*French Language Studies* **10** (2000), 27–44, Copyright © Cambridge University Press

# *Vers une modélisation de l'accentuation du français (seconde partie)*[1]

ALBERT DI CRISTO
*Université d'Aix*

(Received 27 November 1997; revised 15 December 1998)

ABSTRACT

Nous proposons d'effectuer dans cette étude un large tour d'horizon des problèmes liés à l'interprétation de l'accent en général et de l'accentuation du français contemporain en particulier. Cet état des lieux (première partie) débouche sur l'esquisse d'une modélisation du cadre accentuel du français (seconde partie) fondée sur une classification fonctionnelle des catégories d'accentuation et la mise en œuvre généralisée du principe de bipolarisation accentuelle qui permet, selon nous, de rendre compte du caractère probabilitaire de l'accentuation du français et de la variabilité de ses motifs rythmiques attestés dans les formes prosodiques de surface.

## 4 ESQUISSE D'UNE MODÉLISATION DE L'ACCENTUATION DU FRANÇAIS

### 4.1 *Remarques préliminaires*

L'essai de modélisation que nous présentons ici procède d'une démarche qui s'efforce d'intégrer les résultats de nos recherches antérieures sur la phonologie prosodique du français, tels qu'ils ont été exposés notamment dans: Hirst et Di Cristo, 1984; Di Cristo et Hirst, 1993; Di Cristo et Hirst, 1996 et des hypothèses plus générales sur le fonctionnement de la communication dans une perspective cognitive. À cet effet, nous proposons comme point de départ de la discussion une classification des accents fondée sur leur *catégorisation fonctionnelle*. Comme le montre la figure 6, ce classement opère une première distinction entre accents emphatiques et accents non-emphatiques, ces derniers étant répartis en trois catégories majeures correspondant respectivement à l'accentuation lexicale, à l'accentuation rythmique et à l'accentuation nucléaire. D'après la distinction classique établie par Coleman (1914), la catégorie des accents emphatiques se subdivise à son tour en deux groupes: celui des accents emphatiques d'intensification et celui des accents emphatiques de contraste.

[1] *Les pages qui suivent constituent la seconde partie d'un article dont la première partie a fait l'objet d'une publication dans le précédent numéro du* Journal of French Language Studies. *Cette première partie est composée de trois sections majeures dont les contenus concernent respectivement des considérations préliminaires sur l'accent, la représentation formelle de l'accentuation et un panorama des travaux sur l'accentuation du français. La liste des références se trouve à la fin de la première partie de l'article.*



https://doi.org/10.1017/S0959269500000120 Published online by Cambridge University Press

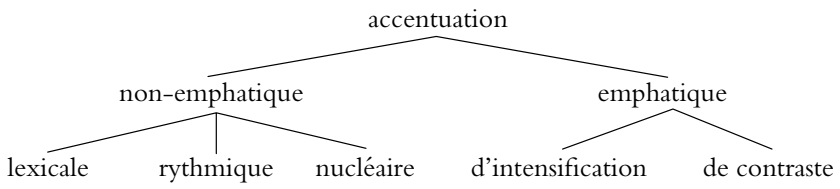



Figure 6. Typologie élémentaire des catégories accentuelles

Il est traditionnellement admis que le terme d'accent lexical désigne l'accent morphologiquement ou lexicalement distinctif qui appartient à la structure du mot et qui doit, de ce fait, figurer dans les entrées lexicales du dictionnaire de la langue. Tel n'est pas le cas du français qui ne possède donc pas, à proprement parler, d'accent lexical, ce qui n'exclut pas pour autant la possibilité d'une accentuation du mot dans cette langue. L'accent lexical participe de l'accentuation objective (non-intentionnelle), pour reprendre l'expression de Milner et Régnault (1987). Il en va de même des accents rythmiques dont le rôle n'est pas d'identifier et/ou de hiérarchiser les morphèmes au sein du mot (Garde, 1965; 1973), mais d'effectuer une partition du discours en fonction des contraintes métriques qui sont propres à la langue. Selon notre opinion, toute langue possède un rythme et toutes les proéminences participent au rythme.

