l'accentuation finale en français ; cf. § « Coexistence de tendances rythmiques duales et niveaux de rythmicité », p. 35).

◆ *Perspectives de recherche*

Au terme de cette recherche qui porte uniquement sur l'étude des indices acoustiques de l'accentuation et du rythme, nous ressentons la nécessité d'aborder une seconde étape concernant la validation perceptive de nos résultats. Nous souhaiterions notamment, en utilisant les paradigmes expérimentaux mis en œuvre par la psycholinguistique, évaluer le rôle de l'accentuation et du rythme dans l'accès au lexique et dans la compréhension des messages en français. Il serait en effet intéressant de voir dans quelle mesure l'accentuation et la structuration rythmique facilitent le décodage des messages, notamment en ce qui concerne la rapidité d'accès au sens. Cette question est d'autant plus intéressante qu'une synthèse récente sur cette problématique souligne l'absence de travaux relatifs aux langues à accentuation fixe comme le français (Cutler et al., 1997).

Plus particulièrement, nous envisageons de concentrer notre recherche sur le rôle de l'accentuation initiale dans l'accès au sens en français. Fónagy (1980) et Di Cristo (sous presse) indiquent que l'accentuation initiale marque les mots lexicaux importants sur le plan sémantique ou les groupes de mots formant une unité sémantique cohésive. Ce type d'accentuation est d'ailleurs particulièrement représentatif du style journalistique, à forte visée communicative, mais se retrouve également dans une grande mesure en parole spontanée. Kohler (1991) note à ce sujet que les différents processus d'encodage induisent essentiellement une différence dans le choix d'une unité rythmique de prédilection : le mot serait une entité plus importante en parole spontanée qu'en parole lue ou en poésie, où la structuration rythmique au delà du mot domine.

Dans cette optique, il se pourrait que l'accentuation initiale ait une fonction non

négligeable dans l'identification des unités lexicales. Il est probable qu'en parole spontanée où il n'existe pas, par définition, de structure textuelle préétablie, les locuteurs aient recours à ce type d'accentuation et aux figures accentuelles qu'elles constituent (les Arcs Accentuels, intégrateurs rythmiques d'unités lexicales et sémantiques), afin de faciliter l'accès au sens. Des recherches récentes en psycholinguistique suggèrent d'ailleurs que la tendance à segmenter le flot de parole à l'initiale des mots lexicaux serait « *une propriété générale des mécanismes de reconnaissance des mots, indépendamment de la langue* », que l'on retrouverait aussi bien en français qu'en anglais (Dumay et al., 1998 : 35). Il serait utile de confronter les résultats issus des recherches en psycholinguistique avec nos propres résultats qui mettent en évidence une stratégie d'accentuation initiale en français. A ce titre également, il serait intéressant de proposer une comparaison entre la prosodie de l'anglais et du français. En effet, nos résultats mettent en évidence l'existence d'une isochronie accentuelle en français sans doute plus prégnante que l'isochronie syllabique, au moins sur le plan phonétique. De même, nous avons montré que l'accentuation initiale, tant emphatique que rythmique, utilisée dans tous les styles de parole, présente une parenté phonétique avec l'accentuation lexicale de l'anglais et participe de l'impression de récurrence régulière des proéminences. Cette dualité rythmique et accentuelle du français nous invite à procéder à des analyses perceptives comparatives entre l'accentuation du français et de l'anglais en procédant à différentes expériences perceptives (tests de désémantisation notamment). Cela nous permettrait de relancer le débat sur les caractéristiques rythmiques à la base de la distinction entre le français et l'anglais, langues traditionnellement opposées sur le plan prosodique.

- *Bibliographie* -

1. Abercrombie, D. (1964) "A phonetician's view of verse structure", *Linguistics*, Vol. 6, 5-13.
2. Allen, G.D. (1975) "Speech rhythm: its relation to performance universals and articulatory timing", *Journal of Phonetics*, Vol. 3, 75-86.
3. Allen, G.D.; Hawkins, S. (1980) "Phonological rhythm: definition and development", in *Child Phonology, Vol. I: Production*, Academic Press, New York, 227-256.
4. Anderson, S.R. (1984) "A metrical interpretation of some traditional claims about quantity and stress", in *Language Sound Structure*, MIT Press, M. Aronoff and R.T. Oehrle (eds.), Cambridge, Mass.
5. Aronoff, M.; Oehrle, R.T. (1984) *Language Sound Structure*, MIT Press, Cambridge, Mass.
6. Arvaniti, A.; Ladd, D.R.; Mennen, I. (1998) "Stability of tonal alignment: the case of Greek prenuclear accents", *Journal of Phonetics*, Vol. 26, 3-25.
7. Astésano, C. (1993) *Étude des interférences rythmiques d'un apprenant anglophone en français et implications didactiques*, Mémoire de DEA, Université d'Aix-Marseille I.
8. Astésano, C. (1998a) "Effects of prosodic constraints on the differential lengthening of syllable constituents in French: a comparison between spontaneous and read speech", *Actes du congrès SPoSS, les sons de la parole spontanée*, La Baume-les-Aix, 24-26 septembre 1998, 143-146.
9. Astésano, C. (1998b) "Différents styles de parole, différents niveaux de rythmicité", *Actes du Colloque « Oralité et Gestualité. Communication multimodale, Interaction »*, Santi, S.; Guaïtella, I.; Cavé, Ch.; Konopczynski, G. (eds.), *Colloque Orage*, 9-11 décembre 1998, Besançon, L'Harmattan, Paris, 297-303.
10. Astésano, C. (1999) *Rythme et Discours : invariance et sources de variabilité des phénomènes accentuels en français*, Thèse de Doctorat, Université de Provence, Aix-en-Provence, 430 pp.
11. Astésano, C.; Bertrand, R.; Brousseau, M.; Chafcouloff, M.; Di Cristo, A.; Ghio, A.; Hirst, D.; Lapierre, S.; Nicolas, P.; Roméas, P.; Sabio, F.; Vincent, M. (1995a) "Le projet PACOMUST, un corpus de PArole COntinue MUltiSTyle: objectifs et choix méthodologiques", *Travaux de l'Institut de Phonétique d'Aix*, Vol. 16, Aix-en-Provence, 9-38.
12. Astésano, C.; Di Cristo, A.; Hirst, D.J. (1995b) "Discourse-based empirical evidence for a multi-class accent system in French", *Proceedings of 13th International Congress of Phonetic Sciences*, Vol. 4, Stockholm, Suède, 630-633.

13. Astésano, C.; Espesser, R.; Hirst, D. J.; Llisterri, J. (1997a) "Stylisation automatique de la fréquence fondamentale : une évaluation multilingue", *4^o Congrès Français d'Acoustique*, 14-18 Avril, Marseille, Vol. 1, 441-443.
14. Astésano, C.; Espesser, R.; Flachaire, E.; Nicolas, P. (1997b) "Tonal configurations of prominence in French", *ESCA Workshop on Intonation*, Athènes, Grèce, 18-20 septembre 1997, 29-32.
15. Astésano, C.; Espesser, R.; Flachaire, E.; Nicolas, P. (1998) "Analyse multi-paramétrique des configurations tonales des proéminences en français", *XXII^{èmes} Journées d'Etude sur la Parole*, Martigny, Suisse, 95-98.
16. Aubergé, V. (1991) *La synthèse de la parole: des règles aux lexiques*, Thèse de Doctorat, Université de Grenoble.
17. Auer, P. (1990) "'Stress-Timing' vs. 'Syllable-Timing' from a Typological Point of View", *Proceedings of LP'90*, B. Palek and P. Janota, Prague, 292-305.
18. Autesserre, D.; Pérennou, G.; Rossi, M. (1989) "Methodology for the Transcription and Labeling of a Speech Corpus", *Journal of the International Phonetic Association*, 19 (1), 2-15.
19. Bacri, N.; Banel, M.H. (1993) "Rhythmic patterns and lexical parsing in French", *Working Papers*, Vol. 41, *ESCA Workshop on Prosody*, Lund, 120-123.
20. Bagshaw, P.C. (1994) *Automatic prosodic analysis for computer-aided pronunciation teaching*, Thèse non publiée, Université d'Edinburgh, 247 pp.
21. Bailly, G.; Benoit, C.; Sawallis, T.R. (1992) *Talking Machines. Theories, Models and Designs*, G. Bailly; C. Benoit & T.R. Sawallis, (eds.), Elsevier Science Publishers, North-Holland, Amsterdam.
22. Bally, Ch. (1926) "Le rythme linguistique et sa signification sociale", *Compte-Rendu du Premier Congrès du Rythme*, Albert Pfrimmer (ed.), Genève, 253-263.
23. Barbosa, P.; Bailly, G. (1994) "Characterisation of rhythmic patterns for text-to-speech synthesis", *Speech Communication*, Vol. 15 (1-2), 127-137.
24. Barry, W.J. (1995) "Phonetics and Phonology of Speaking Styles", *Proceedings of the XIIIth ICPHS*, Vol. 2, Stockholm, 4-10.
25. Bartkova, K. (1988) "On the Use of Segmental Duration in Speaker-Independent Speech Recognition Systems", *Proceedings of Speech '88, 7th FASE Symposium*, Edinburgh, 763-770.
26. Bartkova, K.; Sorin, C. (1987) "A model of segmental duration for speech synthesis

- in French", *Speech Communication*, Vol. 6 (3), 245-260.
27. Beckman, M.E. (1992) "Evidence for Speech Rhythms across Languages", in *Speech Perception, Production and Linguistic Structure*, Y. Tohkura; Vatikiotis-Bateson, E.; Sagisaka, Y. (eds.), Tokyo, 457-463.
 28. Beckman, M.E. (1993) "Modelling the Production of Prosody", *Working Papers*, Vol. 41, *ESCA Workshop on Prosody*, Lund, 258-263.
 29. Beckman, M.E. (1997) "A Typology of Spontaneous Speech", in *Computing Prosody: Computational Models for Processing Spontaneous Speech* (A collection of papers from the Spring '95 Workshop on Computational Approaches to Processing the Prosody of Spontaneous Speech. Kyoto, Japan), Y. Sagisaka; N.W. Campbell; N. Higushi (eds.), Springer Verlag, New York, 7-26.
 30. Beckman, M.E.; Edwards, J. (1992) "Intonational Categories and the Articulatory Control of Duration", in *Speech Perception, Production and Linguistic Structure*, Y. Tohkura; Vatikiotis-Bateson, E.; Sagisaka, Y. (eds.), Tokyo, 359-375.
 31. Behne, D.M. (1989) *Acoustic Effects of Focus and Sentence Position on Stress in English and French*, PhD Dissertation, University of Wisconsin, Madison, USA., 182 p.
 32. Bell, A. (1977) "Accent placement and perception of prominence in rhythmic structures", in *Studies in stress and accent*, Vol. 4, L. Hyman (ed.), Southern California Occasional Papers in Linguistics, 1-13.
 33. Bell-Berti, F. (1991) "Comments on 'some observations on the organisation and rhythm of speech'", *Proceedings of the XIIth ICPHS*, Vol.1, Université de Provence, Aix-en-Provence, 238-242.
 34. Benguerel, A.P. (1971) "Duration of French Vowels in Unemphatic Stress", *Language and Speech*, Vol. 14 (4), 383-391.
 35. Benguerel, A.P. (1973) "Corrélat physiologiques de l'accent en français", *Phonetica*, 27, 21-35.
 36. Benguerel, A.P.; D'Arcy, J. (1986) "Time-warping and the perception of rhythm in speech", *Journal of Phonetics*, Vol. 14, 231-246.
 37. Bertinetto, P.M. (1989) "Reflections on the dichotomy 'stress' vs. 'syllable-timing'", *Revue de Phonétique Appliquée*, Vol. 91-93, 99-130.
 38. Blaauw, E. (1992) "Phonetic differences between read and spontaneous speech", *Proceedings of ICSLP 1992*, Vol. 1, Banff, Canada, 751-754.

39. Blanche-Benveniste, C.; Jeanjean, C. (1987) *Le français parlé. Édition et Transcription*, Paris, Didier Érudition.
40. Blevins, J. (1994) "The Syllable in Phonological Theory", in *The Handbook of Phonological Theory*, J. Goldsmith (ed.), Oxford, Blackwell, 206-244.
41. Bolinger, D.L. (1951) "Intonation: levels versus configurations", *Journal of the linguistic circle of New York*, Vol. 7 (3/1), 199-210.
42. Bolinger, D.L. (1972) "Accent is predictable (if you're a mind-reader)", *Language*, Vol. 48, 633-644.
43. Bolinger, D.L. (1981) *Two kinds of vowels, Two kinds of rhythm*, Indiana University Linguistics Club, Bloomington, Indiana, USA.
44. Bolinger, D.L. (1985) *Intonation and its parts. Melody in spoken English*, Edward Arnold Ltd, London.
45. Bolinger, D.L. (1987) "Two views of Accent", in *On Accent*, Indiana University Linguistics Clubs, Bloomington, 51-107.
46. Bolinger, D.L. (1989) *Intonation and its uses: Melody in Grammar and Discourse*, Edward Arnold Ltd, London.
47. Borel-Maisonny, S. (1967) "Perturbation du rythme dans la parole", *Colloque sur les rythmes*, Lyon, Institut d'audiophonologie de Lyon, SIMEP-Editions, 181-197.
48. Boudreault, M. (1968) *Rythme et mélodie de la phrase parlée en France et au Québec*, Kliensieck, Paris.
49. Brown, G. (1983) "Prosodic Structure and the Given/New Distinction", in *Prosody: Models and Measurements*, A. Cutler & D.R. Ladd (eds.), Heidelberg, 67-77.
50. Brown, G.; Yule, G. (1983) *Discourse Analysis*, Cambridge University Press.
51. Bruce, G. (1977) *Swedish word accents in sentence perspective*, Lund: Gleerup.
52. Bruce, G. (1980) "Tonal and temporal interplay", *Nordic Prosody*, Vol. 2, T. Fretheim (ed.), 63-74.
53. Bruce, G. (1985) "Structures and function of prosody", *Proceedings of the French-Swedish Seminar on Speech*, Guérin, B. & Carré, R. (eds.), Grenoble, 549-559.
54. Bruce, G. (1995) "Modelling Swedish intonation for read and spontaneous speech", *Proceedings of the XIIIth ICPHS*, Vol. 2, KTH and Stockholm University, Stockholm, 28-35.

55. Bruce, G.; Granström, B.; House, D. (1992) "Prosodic phrasing in Swedish speech synthesis", in *Talking Machines: Theories, Models and Designs*, G. Bailly ; C. Benoit & T.R. Sawallis (eds.), Elsevier Science Publishers, North-Holland, Amsterdam, 113-125.
56. Buxton, H. (1983) "Temporal Predictability in the Perception of English Speech", in *Prosody: Models and Measurements*, A. Cutler & D.R. Ladd (eds.), Heidelberg, 111-121.
57. Cambier-Langeveld, T. ; Nespor, M.; van Heuven, V. J. (1997) "The domain of final lengthening in production and perception in Dutch", *Proceedings of the 5th European Conference On Speech Communication and Technology*, Vol.2, ESCA, Rhodes, Greece, 22-25 Septembre 1997, 931-934.
58. Campbell, W.N. (1992) "Syllable-based segmental duration", in *Talking Machines: Theories, Models and Designs*, G. Bailly ; C. Benoit & T.R. Sawallis (eds.), Elsevier Science Publishers, North-Holland, Amsterdam, 211-224.
59. Campbell, W.N. (1993) "Durational Cues to Prominence and Grouping", *Working Papers*, Vol. 41, *ESCA Workshop on Prosody*, Lund., 38-41.
60. Campbell, W.N. (1995a) "From read speech to real speech", *Proceedings of the XIIIth ICPHS*, Vol. 2, KTH and Stockholm University, Stockholm, 20-27.
61. Campbell, W.N. (1995b) "Loudness, Spectral Tilt, and Perceived Prominence in Dialogues", *Proceedings of the XIIIth ICPHS*, Vol. 3, Stockholm, 676-679.
62. Campbell, W.N. (sous presse) "Timing in Speech: a multi-level process", in *Prosody: Theory and Experiment*, M. Horne (ed.), Dordrecht, Kluwer Academic Publishers.
63. Campbell, W.N.; Isard, S.D. (1991) "Segment durations in a syllable frame", *Journal of Phonetics*, Vol. 19, 37-47.
64. Charaudeau, P. (1984) "Description d'un genre: l'interview", in *Aspects du discours radiophonique*, Collection 'Langages, Discours et Sociétés', n°1, P. Charaudeau (ed.), Didier Érudition, 111-115.
65. Charaudeau, P. (1984) *Aspects du discours radiophonique*, Collection 'Langages, Discours et Sociétés', n°1, P. Charaudeau (ed.), Didier Érudition
66. Chomsky, N.; Halle, M. (1968) *The Sound Pattern of English*, New York: Harper & Row.
67. Clynes, M. (1982) *Music, Mind and Brain. The neuropsychology of Music*. 'An expanded version of proceedings from the Third Workshop on the Physical and

- Neuropsychological Foundations of Music, held in August 8-12, 1980, in Ossiach, Austria', Manfred Clynes (ed.), Plenum Press. New York.
68. Couper-Kuhlen, E. (1993) *English Speech Rhythm : Form and function in everyday verbal interaction*, John Benjamins Publishing Company, Amsterdam.
 69. Crompton, A. (1980) "Timing Patterns in French", *Phonetica*, Vol. 37(4), 205-234 .
 70. Crystal, D. (1980) "The Analysis of Nuclear Tones", in *The Melody of Language: Intonation and Prosody*, Linda R. Waugh & C.H. van Schooneveld (eds.), University Park Press, Baltimore, 55-70.
 71. Crystal, T.H.; House, A.S. (1982) "Segmental durations in connected speech signals: Preliminary results", *Journal of the Acoustic Society of America*, Vol. 72 (3), 705-716.
 72. Cutler, A. (1980) "Syllable omission errors and isochrony", in *Temporal Variables in Speech: Studies in Honour of Frieda Goldman-Eisler*, Hans W. Dechert & M. Raupach (eds.), Mouton, The Hague, 183-190.
 73. Cutler, A. (1983) "Speakers' Conceptions of the Function of Prosody", in *Prosody : Models and Measurements*, A. Cutler & D.R. Ladd (eds.), Heidelberg, 79-91.
 74. Cutler, A. (1984) "Stress and Accent in Language Production and Understanding", in *Intonation, Accent and Rhythm: Studies in Discourse Phonology*, D. Gibbon and H. Richter (eds.), Walter de Gruyter, Berlin, 77-90.
 75. Cutler, A. (1992) "The production and perception of word boundaries", in *Speech Perception, Production and Linguistic Structure*, Y. Tohkura; Vatikiotis-Bateson, E.; Sagisaka, Y. (eds.), Tokyo, 419-425.
 76. Cutler, A.; Ladd, D.R. (1983) *Prosody : Models and Measurements*, Heidelberg.
 77. Cutler, A.; Dahan, D.; van Donselae, W. (1997) "Prosody in the comprehension of spoken language : a literature review", *Language and Speech*, Vol. 40 (2), 141-201.
 78. Dabrowska, J. (1967) "Le rythme de l'expression en langue française et polonaise", *Colloque sur les rythmes*, Lyon, Institut d'audiophonologie de Lyon, SIMEP-Éditions, 291-305.
 79. Dahan, D.; Bernard, J.M. (1996) "Interspeaker Variability in Emphatic Accent Production in French", *Language and Speech*, Vol. 39 (4), 341-374.
 80. Darwin, C.J. (1975) "On the dynamic Use of Prosody in Speech Perception", *Status Report on Speech Research*, Vol. 42/43, Haskins Laboratories, 103-115.

81. Dauer, R.M. (1983) "Stress timing and syllable timing reanalysed", *Journal of Phonetics*, Vol. 11, 51-62.
82. Dauer, R.M. (1987) "Phonetic and phonological components of language rhythm", *Proceedings of the XIth ICPHS*, Vol. 5, Tallin, 447-450.
83. Dechert, H ; Raupach, M. (1980) *Temporal Variables in Speech: Studies in Honour of Frieda Goldman-Eisler*, H. W. Dechert & M. Raupach (eds.), Mouton, The Hague.
84. Delais, E. (1994) "Prédiction de la variabilité dans la distribution des accents et les découpages prosodiques en français", *Actes des XX^{èmes} Journées d'Études sur la Parole*, Trégastel, 379-384.
85. Delattre, P. (1962) "Some Factors of Vowel Duration and their Cross-Linguistic Validity", *Journal of the Acoustic Society of America*, Vol. 34, 1141-1143.
86. Delattre, P. (1965) *Comparing the prosodic features of English, German, Spanish and French*, Chilton Books, New York.
87. Delattre, P. (1966a) *Studies in French and comparative phonetics*, Mouton, The Hague.
88. Delattre, P. (1966b) "A comparison of syllable length conditioning among languages", *International Review of Applied Linguistics*, Vol. 4, 183-198.
89. Delattre, P. (1966c) "Les dix intonations de base du français", *the French Review*, Vol. 41 (3), 326-339.
90. Delattre, P. (1981) *Studies in Comparative Phonetics: English, German, Spanish and French*, B. Malmberg (ed.), Julius Groos Verlag, Heidelberg.
91. Dell, F. (1984) "L'accentuation dans les phrases en français", in *Formes sonores du langage*, F. Dell; D.J. Hirst; Vergnaud, J.R. (eds.), Hermann, 65-122.
92. Dell, F. ; Hirst, D.J. ; Vergnaud, J.R. (1984) *Formes sonores du langage*, F. Dell; D.J. Hirst; Vergnaud, J.R. (eds.), Hermann.
93. Di Cristo, A. (1981) "L'intonation est congruente à la syntaxe: une confirmation.", in *L'intonation: de l'acoustique à la sémantique*, Rossi, M.; Di Cristo, A.; Hirst, D.; Martin, Ph.; Nishinuma, Y. (eds.), Klincksieck, Paris, 272-289.
94. Di Cristo, A. (1985) *De la Microprosodie à l'Intonosyntaxe*, Publications de l'Université de Provence.
95. Di Cristo, A. (sous presse) "Le cadre accentuel du français: essai de modélisation", à

paraître dans *Langues*.