Le concept d'accent nucléaire fait problème dans la mesure où il revêt dans la littérature deux significations qui ne sont pas toujours compatibles. Dans la tradition britannique dont Cruttenden (1992) a retracé récemment l'historique, le terme d'accent nucléaire ou de '*nucleus*' désigne depuis Palmer (1922) la dernière syllabe accentuée de l'unité intonative (Tone-group) qui est associée à *l'initialisation* du contour, ce dernier constituant l'élément irréductible de cette unité. Pour la tradition américaine, il est plus courant de considérer l'accent nucléaire comme le *niveau de proéminence le plus élevé* de la phrase ou de l'énoncé, ce qui lui vaut d'être le plus souvent dénommé 'sentence stress' ou 'sentence accent' (Schmerling, 1976; Ladd, 1996). On peut déduire de cette divergence que selon la première interprétation un énoncé est susceptible de posséder plusieurs accents nucléaires, alors que d'après la seconde il ne peut en comporter qu'un seul. En vérité, l'ambiguïté tient également au fait que pour beaucoup l'accentuation nucléaire est synonyme d'accentuation de focalisation. Encore faut-il s'entendre sur ce concept qui est souvent assimilé à celui d'emphase, comme en témoignent les définitions proposées dans les dictionnaires usuels et spécialisés. La notion de focus a déjà fait couler beaucoup d'encre et il n'est pas envisageable, pour préserver l'économie de notre article, que nous nous attaquions ici à cette vaste problématique. Nous tenons cependant à clarifier notre position sur ce point



https://doi.org/10.1017/S0959269500000120 Published online by Cambridge University Press

en soulignant quelques aspects qui nous paraissent essentiels et qui feront l'objet de développements plus conséquents dans une publication ultérieure (Di Cristo, à paraître).

Précisons tout d'abord que notre approche de la question se place dans la perspective d'une conception unificatrice de la communication verbale qui tente d'intégrer et de mettre en relation des aspects qui appartiennent à ce qu'il est convenu d'appeler, depuis Chomsky, la grammaire, des aspects cognitifs parfois qualifiés à tort ou à raison de paralinguistiques, des aspects dynamiques qui relèvent de la planification et du traitement perceptif des unités de la langue (aspects psycholinguistiques) et des contraintes de production et de perception dont la mise en évidence relève plus précisément de l'analyse phonétique. L'idée-force qui sous-tend et justifie cette démarche est que le fonctionnement du langage n'opère pas dans le *vide cognitif* et fait appel à des formes de traitement que l'on a parfois tendance à reléguer par commodité ou par sectarisme dans la 'corbeille à papier' des faits paralinguistiques ou non-linguistiques. Le langage en effet est un phénomène complexe dont l'usage nécessite des *tâches cognitives multiples et diversifiées* (Caplan, 1992; Pinker, 1994). Il est peu gratifiant dans cette optique de vouloir réduire le langage à un code que l'on prétendrait étudier indépendamment de la compréhension de ces tâches cognitives.

On peut s'interroger, toujours dans la même perspective, sur ce que le langage permet de communiquer. La réponse à cette question est manifestement plurielle, dans la mesure où, outre des informations factuelles ('the state of affairs', selon l'expression consacrée chez les anglophones), le langage véhicule des attitudes, des changements d'activité, des émotions et principalement des intentions. Ce point de vue élargi nous invite à prendre en considération, pour constituer notre cadre interprétatif de l'accentuation, à côté des fonctions expressive et impressive traditionnelles (Léon, 1971), des modèles du discours et de la communication qui devraient permettre d'intégrer tous ces aspects, comme, par exemple, le *modèle de discours* de Grosz et Sidner (1986), le *modèle de la contextualisation* inspiré des travaux de Gumpertz (Auer et Di Luzio, 1992) et le *modèle de la pertinence* de Sperber et Wilson (1989). Si l'on admet pareillement que le langage constitue une activité finalisée devant se plier aux exigences et aux attentes du (ou des) récipiendaire(s), on s'efforcera également de tenir compte des modèles de compréhension de la parole développés par les psycholinguistes (Cutler, 1987; Levelt, 1989; McQueen et Cutler, 1997), ainsi que du modèle adaptatif élaboré par Lindblom (1990). C'est donc à la lumière de ces travaux, dont la confrontation fait apparaître de nombreux points de convergence, que nous allons préciser notre position par rapport au problème central de la focalisation en vue d'établir sa relation avec l'accentuation.

Nous distinguons sur une base fonctionnelle, à la suite des propositions de Cullicover et Rochemont (1983), deux types élémentaires de focalisation



https://doi.org/10.1017/S0959269500000120 Published online by Cambridge University Press

qualifiées, respectivement d'information et d'emphase. Nous nous séparons cependant de la conception traditionnelle (Coleman, 1914), telle qu'elle est illustrée par la figure 6, en considérant que la focalisation contrastive appartient primitivement à la catégorie de la focalisation d'information. Conformément à la distinction introduite par Ladd (1980), un focus d'information peut être large (s'il concerne un énoncé complet ou un constituant majeur d'énoncé) ou étroit (s'il s'applique à un constituant mineur, à un item lexical, voire à une syllabe, comme c'est le cas dans la focalisation contrastive).