96. Di Cristo, A.; Hirst, D.J. (1993) "Rythme syllabique, rythme mélodique et représentation hiérarchique de la prosodie du français", *Travaux de l'Institut de Phonétique de Phonétique d'Aix*, Vol. 15, 9-24.
97. Di Cristo, A.; Hirst, D.J. (1996) "Vers une typologie des unités intonatives du français", *Actes des XXI^{èmes} Journées d'Étude sur la Parole*, Avignon, 10-14 juin 1996, 219-222.
98. Di Cristo, A.; Hirst, D.J. (1997) "L'accentuation non emphatique en français: stratégies et paramètres", *Polyphonie pour Ivan Fónagy*, J. Perrot (ed.), Paris, l'Harmattan, 71-101.
99. Ducrot, O.; Schaeffer, J.M. (1995) *Nouveau dictionnaire encyclopédique des sciences du langage*, Éditions du Seuil.
100. Duez, D. (1987) "Hiérarchisation des paramètres acoustiques et identification des frontières", *Actes des XVI^{èmes} Journées d'Étude sur la Parole*, 221-223.
101. Duez, D.; Nishinuma, Y. (1985) "Le rythme en français: alternance des durées syllabiques", *Travaux de l'Institut de Phonétique d'Aix*, Vol. 10, 151-169.
102. Dumay, N. ; Banel, M.H. ; Frauenfelder, U.H ; Content, A. (1998) "Le rôle de la syllabe : segmentation lexicale ou classification ?", *Actes des XXII^{èmes} Journées d'Études sur la Parole*, Martigny, Suisse, 33-36.
103. Durand, J. (1990) *Generative and Non-Linear Phonology*, London, Longman.
104. Eskenazi, M. (1993) "Trends in Speaking Styles Research", *Eurospeech*, Vol.1, Madrid, 501-508.
105. Espesser, R. (1982) "Un système de détection du voisement et de f0", *Travaux de l'Institut de Phonétique d'Aix*, Vol. 8, 241-261.
106. Espesser, R. (1996) "MES: un environnement de traitement du signal", *Actes des XXI^{èmes} Journées d'Études sur la Parole*, Avignon, 10-14 Juin 1996, 447.
107. Fant, G. (1991) "Units of temporal organization: stress groups versus syllables and words", *Proceedings of the XIIth ICPHS*, Vol. 1, Université de Provence, Aix-en-Provence, 247-250.
108. Fant, G.; Nord, L.; Kruckenberg, A. (1987) "Segmental and Prosodic Variabilities in Connected Speech: An Applied Data-Bank Study", *Proceedings of the XIth ICPHS*, Vol.6, Tallin, 102-105.

109. Fant, G.; Kruckenberg, A.; Nord, L. (1990) "Acoustic correlates of rhythmical structures in text reading", *Nordic Prosody*, N°5, K. Wiik & I. Raimo (eds.), University of Turku, 70-86.
110. Fant, G.; Kruckenberg, A.; Nord, L. (1991a) "Language specific patterns of prosodic and segmental structures in Swedish, French and English", *Proceedings of the XIIth ICPhS*, Vol. 4, Université de Provence, Aix-en-Provence, 118-121.
111. Fant, G.; Kruckenberg, A.; Nord, L. (1991b) "Durational correlates of stress in Swedish, French and English", *Journal of Phonetics*, Vol. 19, 351-365.
112. Faure, G.; Hirst, D.J.; Chafcouloff, M. (1980) "Rhythm in English: Isochronism, Pitch, and Perceived Stress", in *The Melody of Language*, Waugh & Schooneveld (eds.), University Park Press, Baltimore, 71-79.
113. Fletcher, J. (1991) "Rhythm and final lengthening in French", *Journal of Phonetics*, Vol. 19(2), 193-212.
114. Fohr, D.; Laprie, Y. (1989) "Snorri: an Interactive Tool for Speech Analysis", *Proceedings of the European Conference on Speech Technology*, Paris.
115. Fónagy, I. (1967) "Les structures rythmiques de la poésie", *Colloque sur les rythmes*, Lyon, Institut d'audiophonologie de Lyon, SIMEP-Editions, 307-323.
116. Fónagy, I. (1980) "L'accent en français: accent probabilitaire", in *L'accent en français contemporain (Studia Phonetica)*, Vol. 15, I. Fónagy & P. Léon (eds.), 123-233.
117. Fónagy, I.; Léon, P. (1980) *L'accent en français contemporain* Studia Phonetica, Vol. 15, I. Fónagy & P. Léon (eds.).
118. Fónagy, I. (1989) "La Cendrillon de Dwight Bolinger", *Semiotica*, Vol. 76 (3-4), 217-264.
119. Fónagy, I.; Fónagy, J. (1976) "Prosodie professionnelle et changements prosodiques", *Le Français Moderne*, n° 3, 193-227.
120. Fougeron, C. (1998) "Articulation consonantique en début de constituants prosodiques de différents niveaux", *Actes des XXII^{èmes} Journées d'Études sur la Parole*, 15-19 juin 1998, Martigny, Suisse, 291-294.
121. Fougeron, C.; Jun, S.A. (1996) "Rate effects on French intonation: Prosodic organization and phonetic realization", *UCLA Working Papers in Phonetics*, N° 94, 26-51.
122. Fougeron, C.; Keating, P.A. (1997) "Articulatory strengthening at edges of prosodic

- domains", *Journal of the Acoustic Society of America*, Vol. 101, 3728-3740.
123. Fraisse, P. (1956) "Les structures rythmiques", *Studia Psychologica*, Publications Universitaires de Louvain.
 124. Fraisse, P. (1967) "Psychologie des rythmes humains", *Colloque sur les rythmes*, Lyon, Institut d'audiophonologie de Lyon, SIMEP-Editions, 23-33.
 125. Fraisse, P. (1974) *Psychologie du Rythme*, Presse Universitaire de France.
 126. Fry, D.B. (1958) "Experiments in the perception of stress", *Language and Speech*, Vol. 1, 126-152.
 127. Fuchs, C. (1997) "Diversité des représentations linguistiques: quels enjeux pour la cognition?", in *Diversité des langues et représentations cognitives*, C. Fuchs & S. Robert (eds.), Ophrys, 5-24.
 128. Fuchs, C. ; Robert, S. (1997) *Diversité des langues et représentations cognitives*, C. Fuchs & S. Robert (eds.), Ophrys.
 129. Gabrielsson, A. (1982) "Perception and Performance of Musical Rhythm", in *Music, Mind and Brain. The neuropsychology of Music*. 'An expanded version of proceedings from the Third Workshop on the Physical and Neuropsychological Foundations of Music, held in August 8-12, 1980, in Ossiach, Austria', Manfred Clynes (ed.), Plenum Press. New York., 159-169.
 130. Garde, P. (1968) *L'accent*, Presse Universitaire de France., Collection SUP.
 131. Ghio, A. (1997) *ACHILE: un dispositif de décodage ACooustico-pHonétique et d'Identification LExicale indépendant du locuteur à partir de modules mixtes*, Thèse de Doctorat, Université d'Aix-Marseille III, 231 pp.
 132. Gibbon, D. ; Richter, H. (1984) *Intonation, Accent and Rhythm: Studies in Discourse Phonology*, D. Gibbon and H. Richter (eds.), Walter de Gruyter, Berlin,
 133. Goldman-Eisler, F. (1968) *Psycholinguistics: Experiments in spontaneous speech*, Academic Press, New York.
 134. Goldsmith, J. (1994) *The Handbook of Phonological Theory*, Blackwell, Oxford.
 135. Grosjean, F. (1980) "Temporal Variables within and between Languages", in *Temporal Variables in Speech: Studies in Honour of Frieda Goldman-Eisler*, Hans W. Dechert & M. Raupach (eds.), Mouton, The Hague, 39-53.
 136. Grosjean, F.; Deschamps, A. (1972) "Analyse des variables temporelles du français spontané", *Phonetica*, 26, 3, 129-156.

137. Grosjean, F.; Deschamps, A. (1973) "Analyse des variables temporelles du français spontané: comparaison du français oral dans la description avec l'anglais et avec le français", *Phonetica*, 28, 3/4, 191-226.
138. Grosjean, F.; Deschamps, A. (1975) "Analyse contrastive des variables temporelles de l'anglais et du français: vitesse de la parole et variables composantes, phénomènes d'hésitation", *Phonetica*, 31, 144-184.
139. Guaitella, I. (1986) *Considérations sur le rythme d'une structure prosodique à un seul niveau*, DEA, Aix-en-Provence.
140. Guaitella, I. (1988) "Variations de durée en syllabe accentuée", *Travaux de l'Institut de Phonétique d'Aix*, 12, 189-203.
141. Guaitella, I. (1990) "Propositions pour une méthode d'analyse de l'intonation en parole spontanée", *1er Congrès Français d'Acoustique*, Vol. 2 (sup.), 515-518.
142. Guaitella, I. (1991) *Rythme et Parole : Comparaison critique du rythme de la lecture oralisée et de la parole spontanée*, Thèse de Doctorat, Université de Provence I, 423 pp.
143. Guaitella, I. (1995) "Mélodie du geste, mimique vocale?", *Semiotica*, Vol. 103 (3/4), Mouton de Gruyter, Berlin. New York, 253-276.
144. Guaitella, I. (1996) "Parole spontanée et lecture oralisée: Activités cognitives différentes, organisations rythmiques différentes", *Travaux de l'Institut de Phonétique d'Aix*, Vol. 17, 9-30.
145. Guaitella, I.; Santi, S. (1992) "The punctuation and perception of read and spontaneous prosody: An application to speech synthesis", in *Talking Machines: Theories, Models, and Designs*, G. Bailly, C. Benoît, & T.R. Sawallis (eds.), Elsevier Science Publishers, 351-365.
146. Gussenhoven, C. (1987) "Focus, Mode and the Nucleus", in *On Accent*, Indiana University Linguistics Club, Bloomington, 1-50.
147. Gussenhoven, C.; Bolinger, D.; Keijsper, C. (1987) *On Accent*, Indiana University Linguistics Club, Bloomington.
148. Halle, M.; Vergnaud, J.R. (1987) *An Essay On Stress*, MIT Press, Cambridge, Mass..
149. Halle, M.; Idsardi, W. (1994) "General Properties of Stress and Metrical Structure", in *The Handbook of Phonological Theory*, John Goldsmith (ed.), Blackwell, Oxford, 403-443.

150. Handel, S. (1989) *Listening. An Introduction to the Perception of Auditory Events*, The MIT Press, Cambridge, Mass.
151. Hardcastle, W.J; Marchal, A. (1990) *Speech Production and Speech Modelling*, Kluwer Academic Publishers, Dordrecht.
152. Hayes, B. (1989) "The prosodic hierarchy in meter", in *Phonetics and Phonology : Rhythm and Meter*, Vol. 1, Academic Press, INC., Paul Kiparsky & Gilbert Youmans (eds.), San Diego, U.S.A., 201-260.
153. Hennebert, P. (1967) "Introduction biologique aux problèmes des rythmes", *Colloque sur les rythmes*, Lyon, Institut d'audiophonologie de Lyon, SIMEP-Editions, 47-58.
154. Hermes, D.J.; Van Gestel, J.C. (1991) "the frequency scale of speech intonation", *Journal of the Acoustic Society of America*, Vol. 90(1), 97-102.
155. Heuft, B.; Portele, T. (1996) "Proéminence de syllabes et de frontières en synthèse vocale de l'allemand", *Actes des XXI^{èmes} Journées d'Études sur la Parole*, Avignon, 195-198.
156. Hirschberg, J. (1993) "Studies in Intonation and Discourse", *Working Papers*, Vol. 41, *ESCA Workshop on Prosody*, Lund, 90-95.
157. Hirschberg, J. (1995) "Prosodic and Other Acoustic Cues to Speaking Style in Spontaneous and Read Speech", *Proceedings of the XIIIth ICPhS*, Vol. 2, KTH and Stockholm University, Stockholm, 36-43.
158. Hirschberg, J. (sous presse) "A Corpus-Based Approach to the Study of Speaking Style", *Prosody: Theory and Experiment*, M. Horne (ed.), Dordrecht, Kluwer Academic Publishers.
159. Hirst, D.J. (1983) "Structures and Categories in Prosodic Representations", in *Prosody: Models and Measurements*, A. Cutler & D.R. Ladd (eds.), Heidelberg, 93-109.
160. Hirst, D.J. (1988) "Intonation et représentations mentales", *Revue de Phonétique Appliquée*, Vol. 87-88-89, 209-221.
161. Hirst, D.J. (1992) "Prediction of prosody: An overview", in *Talking Machines: Theories, Models and Designs*, G. Bailly ; C. Benoit & T.R. Sawallis (eds.), Elsevier Science Publishers, North-Holland, Amsterdam
162. Hirst, D.J. (1993) "Peak, boundary and cohesion characteristics of prosodic grouping", *Working Papers*, Vol. 41, *ESCA Workshop on Prosody*, Lund.

163. Hirst, D.J.; Di Cristo, A. (1984) "French Intonation: a Parametric Approach", *Die Neueren Sprachen*, Vol. 83(5), 554-569.
164. Hirst, D.J.; Espesser, R. (1993) "Automatic modelling of fundamental frequency with a quadratic spline function", *Travaux de l'Institut de Phonétique d'Aix*, Vol. 15, 71-85.
165. Hirst, D.J.; Di Cristo, A. (1996) "Y a-t-il des unités tonales en français?", *Actes des XXI^{èmes} Journées d'Etude sur la Parole*, Avignon, 10-14 juin 1996, 223-226.
166. Hirst, D.J.; Astésano, C.; Di Cristo, A. (1998) "Differential lengthening of syllabic constituents in French: the effect of accent type and speaking style", *ICSLP'98*, Australia.
167. Hirst, D.J.; Di Cristo, A.; Espesser, R. (sous presse) "Levels of description and levels of representation in the analysis of intonation", in *Prosody: Theory and Experiment*, M. Horne (ed.), Kluwer Academic Press.
168. Hofhuis, E. (1993) "Establishing Prosodic Structure by Measuring Segment Duration", *Working papers*, Vol. 41, *ESCA Workshop on Prosody*, Lund, 136-139.
169. Hofhuis, E.; Gussenhoven, C.; Rietveld, A.C.M. (1995) "Final lengthening at prosodic Boundaries in Dutch", *Proceedings of the XIIIth ICPhS*, Vol. 1, Stockholm, 154-157.
170. Horne, M. (sous presse) *Prosody: Theory and Experiment*, M. Horne (ed.), Dordrecht, Kluwer Academic Publishers.
171. House, D.; Hermes, D.; Beaugendre, F. (1997) "Temporal-alignment categories of accent-lending rises and falls ", *Proceedings of the 5th European Conference On Speech Communication and Technology*, Vol.2, ESCA, Rhodes, Greece, 22-25 Septembre 1997, 879-882.
172. Hyman, L.M. (1975) *Phonology: Theory and Analysis*, Holt, Rinehart & Winston.
173. Hyman, L.M. (1977) *Studies in stress and accent*, Vol. 4, L. Hyman (ed.), Southern California Occasional Papers in Linguistics,
174. Jankowski, L. (en cours) *Langage, communication et structuration temporelle*, Thèse de 3ième cycle, en préparation, Université de Provence I.
175. Jankowski, L.; Astésano, C.; Di Cristo, A. (1999) "The initial rhythmic accent in French: acoustic and perceptual prosodic cues", *ICPhS'99*, 1-7 août, San Francisco, USA, 257-260.
176. Jassem, W.; Hill, D.R.; Witten, I.H. (1984) "Isochrony in English Speech: its

Statistical Validity and Linguistic Relevance", in *Intonation, Accent and Rhythm: Studies in Discourse Phonology*, D. Gibbon and H. Richter (eds.), Walter de Gruyter, Berlin, 203-225.

177. Jun, S.A.; Fougeron, C. (1995) "The accentual phrase and the prosodic structure of French", *Proceedings of the XIIIth ICPPhS*, Vol. 2, KTH and Stockholm University, Stockholm, 722-725.
178. Jun, S.A.; Fougeron, C. (1997) "A phonological model of French intonation: introducing an intermediate phrase", *ESCA Workshop on Intonation*, Athènes, Grèce, 18-20 septembre 1997, (non inclus dans les actes).
179. Kager, R. (1994) "The Metrical Theory of Word Stress", in *The Handbook of Phonological Theory*, John Goldsmith (ed.), Blackwell, Oxford, 367-402.
180. Kiparsky, P. ; Gilbert, Y. (1989) *Phonetics and Phonology : Rhythm and Meter*, Vol. 1, Academic Press, INC., Paul Kiparsky & Gilbert Youmans (eds.), San Diego, U.S.A.
181. Klatt, D.H. (1976) "Linguistic uses of segmental duration in English: acoustic and perceptual evidence", *Journal of the Acoustic Society of America.*, Vol. 59, 5, 1208-1221.
182. Kohler, K.J. (1986) "Invariance and Variability in Speech Timing: From Utterance to Segment in German", in *Invariance and Variability in Speech Processes*, J.S. Perkell and D.H. Klatt (eds.), MIT., Lawrence Erlbaum Associates Publishers, New Jersey, 268-289.
183. Kohler, K.J. (1987) "Categorical Pitch Perception", *Proceedings of the XIth ICPPhS*, Vol. 5, Tallin, 331-333.
184. Kohler, K.J. (1991) "Isochrony, units of rhythmic organization and speech rate", *Proceedings of the XIIth ICPPhS*, Vol.1, Université de Provence, Aix-en-Provence, 257-261.
185. Konopczynski, G. (1986) *Du prélangage au langage: acquisition de la structuration prosodique*, Thèse de Doctorat d'État, Université de Strasbourg.
186. Konopczynski, G. (1991) "Acquisition de la proéminence dans le langage émergent", *Proceedings of the XIIth ICPPhS*, Vol.1, Université de Provence, Aix-en-Provence, 333-337.
187. Kuhl, P.K (1992) "Speech Prototypes: Studies on the Nature, Function, Ontogeny and Phylogeny of the "Centers" of Speech Categories", in *Speech Perception, Production and Linguistic Structure*, Y. Tohkura; Vatikiotis-Bateson, E.; Sagisaka, Y. (eds.), Tokyo, 239-264.

188. Ladd, D.R. (1992) "An introduction to intonational phonology", *Papers in Laboratory Phonology II. Gesture, Segment, Prosody*, G.J. Docherty & D.R. Ladd (eds.), Cambridge University Press, 321-334.
189. Ladd, D.R. (1996) *Intonational Phonology*, Cambridge Studies in Linguistics, Cambridge University Press.
190. Ladd, D.R. (sous presse) "Tones and turning points: Bruce, Pierrehumbert, and the elements of intonational phonology", in *Prosody: Theory and Experiment*, M. Horne (ed.), Dordrecht, Kluwer Academic Publishers.
191. Ladd, D.R.; Campbell, W.N. (1991) "Theories of prosodic structure: evidence from syllable duration", *Proceedings of the XIIth ICPHS*, Vol. 2, Université de Provence, Aix-en-Provence, 290-293.
192. Ladd, D.R.; Morton, R. (1997) "The perception of intonational emphasis: continuous or categorical?", *Journal of Phonetics*, Vol. 25, 313-342.
193. Landercy, A. ; Renard, R. (1977) *Éléments de Phonétique*, Didier, Bruxelles.
194. Lecuit, V.; Demolin, D. (1998) "Relation entre pression sous glottique et intensité: étude des voyelles du français", *Actes des XXII^{èmes} Journées d'Etude sur la Parole*, Martigny, 15-19 Juin 1998, 299-302.
195. Lehiste, I. (1972) "The Timing of Utterances and Linguistic Boundaries", *Journal of the Acoustic Society of America*, Vol. 51 (6-2), 2018-2024.
196. Lehiste, I. (1973) "Rhythmic units and syntactic units in production and perception", *Journal of the Acoustical Society of America*, Vol. 54 (5), 1228-1234.
197. Lehiste, I. (1983) "Semantics, Syntax and Prosody", *Proceedings of the Xth ICPHS*, Vol. 1, A. Cohen & M .P. v.d. Broecke (eds.), 1-6 August 1983, Utrecht, The Netherlands, 69-74.
198. Léon, P. (1992) *Phonétisme et prononciations du français*, Nathan Université, Paris.
199. Léon, P. (1993) *Précis de Phonostylistique: Parole et Expressivité*, Nathan Université.
200. Levac, L.; Cedergren, H.J.; Perreault, H. (1993) "Phonetic Evidence of Narrow and Wide Temporal Scope for Prosodic Constituents in French", *Working Papers*, Vol. 41, *ESCA Workshop on Prosody*, Lund, 54-57.
201. Levin, H.; Schaffer, C.A.; Snow, C. (1982) "The prosodic and paralinguistic features of reading and telling stories", *Language and Speech*, Vol. 25, n° 1, 43-54.

202. Levitt, A. (1992) "Reiterant speech as a test of non native speakers' mastery of the timing of French", *Haskins Laboratories Status Report on Speech Research*, Vol. 109-110, 59-72.
203. Lévy, E. (1926) "Métrique et Rythmique", *Compte-Rendu du Premier Congrès du Rythme*, Albert Pfrimmer (ed.), Genève, 76-81.
204. Liberman, M. (1975) *The Intonational System of English*, PhD. Dissertation, Cambridge, Mass. MIT; published (1979), New York: Garland.
205. Liberman, M.; Prince, A. (1977) "On stress and linguistic rhythm", *Linguistic Inquiry*, Vol. 2, 249-336.
206. Lindblom, B. (1978) "Final lengthening in speech and music", *Nordic Prosody*, E. Garding (ed.), Lund, 85-101.
207. Lindblom, B. (1990) "Explaining Phonetic Variation: A Sketch of the H&H Theory", in *Speech Production and Speech Modelling*, Kluwer Academic Publishers, W.J. Hardcastle and A. Marchal (eds.), Dordrecht, 403-439.
208. Llorca, R. (1987) *Éléments d'analyse du rythme de la parole en français*, Thèse nouveau régime, Université de Franche-Comté, Besançon, 378 pp.
209. Lucci, V. (1973) "Étude phonostylistique du rythme et de la variabilité de la longueur en français parlé et français lu", *Bulletin de l'Institut de Phonétique de Grenoble*, Vol. II, 139-161.
210. Lucci, V. (1983) *Étude phonétique du français contemporain à travers la variation situationnelle*, Publications de l'Université des Langues et Lettres de Grenoble, 360pp.
211. Malmberg, B. (1967) "Le rythme comme phénomène linguistique et phonétique", *Colloque sur les rythmes*, Lyon, Institut d'audiophonologie de Lyon, SIMEP-Éditions, 35-45.
212. Martin, Ph. (1980) "Une théorie syntaxique de l'accentuation en français", in *L'accent en français contemporain (Studia Phonetica)*, Vol. 15, I. Fónagy & P. Léon (eds.), 1-12.
213. Martin, Ph. (1981a) "Pour une théorie de l'intonation", in *L'intonation: de l'acoustique à la sémantique*, Rossi, M.; Di Cristo, A.; Hirst, D.; Martin, Ph.; Nishinuma, Y. (eds.), Klincksieck, Paris, 234-271.
214. Martin, Ph. (1981b) "Extraction de la fréquence fondamentale par intercorrélation avec une fonction peigne", *Actes des XII^{èmes} Journées d'Études sur la Parole*, Montréal, 221-232.