La focalisation d'emphase relève selon nous des fonctions impressive et expressive. Elle équivaut de ce point de vue à une focalisation d'intensification, c'est-à-dire à une forme de focalisation marquée (Monaghan, 1993) à laquelle Bolinger (1985) associe pour l'anglais une accentuation dite de 'puissance' (accents of power), alors qu'il attribue à la focalisation d'information une accentuation dite 'd'intérêt' (accents of interest). Ce dernier terme revêt cependant une signification plus générale pour cet auteur, dans la mesure où il recouvre à ses yeux une réalité plus large que le simple donné informationnel (habituellement représenté en termes d'oppositions binaires du type: support/apport, donné/nouveau ou thème/rhème) et où il permet ainsi d'intégrer à l'analyse des aspects de l'énonciation et de la contextualisation qui ne peuvent pas être pris en compte dans ce cadre trop réducteur (*Cf.* pour une discussion de ces aspects: Kerbrat-Orecchioni,1980; Blakemore, 1992; Auer et Di Luzio, 1992 et Vallduví, 1994).

Le cadre interprétatif de la focalisation d'information est variable selon les écoles et les auteurs. D'après Chomsky et Halle (1968), la marque formelle du focus, le nucleus, est attribuée sur la base de contraintes lexico-syntaxiques par la mise en œuvre de la célèbre 'règle d'accentuation nucléaire' (Nuclear Stress Rule). De son côté, Bolinger (1985) soutient l'hypothèse de la mise en relief sémantique ('semantic highlighting') selon laquelle les items lexicaux sont mis en focus en fonction du poids sémantique qu'ils assument dans l'énoncé. Enfin, pour Gussenhoven (1983) si le domaine focal fait bien l'objet d'un choix du locuteur, il se plie néanmoins à certaines contraintes imposées par les catégories logico-grammaticales. À la vérité, il ne nous semble pas que ces positions soient mutuellement exclusives si l'on envisage la focalisation dans le cadre d'une théorie comme celle de l'Optimalité (*Cf.* supra 1.4, pour les références). Dans une perspective similaire, Monaghan (1993) propose de tester l'hypothèse selon laquelle les facteurs qui déterminent l'attribution de l'accent focal constitueraient une hiérarchie incluant, dans un ordre d'importance décroissant, l'intention, le contexte, le poids sémantique des items lexicaux et l'ordre des mots. Nous nous rallions également à l'opinion de cet auteur selon laquelle le rôle du contexte est déterminant pour distinguer notamment une focalisation d'information contrastive d'une focalisation d'information non-contrastive (*Cf.* infra). Dans tous les cas, la sélection du focus doit satisfaire aux conditions de bonne formation et de cohérence du



https://doi.org/10.1017/S0959269500000120 Published online by Cambridge University Press

discours qui ne peuvent être évaluées que par référence au contexte dans lequel tout message trouve sa véritable raison d'être.

Dans notre approche de la focalisation, nous adoptons, en première analyse, comme cadre interprétatif, le modèle de la pertinence développé par Sperber et Wilson (1986). Ce modèle cognitiviste de la communication verbale se fonde sur la mise en œuvre de deux composantes interactives: la composante codée (la grammaire de la langue) et la composante ostensive-inférentielle, par lesquelles s'accomplit la visée de pertinence, c'est-à-dire la recherche d'un effet cognitif optimal pour un effort minimal de traitement mental. Rendre efficace la communication ostensive-inférentielle consiste, pour un locuteur, à rendre manifeste des informations permettant à l'allocutaire d'effectuer des inférences. L'inférence peut être définie comme un processus mental qui s'appuie sur un ensemble de prémisses supposées vraies pour aboutir à une conclusion probable validatrice d'une hypothèse (ou de plusieurs hypothèses) retenue comme pertinente parce qu'elle s'insère dans un contexte manifeste. Selon cette optique, la focalisation a valeur de déictique, en ce sens qu'elle attire l'attention de l'interlocuteur sur la partie du message qui contribue le plus efficacement à 'élargir l'environnement cognitif mutuel' des protagonistes, cet objectif constituant pour les auteurs l'essence même de la communication. Il est intéressant de constater par ailleurs, à la lumière des travaux des psycholinguistes et des phonéticiens, que le focus joue un rôle essentiel dans la reconstruction du contenu propositionnel du message en ce sens que, d'une part, la vitesse de traitement de l'information sémantique dépend fortement de la rapidité avec laquelle le focus est identifié (Cutler et Foss, 1977 ) et que, d'autre part, une accentuation inappropriée aux exigences du contexte discursif perturbe significativement la reconnaissance du message (Birch et Clifton, 1995).