215. Martin, Ph. (1987) "Structure rythmique de la phrase française. Statut théorique et données expérimentales", *XVI^{èmes} Journées d'Études sur la Parole*, 255-258.
216. Massaro, D.W. (1975) *Understanding Language: an Information-Processing Analysis of Speech Perception, Reading and Psycholinguistics*, D.W. Massaro (ed.).
217. Mc Fadden, H. (1974) "Conditional Logit analysis of qualitative choice behaviour", in *Frontiers in Econometrics*, P. Zarembka (ed.), New York Academic Press.
218. McAdams, S. (1997) "L'organisation perceptive de l'environnement sonore", *Rencontres IPSEN en ORL, 1997*, (tiré du site *Web de l'Ircam*).
219. Mehler, J.; Christophe, A. (1992) "Speech processing and segmentation in Romance Languages", in *Speech Perception, Production and Linguistic Structure*, Y. Tohkura; Vatikiotis-Bateson, E.; Sagisaka, Y. (eds.), Tokyo, 221-238.
220. Mertens, P. (1992) "L'accentuation de syllabes contiguës", in *Études de Linguistique Générale et Romane*, Vol. 1, N. Delbecque & P. Swiggers (eds.), 145-165.
221. Mertens, P. (1993) "Intonational grouping, boundaries and syntactic structure in French", *Working Papers*, Vol. 41, *ESCA Workshop on Prosody*, Lund, 156-159.
222. Meunier, C. (1994) *Les groupes de consonnes. Problématique de la segmentation et Variabilité acoustique*, Thèse de Doctorat, Université de Provence, 475 pp.
223. Miller, G. (1956) "The magical number seven, plus or minus two: some limits on our capacity for processing information", *Psychological Review*, Vol. 63, 81-96.
224. Monnin, P.; Grosjean, F., (1993) "Les structures de performance en français: caractérisation et prédiction", *L'Année Psychologique*, Vol. 93, 9-30.
225. Nagelkerke, N.J.D. (1991) "A note on a general definition of the coefficient of determination", *Biometrika*, Vol. 78(3), 691-692.
226. Nespor, M.; Vogel, I. (1983) "Prosodic Structure Above the Word", in *Prosody: Models and Measurements*, A. Cutler & D.R. Ladd (eds.), Heidelberg, 123-156.
227. Nicolas, P. (1995) *Organisation intonative du texte lu en français*, Thèse de Doctorat, Université de Provence I, Aix-en-Provence, 338 pp.
228. Nishinuma, Y.; Duez, D. (1987) "Étude perceptive de l'organisation temporelle de l'énoncé en français", *Travaux de l'Institut de Phonétique d'Aix*, 11, 184-201.
229. Nooteboom, S.G. (1991) "Some observations on the temporal organisation and rhythm of speech", *Proceedings of the XIIth ICPHS*, Vol. 1, Université de Provence, Aix-en-Provence, 228-237.

230. Nord, L.; Kruckenberg, A.; Fant, G. (1990) "Some aspects of rhythm in prose, poetry and music", *Nordic Prosody*, Vol. 5, K. Wiick & I. Raimo (eds.), University of Turku, 256-265.
231. O'Shaughnessy, D. (1981) "A study of French vowel and consonant durations", *Journal of Phonetics*, Vol. 9, 385-406.
232. O'Shaughnessy, D. (1984) "A multispeaker analysis of durations in read French paragraphs", *Journal of the Acoustic Society of America*, Vol. 76 (6), 1664-1672.
233. Ouellon, C. (1991) "L'énoncé en discours oral spontané", *Revue Québécoise de Linguistique Théorique et Appliquée*, Vol. 10 (3), H. Wittmann (ed.), 61-84.
234. Park, Y.M. (1986) *Étude sur le rythme en français*, DEA, Aix en Provence.
235. Padeloup, V. (1987) "Analyse acoustique de la structuration rythmique du français oral", *XVI^{èmes} Journées d'Études sur la Parole*, 259-262.
236. Padeloup, V. (1990) *Modèle de règles rythmiques du français appliqué à la synthèse de la parole*, Thèse de Doctorat de 3^{ième} cycle, Université de Provence I, 386 pp.
237. Perkell, J.S. ; Klatt, D.H. (1986) *Invariance and Variability in Speech Processes*, J.S. Perkell and D.H. Klatt (eds.), MIT., Lawrence Erlbaum Associates Publishers, New Jersey.
238. Perrot, J. (1997) *Polyphonie pour Ivan Fónagy*, Jean Perrot (ed.), L'Harmattan, Paris.
239. Piattelli-Palmarini, M. (1979) *Théories du langage, Théories de l'apprentissage. Le débat entre Jean Piaget et Noam Chomsky*, M. Piattelli-Palmarini (ed.), Centre Royaumont pour une science de l'homme. Éditions du Seuil, 533 pp.
240. Pierrehumbert, J. (1980) *The Phonology and Phonetics of English Intonation*, PhD. Thesis, Cambridge: MIT, publiée (1988): Indiana University Linguistic Club.
241. Pierrehumbert, J. (sous presse) "Tonal elements and their alignment", in *Prosody: Theory and Experiment*, M. Horne (ed.), Dordrecht, Kluwer Academic Publishers.
242. Pike, K.L. (1945) *The intonation of American English*, the University of Michigan Press, Ann Arbor.
243. Pompino-Marschall, B. (1989) "On the psychoacoustic nature of the P-center phenomenon", *Journal of Phonetics*, Vol. 17, 175-192.
244. Pompino-Marschall, B. (1990) "Phonetic Events and Boundaries: the Syllable's

- Center of Gravity and Its Consonantal Margins", *Proceedings of LP'90*, B. Palek and P. Janota (eds.), Pragues, 350-363.
245. Portele, T.; Heuft, B. (1997) "Towards a prominence-based synthesis system", *Speech Communication*, Vol. 21, 61-72.
246. Repp, B. (1992) "Probing the cognitive representation of musical time: structural constraints on the perception of timing perturbations.", *Haskins Laboratories Status Report on Speech Research*, 111/112, New Haven, Connecticut, 293-320.
247. Rietveld, T.; Gussenhoven, C., (1985) "On the relation between pitch excursion size and prominence", *Journal of Phonetics*, Vol. 13, 299-308.
248. Rietveld, T.; Gussenhoven, C., (1997) "The influence of phrase boundaries on perceived prominence in two-peak intonation contours ", *Proceedings of the 5th European Conference On Speech Communication and Technology*, Vol.2, ESCA, Rhodes, Greece, 22-25 Septembre 1997, 859-862.
249. Robert, S. (1997) "Variation des représentations linguistiques: des unités à l'énoncé", in *Diversité des langues et représentations cognitives*, C. Fuchs & S. Robert (eds.), Ophrys, 25-39.
250. Rossi, M. (1967a) "Sur la hiérarchie des paramètres de l'accent", *VIth ICPHS*, Pragues, 7-13 Sept., 779-786.
251. Rossi, M. (1967b) "L'accent, le mot et ses limites", *Actes du X^o Congrès International des Linguistes*, Bucarest, 28 Août-2 Sept., 175-180.
252. Rossi, M. (1971) "Le seuil de glissando ou seuil de perception des variations tonales pour la parole", *Phonetica*, Vol. 24, 129-161.
253. Rossi, M. (1972) "Le seuil différentiel de durée", in *Papers in Linguistics and Phonetics to the Memory of Pierre Delattre*, Vol. 54, A. Valdman (ed.), Collection Janua Linguarum, Mouton, The Hague, Indiana University.
254. Rossi, M. (1980) "Le français, langue sans accent ?", in *L'accent en français contemporain (Studia Phonetica)*, Vol. 15, I. Fónagy & P. Léon (eds.), 13-51.
255. Rossi, M. (1981a) "Intonation, énonciation, syntaxe", in *L'intonation: de l'acoustique à la sémantique*, Rossi, M.; Di Cristo, A.; Hirst, D.; Martin, Ph.; Nishinuma, Y. (eds.), Klincksieck, Paris, 184-233.
256. Rossi, M. (1981b) "l'intonation n'est pas congruente à la syntaxe: une explication", in *L'intonation: de l'acoustique à la sémantique*, Rossi, M.; Di Cristo, A.; Hirst, D.; Martin, Ph.; Nishinuma, Y. (eds.), Klincksieck, Paris, 290-296.

257. Rossi, M. (1985a) "De la quiddité des variables ", *Actes du séminaire 'Variabilité et spécificité du locuteur'*, Société Française d'Acoustique, Marseille-Luminy, 11-27.
258. Rossi, M. (1985b) "L'intonation et l'organisation de l'énoncé", *Phonetica*, Vol. 42, 135-153.
259. Rossi, M. (1987) "Peut-on prédire l'organisation prosodique du langage spontané?", *Études de Linguistique Appliquée. Aspects prosodiques de la communication*, M. Callamand (ed.), 20-48.
260. Rossi, M.; Di Cristo, A.; Hirst, D.; Martin, Ph.; Nishinuma, Y. (1981) *L'intonation: de l'acoustique à la sémantique*, Klincksieck, Paris.
261. Sagisaka, Y.; Campbell, N.W. ; Higushi, N. (1997) *Computing Prosody: Computational Models for Processing Spontaneous Speech*, Sagisaka, Y.; Campbell, N.W. ; Higushi, N. (eds.), (A collection of papers from the Spring '95 Workshop on Computational Approaches to Processing the Prosody of Spontaneous Speech. Kyoto, Japan). Springer Verlag, New York.
262. Santi, S. (1992) *Synthèse vocale des sons du français: modélisation acoustique et évaluation perceptive*, Thèse de Doctorat, Université de Provence.
263. Santi, S. ; Guaïtella, I. ; Cavé, Ch. ; Konopczynski, G. (1998) *Oralité et Gestualité. Communication multimodale, Interaction*, Santi, S. ; Guaïtella, I. ; Cavé, Ch. ; Konopczynski, G. (eds.), *Colloque Orage*, 9-11 décembre 1998, Besançon, L'Harmattan, Paris.
264. Scott, D.R.; Isard, S.D.; de Boysson-Bardies, B. (1985) "Perceptual isochrony in English and French", *Journal of Phonetics*, Vol. 13, 155-162.
265. Séguinot, A. (1976) "L'accent d'insistance en français standard", *Studia Phonetica*, Vol. 12, 1-58.
266. Seiler, H. (1997) "Localisation et prédication: grec ancien et langues diverses", in *Diversité des langues et représentations cognitives*, C. Fuchs & S. Robert (eds.), 106-118.
267. Selkirk, E. (1980) "On Prosodic Structure and its Relation to Syntactic Structure", *Nordic Prosody*, Vol. 2, T. Fretheim (ed.), 111-140.
268. Selkirk, E. (1984) "On the Major Class Features and Syllable Theory", in *Language Sound Structure*, The MIT Press, M. Aronoff and R.T. Oehrle (eds.), Cambridge, Mass., 107-136.
269. Shattuck-Hufnagel, S. (sous presse) "Phrase-level phonology in speech production planning: Evidence for the role of prosodic structure", in *Prosody: Theory and*

Experiment, M. Horne (ed.), Dordrecht, Kluwer Academic Publishers.

270. Swerts, M. (1993) "On The Prosodic Prediction of Discourse Finality", *Working Papers*, Vol. 41, *ESCA Workshop on Prosody*, Lund, 96-99.
271. Swerts, M. (1995) "Combining statistical and phonetic analyses of spontaneous discourse segmentation", *Proceedings of the XIIIth ICPHS*, Vol. 4, Stockholm, 208-211.
272. Swerts, M.; Geluykens, R. (1994) "Prosody as a marker of information flow in spoken discourse", *Language and Speech*, Vol. 37(1), 21-43.
273. Terken, J. (1991a) "Production and perception of prosodic prominence", *Proceedings of the XIIth ICPHS*, Vol.1, Université de Provence, Aix-en-Provence, 288-293.
274. Terken, J. (1991b) "Fundamental frequency and perceived prominence of accented syllables", *Journal of the Acoustic Society of America*, Vol. 89(4), 1768-1776.
275. Terken, J. (1993) "Issues in the Perception of Prosody", *Working Papers*, Vol. 41, *ESCA Workshop on Prosody*, Lund, 228-233.
276. Terken, J. (1997) "Variation of Accent Prominence within the Phrase: Models and Spontaneous Speech data", in *Computing Prosody: Computational Models for Processing Spontaneous Speech* (A collection of papers from the Spring '95 Workshop on Computational Approaches to Processing the Prosody of Spontaneous Speech. Kyoto, Japan), Y. Sagisaka; N.W. Campbell; N. Higushi (eds.), Springer Verlag, New York, 95-116.
277. Terken, J.; Hermes, D. (sous presse) "The perception of prosodic prominence", in *Prosody: Theory and Experiment*, M. Horne (ed.), Dordrecht, Kluwer Academic Publishers.
278. t'Hart, J.; Collier, R.; Cohen, A. (1990) *A perceptual study of intonation: an experimental-phonetic approach*, Cambridge : Cambridge University Press.
279. Tohkura, Y. ; Vatikiotis-Bateson, E.; Sagisaka, Y. (1992) *Speech Perception, Production and Linguistic Structure*, Y. Tohkura; Vatikiotis-Bateson, E.; Sagisaka, Y. (eds.), Tokyo.
280. Touati, P. (1987) "Structures prosodiques du suédois et du français : profils temporels et configurations tonales.", *Travaux de l'Institut de Linguistique de Lund*, Vol. 21, Lund University Press.
281. Trager, G.; Smith, H.L. (1951) *An Outline of English Structure*, Bettenburg Press.

282. Turk, A.E.; Sawusch, J.R. (1997) "The Domain of Accentual Lengthening in American English", *Journal of Phonetics*, Vol. 25, 25-41.
283. Vaissière, J. (1974a) "On French Prosody", *Quarterly Program Report*, n° 114, MIT, 212-223.
284. Vaissière, J. (1974b) "Fréquence fondamentale des phrases déclaratives en français", *Actes des V^{èmes} Journées d'Études sur la Parole*, Vol. 1, Mai 1974, Orsay, 59-68.
285. Vaissière, J. (1975) "Caractérisation des variations de la fréquence du fondamental dans les phrases françaises", *Actes des VI^{èmes} Journées d'Étude sur la Parole*, Mai 1975, Toulouse, 40-50.
286. Vaissière, J. (1977) "Premiers essais d'utilisation de la durée pour la segmentation en mots dans un système de reconnaissance", *VIII^{èmes} Journées d'Étude sur la Parole*, 25-27 Mai 1977, Aix-en-Provence, 345-352.
287. Vaissière, J. (1983) "Language-Independent Prosodic Features", in *Prosody: Models and Measurements*, A. Cutler & D.R. Ladd (eds.), Heidelberg, 53-66.
288. Vaissière, J. (1990) "Rhythm, accentuation and final lengthening in French", *Music, Brain and Language*, Stockholm.
289. Vaissière, J. (1991) "Perceiving rhythm in French?", *Proceedings of the XIIth ICPhS*, Vol. 4, Université de Provence, Aix-en-Provence, 258-261.
290. Vaissière, J. (1997) "Ivan Fónagy et la notation prosodique", in *Polyphonie pour Ivan Fónagy*, Jean Perrot (ed.), L'Harmattan, Paris, 479-487.
291. van Heuven, V.J. (1993) "On the temporal domain of focal accent", *Working Papers*, Vol. 41, *ESCA Workshop on Prosody*, Lund, 132-135.
292. van Santen, J.P.H. (1997) "Segmental Duration and Speech Timing", in *Computing Prosody: Computational Models for Processing Spontaneous Speech* (A collection of papers from the Spring '95 Workshop on Computational Approaches to Processing the Prosody of Spontaneous Speech. Kyoto, Japan), Y. Sagisaka; N.W. Campbell; N. Higushi (eds.), Springer Verlag, New York, 225-249.
293. Verluyten, P. (1983) "Phonetic Reality of Linguistic Structures: the Case of (Secondary) Stress in French", *Proceedings of the Xth ICPhS*, Vol. 2, Utrecht, 522-525.
294. Vihanta, V.V. (1991) "Signalisation prosodique de la structure informationnelle dans le discours radiophonique en finnois et en français", *Proceedings of the XIIth ICPhS*, Vol. 2, Université de Provence, Aix-en-Provence, 422-425.

295. Vihanta, V.V. (1993) "Focalisations et autres proéminences en français lu et spontané", in *Mélanges d'Études Romanes offerts à Lauri Lindgren*, Turun Yllöpisto, 258-289.
296. Vion, R. (1992) *La communication verbale. Analyse des interactions*, Paris, Hachette Supérieur.
297. Waugh, L.R.; Schooneveld, C.H. (1980) *The Melody of Language: Intonation and Prosody*, Linda R. Waugh & C.H. van Schooneveld (eds.), University Park Press, Baltimore,
298. Wells, J.; Barry, W.; Grice, M.; Fourcin, A.; Gibbon, D. (1992) "Standard Computer-Compatible Transcription. SAM stage report Sen.3 SAM-UCL-037", in *SAM ESPRIT PROJECT 2589, Multilingual Speech Input/Output Assessment, Methodology and Standardisation, Final Report, Year Three I.III.91.28.II.1992. Ref. SAM-ICL-G004*. London: University College London.
299. Wenk, B.J. (1983) "Effets de rythme dans le français parlé", *Recherches sur le français parlé*, Vol. 5, Publications de l'Université de Provence, Aix-en-Provence, 147-162.
300. Wenk, B.J. ; Wioland, F. (1982) "Is French really syllable-timed?", *Journal of Phonetics*, Vol. 10(2), 193-216.
301. Wenk, B.J.; Wioland, F. (1984) "Aspects rythmiques du discours radiophonique", in *Aspects du discours radiophonique*, Collection 'Langages, Discours et Sociétés', n°1, P. Charaudeau (ed.), Didier Érudition, 71-80.
302. Wichmann, A. (1993) "F0 Troughs and Prosodic Phrasing", *ESCA Workshop on Prosody*, Lund, 50-53.
303. Wichmann, A.; Knowles, G. (1995) "How determinable are Intonation Units?", *Proceedings of ICPhS'95*, Vol. 2, Stockholm, 14-19 August 1995, 222-225.
304. Wioland, F. (1984) "Organisation temporelle des structures rythmiques du français parlé", *Bulletin des rencontres régionales de linguistique*, Lausanne, 7-9 juin 1984, 293-322.
305. Wioland, F. (1985) *Les structures syllabiques du français. Fréquence et distribution des phonèmes consonantiques. Contraintes idiomatiques dans les séquences consonantiques*, Vol. 31, Collection Travaux de Linguistique Quantitative, Slatkine-Champion, Genève, Suisse.

- *Table des Matières* -

- Introduction -	7
Chapitre 1 - le Rythme -	19
1. Introduction	21
1.1. Eléments de réflexion	21
1.1.1. Le rythme : réalité perceptive	21
1.1.2. Différents niveaux d'actualisation du rythme	21
1.1.3. Conceptions duales du rythme	22
1.1.4. Conception unificatrice du rythme	23
1.1.5. Place du rythme dans la prosodie	24
1.2. Différentes approches de l'analyse du rythme parolier	24
2. Le rythme : quelques notions fondamentales	27
2.1. Psychologie du rythme	27
2.1.1. Le rythme biologique	27
2.1.2. Périodicité et structure	29
2.1.3. Loi d'assimilation et de dissimilation	29
2.1.4. Opposition entre 'mètre ' et 'rythme'	30
2.1.5. Théorie de la Gestalt et rythme parolier	31
2.2. Rythmicité accentuelle vs. rythmicité syllabique	32
2.2.1. Introduction : le mythe de l'isochronie et de l'isosyllabité	32
2.2.2. Caractéristiques prosodiques : niveaux de rythmicité syllabique et accentuelle	33
2.2.3. Caractéristiques phonétiques distinguant les langues	38
2.2.4. Caractéristiques physiologiques et acoustiques de l'accentuation : incidence sur le rythme	39
2.2.5. Conclusion sur la rythmicité des langues	41
2.3. La syllabe, unité rythmique	42
2.3.1. La syllabe, une unité d'encodage et de décodage	42
2.3.2. Structure syllabique	43
2.4. La performance rythmique comme expression de l'esthétique humaine	45
2.5. Perception du rythme	47
3. Résumé	49
Chapitre 2 - Les paramètres du rythme -	51
1. L'accentuation	53
1.1. Corrélats phonétiques de l'accentuation non emphatique	53
1.1.1. Paramètres d'intensité et de hauteur	53
1.1.2. La durée, paramètre central de l'accentuation du français	54
1.2. Place et fonction de l'accentuation non emphatique	57
1.2.1. Cas particulier : l'allongement final, un phénomène universel	58
1.3. Corrélats phonétiques de l'accentuation emphatique : nature, place et fonction	61
1.3.1. Nature de l'accent emphatique	61
1.3.2. Fonctions de l'accent emphatique	62
1.3.3. Distribution et domaine de l'accent emphatique	62
1.3.4. Mise en doute de l'emphase comme procédé accentuel	64
1.4. Perception des phénomènes accentuels	65

1.5.	Approches phonologiques de l'accentuation	68
1.5.1.	Définitions de l'accent	68
1.5.2.	Fonctions de la prosodie : le débat sur le lien avec la syntaxe	70
1.5.3.	Discussion sur la pertinence phonologique de l'accentuation du français	75
1.6.	L'organisation accentuelle en français	78
1.6.1.	Introduction	78
1.6.2.	Le phénomène de désaccentuation en français	79
1.6.3.	L'accentuation initiale : évolution du système accentuel en français	80
1.6.4.	Approches métriques de l'accentuation	85
1.6.5.	Limites de la représentation métrique	89
1.6.6.	Un modèle de l'accentuation du français	90
2.	Le timbre	97
3.	Le débit de parole	98
4.	Les pauses	99
4.1.	Nature des pauses	99
4.2.	Distribution et fonctions des pauses	100
4.3.	Durée des pauses	101
4.4.	Pauses et perception des unités linguistiques	102
5.	Résumé	104
Chapitre 3 - Variabilité et Discours -		105
1.	Introduction	107
2.	Invariance et variabilité	107
3.	Caractérisation des types de discours	109
3.1.	Parole lue vs. parole spontanée	110
3.1.1.	Parole lue et parole spontanée : deux processus d'encodage différents	110
3.1.2.	Comment définir la parole spontanée ?	111
3.1.3.	Intérêts de l'analyse de types de discours spontanés	112
3.1.4.	Style journalistique : hétérogénéité linguistique	113
3.2.	Caractéristiques phonétiques et prosodiques des différents styles de parole	116
3.2.1.	Principales caractéristiques phonétiques et prosodiques distinguant les styles de parole	117
3.2.2.	Critique des méthodes traditionnelles de comparaison des types de discours	119
4.	Résumé	122
Chapitre 4 - Aspects Méthodologiques -		123
1.	Objectifs de l'étude et cadre méthodologique	125
2.	Paramètres retenus pour l'analyse de l'accentuation	126
2.1.	Organisation paradigmatique de l'accentuation	127
2.1.1.	Durée : quelles unités d'observation ?	128
2.1.2.	Hauteur : unités et paramètres d'observation	128
2.2.	Organisation temporelle de l'accentuation	129
3.	Protocole expérimental	129
3.1.	Description du corpus	129

3.1.1.	Critères linguistiques de sélection du corpus	129
3.1.2.	Choix des locuteurs et des types de discours	135
3.2.	Niveaux de transcription et d'étiquetage du corpus	140
3.2.1.	Considérations techniques	140
3.2.2.	Quels types de transcriptions pour quels types d'études ?	140
3.2.3.	Niveaux de transcription adaptés à notre étude	141
3.2.4.	Segmentation phonémique	143
3.2.5.	Base de données phonémique de notre corpus	145
3.2.6.	Classification des phonèmes et affranchissement par rapport au contexte	146
3.2.7.	Syllabation et étiquetage des phonèmes en Attaque/Noyau/Coda	151
3.2.8.	Étiquetage prosodique	153
4.	Résumé	160
Chapitre 5 - Durées segmentales -		165
1.	Problématique des durées segmentales : Introduction	167
2.	État des connaissances sur la durée segmentale en français	169
2.1.	Diversité des études	169
2.2.	Caractéristiques des durées segmentales en français	170
2.2.1.	Variabilité induite par les événements articulatoires (durées intrinsèques)	170
2.2.2.	Variabilité induite par le contexte phonétique (durées co-intrinsèques)	171
2.2.3.	Variabilité de durée induite par le contexte prosodique	172
3.	Tendances d'allongement différentiel des segments	172
4.	Modélisation des durées segmentales dans un cadre syllabique : le modèle de Campbell	173
4.1.	Quelle unité phonologique pour rendre compte des phénomènes temporels ?	173
4.1.1.	La syllabe, interface rythmique	174
4.1.2.	Avantages du cadre syllabique	174
4.2.	Normalisation des durées segmentales	175
4.2.1.	Pourquoi normaliser ?	175
4.2.2.	La normalisation gomme-t-elle les variations de durée pertinentes linguistiquement ?	176
4.2.3.	Comment normaliser ?	176
4.2.4.	Normalisation par transformée Z	177
4.3.	Choix d'une méthode de normalisation adaptée à notre objet d'étude	179
5.	Analyses Empiriques	181
5.1.	Objectifs de recherche	181
5.2.	Plan expérimental	182
5.2.1.	Limites méthodologiques	183
5.2.2.	Déroulement des analyses	183
5.3.	Exposé des résultats par niveaux de constituance	184
5.3.1.	Niveau de constituance syllabique	184
5.3.2.	Niveaux infrasyllabiques : remarques méthodologiques préliminaires	193
5.3.3.	Niveau de constituance infrasyllabique en A/Rime	194
5.3.4.	Niveau de constituance syllabique en A, N et C	203
5.3.5.	Comparaison des stratégies individuelles et stylistiques par catégories accentuelles au niveau infrasyllabique des ANC	213

6. Discussion et conclusion générale sur les tendances d'allongement différentiel en fonction des contraintes prosodiques et des styles de parole	225
7. Résumé	233
Chapitre 6 - Configurations Tonales -	235
1. Introduction	237
1.1. Le débat sur l'analyse en configurations vs. l'analyse en niveaux tonals	238
2. Caractérisation des configurations tonales des proéminences en français	240
2.1. Procédure expérimentale	240
2.1.1. Méthodologie statistique	242
2.1.2. Justification de l'extraction automatique des paramètres et discussion sur la variabilité observée	244
2.2. Résultats	246
2.2.1. Analyse de Variance	246
2.2.2. Modèle Logit Multinomial	248
2.3. Discussion	249
3. Conclusion	255
4. Résumé	257
Chapitre 7 - l'organisation temporelle de la parole -	259
1. Introduction	261
2. Plan expérimental	263
2.1. Cadre méthodologique et procédure expérimentale	264
2.1.1. Définition des unités d'observation	265
3. Résultats	268
3.1. Niveaux de rythmicité	268
3.1.1. Constituance et variabilité rythmique	268
3.2. Unités de référence de l'organisation rythmique	270
3.2.1. Syllabes inaccentuées, syllabes de référence	270
3.2.2. Pauses, unités rythmiques de référence	272
3.3. Organisation temporelle des Mots Prosodiques	273
3.3.1. Taille des Mots Prosodiques et variabilité stylistique	273
3.3.2. Densité accentuelle	274
3.3.3. Taille des Mots Prosodiques en fonction des catégories accentuelles	276
3.4. Gestion stylistique des phénomènes de pause	277
3.4.1. Nature et durée des pauses	277
3.4.2. Gestion des pauses et débit de parole	279
3.4.3. Corrélation entre les durées de pauses et les durées des syllabes accentuées précédentes	280
4. Discussion et Conclusion	283
5. Résumé	287
- Conclusions et perspectives -	289
- Bibliographie -	297
Table des matières	323
Table des illustrations	329

- *Table des Illustrations* -

Figure 1 : Courbe rythmique (tiré de Wenk, 1983 : 157).	40
Figure 2 : Structure interne de la syllabe, tiré de J. Blevins, 1994, p. 216.	44
Figure 3 : Représentation en arbre métrique du mot 'Alabama'	86
Figure 4 : Représentation en grille métrique du mot 'Alabama'	86
Figure 5 : Représentation en grille métrique parenthésée de l'énoncé français : « les amis de mes voisins » (tiré de Di Cristo, sous presse).	89
Figure 6 : Représentation métrique du mot « Félicité » (tiré de Di Cristo, sous presse).	96
Figure 7 : Représentation métrique d'un énoncé et des constituants métriques de la phrase « il a rencontré les écrivains de la contestation » (Tiré de Di Cristo, sous presse).	96
Figure 8 : Typologie de l'interaction adaptée du modèle de Vion (1992) pour les besoins du projet PACOMUST. Ce tableau est issu de Astésano et al. (1995), p.37.	134
Figure 9 : Pourcentages d'occurrences des voyelles et des consonnes par locuteur, comparés à ceux de la base de données de Wioland (1985).	146
Figure 10 : Durées en millisecondes des voyelles et des consonnes par locuteur	146
Figure 11 : Durées moyennes brutes (ms) des macro-classes, tous locuteurs confondus	149
Figure 12 : Durées moyennes brutes (ms) des macro-classes par locuteur.	150
Figure 13 : Constituants immédiats de la syllabe : Attaque (A), Noyau (N), Coda (C). (σ = syllabe ; X = phonèmes)	153
Figure 14 : Représentation du mot « droit » en constituants syllabiques	153
Figure 15: Distribution et fonctions des catégories accentuelles du modèle (Di Cristo & Hirst, 1997 ; Di Cristo, sous presse)	157
Figure 16 : Tendances générales de l'allongement syllabique en fonction des catégories accentuelles, tous locuteurs confondus. L'axe des abscisses représente les catégories accentuelles et l'axe des ordonnées indique les durées normalisées (Transformées Z). (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique).	186
Figure 17 : Durées syllabiques normalisées (Z) par catégorie accentuelle, et par locuteur. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique)	187
Figure 18 : Durées syllabiques normalisées par catégorie accentuelle, pour les locuteurs du style Interview. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique)	188
Figure 19 : Durées syllabiques normalisées par catégorie accentuelle, pour les locuteurs du style Lecture. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique)	190
Figure 20 : Durées syllabiques normalisées par catégorie accentuelle, pour les locuteurs du style Journal. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI =	

- accent initial rythmique ; EMP = accent initial emphatique) _____ 191
- Figure 21 : Tendances générales observées pour la hiérarchie des constituants syllabiques sur la base des durées normalisées par catégorie accentuelle, tous locuteurs et styles de parole confondus. L'axe des abscisses représente les catégories accentuelles ; sur l'axe des ordonnées se trouvent les durées normalisées (Transformée Z) ; A = Attaque. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique) __ 195
- Figure 22 : Durée moyenne des Attaques et des Rimes (A/Rimes) par locuteur, toutes catégories accentuelles confondues. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique) __ 196
- Figure 23 : Durée moyenne des Attaques et des Rimes (A/Rimes) pour le style Interview, en fonction des catégories accentuelles. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique) __ 198
- Figure 24 : Durée moyenne des Attaques et des Rimes (A/Rimes) pour le style Lecture, en fonction des catégories accentuelles. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique) __ 199
- Figure 25 : Durée moyenne des Attaques et des Rimes (A/Rimes) pour le style Journal, en fonction des catégories accentuelles. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique) __ 201
- Figure 26 : Tendances générales observées pour la hiérarchie des constituants syllabiques sur la base des durées normalisées par catégorie accentuelle, tous locuteurs et styles de parole confondus. L'axe des abscisses représente les catégories accentuelles ; sur l'axe des ordonnées se trouvent les durées normalisées (Transformée Z) ; (A = Attaque ; N = Noyau ; C = Coda). (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique) _____ 204
- Figure 27 : Durée moyenne des Attaques, Noyaux et Codas (ANC) par locuteur, toutes catégories accentuelles confondues _____ 205
- Figure 28 : Durée moyenne des Attaques, Noyaux et Codas (ANC) par catégorie accentuelle, pour les locuteurs du style Interview. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique) _____ 208
- Figure 29 : Durée moyenne des Attaques, Noyaux et Codas (ANC) par catégorie accentuelle, pour les locuteurs du style Journal. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique) _____ 209

<i>Figure 30 : Durée moyenne des Attaques, Noyaux et Codas (ANC) par catégorie accentuelle, pour les locuteurs du style Lecture. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique)</i>	211
<i>Figure 31 : Durée moyenne des Attaques, Noyaux et Codas (ANC) pour la catégorie accentuelle AFS, et pour les 3 styles de parole</i>	214
<i>Figure 32 : Durée moyenne des Attaques, Noyaux et Codas (ANC) pour la catégorie accentuelle AI, et pour les 3 styles de parole</i>	215
<i>Figure 33 : Durée moyenne des Attaques, Noyaux et Codas (ANC) pour la catégorie accentuelle UI-NT, et pour les 3 styles de parole</i>	217
<i>Figure 34 : Durée moyenne des Attaques, Noyaux et Codas (ANC) pour la catégorie accentuelle UI-T, et pour les 3 styles de parole</i>	219
<i>Figure 35 : Durée moyenne des Attaques, Noyaux et Codas (ANC) pour la catégorie accentuelle EMP, et pour les 3 styles de parole</i>	220
<i>Figure 36 : Durée moyenne des Attaques, Noyaux et Codas (ANC) pour la catégorie accentuelle INA, et pour les 3 styles de parole</i>	222
<i>Figure 37 : Allongement différentiel global (toutes catégories accentuelles confondues) pour les 3 styles de parole. (A = Attaque ; N = Noyau ; C = Coda).</i>	229
<i>Figure 38 : Modélisation de la f0 d'une proéminence perçue. 3 points cibles définissent la configuration tonale dont sont extraits les 8 paramètres.</i>	241
<i>Figure 39 : Hypothèse : les catégories accentuelles peuvent être distinguées sur la base des paramètres tonals sélectionnés (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique).</i>	244
<i>Figure 40 : Résultats : Le modèle général explique 27% de la variabilité : les paramètres tonals semblent donc avoir un effet explicatif non négligeable dans la prédiction d'appartenance d'une proéminence à une catégorie accentuelle. (AFS = accent final de syntagme ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique).</i>	249
<i>Figure 41 : Configurations tonales des proéminences de type AFS (Final de mot Rythmique) et AI (Initial de mot Rythmique). Les configurations sont déterminées par les Points Cibles Central, Gauche et Droit. L'axe des ordonnées est exprimé en ERB * 1000. L'axe du temps (t) est purement indicatif.</i>	250
<i>Figure 42 : Configurations tonales des proéminences de type AI (Initial de mot Rythmique) et EMP (Initial de mot Emphatique). Les configurations sont déterminées par les Points Cibles Central, Gauche et Droit. L'axe des ordonnées est exprimé en ERB * 1000. L'axe du temps (t) est purement indicatif.</i>	252
<i>Figure 43 : Configurations tonales des proéminences de type AFS (Final de mot Rythmique) et UI-T (Final d'Unité Intonative Terminale). Les configurations sont déterminées par les Points Cibles Central, Gauche et Droit. L'axe des ordonnées est exprimé en ERB * 1000. L'axe du temps (t) est purement indicatif.</i>	254
<i>Figure 44 : Représentation des niveaux de constituance métrique d'une partie d'énoncé</i>	

produit par la locutrice Interview F, selon la représentation en grille métrique parenthésée de Halle & Vergnaud (1987). Les Mots Prosodiques complexes représentent, dans cet exemple, une bipolarisation accentuelle au niveau du mot lexical. _____ 266

Figure 45 : Coefficients de variations (écart-types / moyennes) des constituants prosodiques par style de parole (INA = syllabes inaccentuées ; MP Simples = Mots Prosodiques contenant un accent final ; Arcs acc. = Mots Prosodiques contenant un accent final et un accent initial rythmique (AI) ou emphatique (EMP) ; Groupes de souffles = unités de performance comprises entre deux pauses respiratoires.) __ 270

Figure 46 : Corrélation entre la durée brute (ms) des syllabes accentuées précédant les pauses, et la durée brute (ms) des pauses, par style de parole _____ 281

Figure 47 : Corrélation entre la durée brute (ms) des syllabes accentuées précédant les pauses, et la durée brute (ms) des pauses, par locuteur _____ 282



Tableau 1 : Principales caractéristiques des syllabes accentuées et inaccentuées en codachronie et en capochronie. (tiré de Wenk, 1983:156) _____ 41

Tableau 2: Les phonèmes du français (Ghio, 1997 : 62) _____ 147

Tableau 3 : Durées moyennes brutes (ms) par macro-classe, tous locuteurs confondus _ 148

Tableau 4 : Débit hors pause par locuteur (nombre de syllabes par seconde) _____ 150

Tableau 5 : Tableau d'Anova générale pour les durées normalisées des syllabes, par catégorie accentuelle, pour les 6 locuteurs. (DDL = Degré De Liberté). _____ 185

Tableau 6 : Tableau d'Anova générale pour les durées normalisées des syllabes, par catégorie accentuelle, pour les locuteurs du style Interview. _____ 188

Tableau 7 : Tableau d'Anova générale pour les durées normalisées des syllabes, par catégorie accentuelle, pour les locuteurs du style Lecture _____ 189

Tableau 8 : Tableau d'Anova générale pour les durées normalisées des syllabes, par catégorie accentuelle, pour les locuteurs du style Journal _____ 190

Tableau 9 : Résumé des résultats des tests Anova : les rapports entre les différentes composantes syllabiques (ANC ; et A/Rime) sont présentés par catégorie accentuelle et par locuteur. Une synthèse par style est proposée. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique) _____ 194

Tableau 10 : Tableau d'Anova générale pour les durées normalisées des A/Rimes, par catégorie accentuelle, pour les 6 locuteurs _____ 196

Tableau 11 : Tableau d'Anova générale pour les durées normalisées des A/Rimes, par catégorie accentuelle, pour les locuteurs du style Interview _____ 197

Tableau 12 : Tableau d'Anova générale pour les durées normalisées des A/Rimes, par catégorie accentuelle, pour les locuteurs du style Lecture _____ 199

Tableau 13 : Tableau d'Anova générale pour les durées normalisées des A/Rimes, par

<i>catégorie accentuelle, pour les locuteurs du style Journal</i>	200
<i>Tableau 14 : Tableau d'Anova générale pour les durées normalisées des A N C, par catégorie accentuelle, pour les 6 locuteurs</i>	205
<i>Tableau 15 : Tableau d'Anova générale pour les durées normalisées des A N C, par catégorie accentuelle * type de constituant, pour les locuteurs du style Interview</i>	207
<i>Tableau 16 : Tableau d'Anova générale pour les durées normalisées des A N C, par catégorie accentuelle * type de constituant, pour les locuteurs du style Journal</i>	209
<i>Tableau 17 : Tableau d'Anova générale pour les durées normalisées des A N C, par catégorie accentuelle * type de constituant, pour les locuteurs du style Lecture</i>	210
<i>Tableau 18 : Tableau d'Anova générale pour les durées normalisées des A N C pour la catégorie accentuelle AFS, pour les trois styles</i>	213
<i>Tableau 19 : Tableau d'Anova générale pour les durées normalisées des A N C pour la catégorie accentuelle AI, pour les trois styles</i>	215
<i>Tableau 20 : Tableau d'Anova générale pour les durées normalisées des A N C pour la catégorie accentuelle UI-NT, pour les trois styles</i>	216
<i>Tableau 21 : Tableau d'Anova générale pour les durées normalisées des A N C pour la catégorie accentuelle UI-T, pour les trois styles</i>	218
<i>Tableau 22 : Tableau d'Anova générale pour les durées normalisées des A N C pour la catégorie accentuelle EMP, pour les trois styles</i>	220
<i>Tableau 23 : Tableau d'Anova générale pour les durées normalisées des A N C pour la catégorie accentuelle INA, pour les trois styles</i>	221
<i>Tableau 24 : Tableau récapitulatif des stratégies d'allongement global et d'allongement différentiel par style de parole, en fonction des catégories accentuelles (S = significatif ; N.S. = non significatif): seuls les résultats statistiquement significatifs sont indiqués dans les tendances d'allongement global et différentiel. (Lec. = Lecture ; Int. = Interview ; Jour. = Journal).</i>	224
<i>Tableau 25 : Pourcentages de configurations tonales par type de configuration et par catégorie accentuelle. (AFS = accent final de syntagme ; UI-NT = accent final d'unité intonative non terminale ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique).</i>	245
<i>Tableau 26 : Paramètres non pertinents pour la description des proéminences (Interaction = interaction entre les facteurs 'catégories accentuelles' et 'locuteurs' ; P-val. = valeur de P ; F-val. = valeur de F).</i>	247
<i>Tableau 27 : Paramètres pertinents pour la description des proéminences, à inclure dans le modèle Logit. (Interaction = interaction entre les facteurs 'catégories accentuelles' et 'locuteurs' ; P-val. = valeur de P ; F-val. = valeur de F).</i>	248
<i>Tableau 28 : Résultats des pseudo-R² (AFS = accent final de syntagme ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique).</i>	248
<i>Tableau 29 : Résumé des caractéristiques tonales des catégories accentuelles (AFS = accent final de syntagme ; UI-T = accent final d'unité intonative terminale ; AI = accent initial rythmique ; EMP = accent initial emphatique).</i>	254

<i>Tableau 30 : Coefficient de variation (rapport de l'écart type sur la moyenne de durée brute, exprimé en pourcentage) et niveaux de constituance en fonction du style de parole. (INA = syllabes inaccentuées ; -/Finaux = Mots Prosodiques Simples délimités par des accents finaux de syntagme (AFS), ou d'unité intonative terminale (UI-T) ou non terminale (UI-T) ; AI/Finaux = Mots Prosodiques Complexes contenant un PM défini par un Accent Initial rythmique (AI) ou emphatique (EMP)).</i>	269
<i>Tableau 31 : Rapport de durée (ms) entre les syllabes accentuées et inaccentuées par catégorie accentuelle, en fonction des locuteurs et des styles de parole. (INA = syllabes inaccentuées ; AFS = accent final de syntagme ; UI-NT = accent final d'Unité Intonative non terminale ; UI-T = accent final d'unité intonative terminale ; EMP = accent initial emphatique ; AI = accent initial rythmique ; ACC = syllabes accentuées, toutes catégories confondues).</i>	272
<i>Tableau 32 : Durées des Pauses Respiratoires par locuteur (Seulement Pause pour les locuteurs du style Lecture) ; durées des Intervalles Interaccentuels définis par des AFS ; rapports entre la durée des IIA AFS et la durée de l'entité rythmique 'Pauses Respiratoires + syllabes accentuées pré-pausales'. (IIA AFS = Intervalle InterAccentuel formé par un Accent Final de Syntagme ; Durée syllabique UI = durée moyenne des syllabes finales d'Unité Intonative Non Terminale ; PR = Pauses Respiratoires)</i>	273
<i>Tableau 33 : Pourcentages des groupes rythmiques en fonction du nombre de syllabes les constituant, par style de parole (Mono = Monosyllabiques ; Bi = Bisyllabiques ; Tri = Trisyllabiques ; Quadri = Quadrisyllabiques ; 5 à 7 = groupes de 5 à 7 syllabes inclus ; 8 à 10 = groupes de 8 à 10 syllabes inclus ; 5 à 10 = total des groupes de 5 à 10 syllabes).</i>	274
<i>Tableau 34 : Densité accentuelle (%) et pourcentages de syllabes accentuées par catégories accentuelles, et selon leurs fonctions (rythmiques vs. sémantico-pragmatiques) et leur distribution (initiales vs. finales). (AFS = accent final de syntagme ; AI = accent initial rythmique ; UI = accent final d'Unité Intonative terminale ou non terminale ; EMP = accent initial emphatique).</i>	275
<i>Tableau 35 : Types de pauses par locuteur (en occurrence et en pourcentage (%)). Les pourcentages indiquent le rapport entre les deux macro-classes 'Hésitations Vocales' et 'Pauses Silencieuses', ainsi que le rapport entre les 'Pauses Respiratoires' et les 'Pauses Vides' à l'intérieur de la macro-classe des 'Pauses Silencieuses'.</i>	277
<i>Tableau 36 : Durée des pauses par locuteur et par style de parole (en ms)</i>	278
<i>Tableau 37 : Débit hors pauses (vitesse d'articulation), débit pauses incluses et influence des pauses sur la vitesse d'articulation (Rapport en pourcentage entre le débit hors pauses et le débit pauses incluses). Les résultats sont présentés par locuteur et par style de parole.</i>	280
<i>Tableau 38 : Caractéristiques rythmiques des styles de parole. Le sigle φ indique les paramètres rythmiques pour lesquels il existe des particularités liées à l'appartenance sexuelle.</i>	285

Quelles sont les caractéristiques accentuelles du français ? Le français est-il réellement prototypique des langues dites à rythmicité syllabique ? Le système accentuel du français est-il une variable dépendante de la situation de communication ? Autrement dit, existe-t-il autant de systèmes prosodiques qu'il y a de 'styles' de parole ? L'auteur propose de répondre à ces questions à travers l'analyse de divers types de discours. Elle adopte pour cela une approche combinant les concepts issus de la théorie métrique et ceux issus de la mouvance fonctionnaliste (aspects pragmatiques et énonciatifs). La description du système prosodique dans toute sa *variabilité*, par l'analyse de parole continue en situation de communication, lui permet de dégager des tendances accentuelles et rythmiques en rupture avec les descriptions traditionnelles de la prosodie du français. Par l'analyse des durées différentielles des composantes syllabiques, des aspects dynamiques de f0, ainsi que de l'organisation temporelle des proéminences, l'auteur montre que le français, longtemps considéré comme une langue à isochronie syllabique stricte, possède en réalité une double identité accentuelle et rythmique. D'une part, il apparaît que les accents en français tendent à se produire à intervalles réguliers, indiquant que le rythme syllabique coexiste avec un rythme accentuel. D'autre part, le début des mots ou des groupes de mots est souvent marqué par un accent initial, qui contribue à la perception globale d'une récurrence régulière subjective des accents. Cette accentuation initiale, combinée à l'accentuation finale, contribue également à l'émergence d'une unité phonologique (le Mot Prosodique Complexe) favorisant la cohésion rythmique. Cette unité de référence rythmique est particulièrement prégnante en parole spontanée, indiquant que l'accentuation initiale joue un rôle non négligeable dans l'encodage de la parole. Cette étude a des implications directes pour l'amélioration des systèmes de synthèse vocale en français, ainsi que pour les recherches en psycho- et neurolinguistique sur le décodage de la parole.

Corine Astésano est actuellement chercheur-associé à l'Université d'Edinburgh. Tout en poursuivant ses recherches sur les caractéristiques stylistiques et systémiques de l'organisation accentuelle en français, elle s'intéresse plus particulièrement aujourd'hui au rôle de la prosodie dans la perception et la segmentation de la parole, en utilisant les techniques d'investigation issues de la psycholinguistique et de la neurolinguistique.