Le modèle de la pertinence analyse la focalisation sur la base d'une distinction fondamentale entre implication de premier plan (foreground) et implication de second plan (background), sur laquelle repose la construction de l'échelle focale. Cette façon d'aborder la problématique de l'organisation informationnelle nous a semblé être plus 'opérationnelle' par sa flexibilité que le recours à des oppositions binaires du type: 'thème/rhème' ou 'connu/nouveau' dont l'emploi systématique se heurte à de multiples difficultés (Prince, 1981; Chafe, 1987; Dryer, 1996). De surcroît, la distinction premier plan/second plan ne recoupe pas nécessairement la dichotomie nouveau/connu, comme l'illustre l'exemple (3), ci-après. Imaginons une scène se déroulant dans un parc public. Une personne est assise sur un banc. Deux hommes se dirigent vers la personne assise et l'un d'entre eux s'adresse à cette personne en lui disant:

↘ → →

(3) *On voudrait vous parler*, mon frère et moi.

Nous avons mis en italiques ici le segment du message qui constitue dans ce



https://doi.org/10.1017/S0959269500000120 Published online by Cambridge University Press

contexte, selon l'échelle focale, le focus informationnel (implication de premier plan) de ce message (noter que flèche pointant vers le bas signale la chute mélodique spécifique de l'intonation terminale et les flèches horizontales, la mélodie plate de l'intonation dite 'parenthétique'). S'il s'agit, au moment où se déroule l'action, d'une première rencontre entre les trois personnages mis en scène, il va sans dire que la personne assise ignorait jusqu'alors que les deux personnages qui l'abordent appartiennent à la même famille. En ce sens, la totalité du message doit être considérée comme apportant de l'information nouvelle. Il semble, d'autre part, que le segment identifié ici comme le focus n'a pas, dans ce contexte, une quelconque valeur contrastive. Cette observation nous permet de situer notre position par rapport au débat selon lequel, à la différence de l'anglais, il serait impossible en français d'opérer un détachement de la sorte sans impliquer un contraste (Ronat, 1984; Sperber et Wilson, 1989; Hirst, 1989). À cet égard, nous estimons que le point de vue qu'exprime Monaghan (1993) au sujet de l'anglais vaut également pour le français, à savoir que la valeur contrastive du focus est le plus souvent inférable du contexte, ce dernier terme étant pris ici dans son acception la plus large qui inclut à la fois la situation et le contexte. Ainsi, si l'on considère les échanges (4a) et (4b):

(4a)
– C'est quoi, ce bruit?
↘ → →
– *C'est Jean-Pierre*, qui vient de rentrer.

(4b)
– Tu as entendu ce bruit? Je crois que Marie vient de rentrer.
↘ → →
– *Non, Jean-Pierre*, vient de rentrer.

Nous estimons que dans (4a) 'C'est Jean-Pierre' est interprétable comme une focalisation d'information non-contrastive malgré l'usage de la structure clivée, alors que dans (4b), même sans le recours à cette structure, 'Jean-Pierre' est identifiable comme une focalisation d'information contrastive. Nous ferons également remarquer qu'aucun des exemples précédents ne fait nécessairement intervenir la focalisation d'emphase qui selon nous relève, comme nous l'avons déjà noté, de la mise en relief d'un item lexical (ou d'un énoncé, dans certaines formes d'emphases cumulatives) à des fins impressives ou expressives et qui est révélatrice à la fois de l'intention et de l'attitude du locuteur. De ce point de vue, un segment de discours repéré comme un focus d'information peut fort bien contenir un ou plusieurs items porteurs d'une focalisation d'emphase.

Pour conclure ce développement consacré au thème de la focalisation que nous avons envisagée jusqu'ici dans une perspective strictement fonctionnelle, nous suggérons que la mise en forme prosodique du message joue un rôle déterminant dans le dispositif ostensif-inférentiel et l'expression du pacte de pertinence, dans la mesure où elle facilite considérablement l'émergence des



https://doi.org/10.1017/S0959269500000120 Published online by Cambridge University Press

implicitations contextuelles tout en réduisant de façon significative le coût du traitement cognitif. Nous n'utiliserons qu'un seul exemple pour illustrer ce dernier point. Soit le dialogue suivant:

(5)

– Tu as vu le film 'Les anges de la nuit'?

↘ → →

– *Je ne vais jamais voir* les films débiles.

Dans la réplique l'attention de l'auteur de la question est attirée sur le focus d'information que signale le déplacement de l'accent nucléaire de la dernière syllabe de l'énoncé (accentuation focale par défaut) à la syllabe 'voir'. Ce mouvement de l'accent nucléaire entraîne le déplacement du patron intonatif conclusif qui lui est associé et qui délimite ainsi la portée de ce focus. Le segment post-focal: 'les films débiles', privé de toute forme de relief mélodique, est généralement réalisé avec une mélodie uniformément plate dans un registre grave, ce qui correspond, dans les descriptions traditionnelles, à la manifestation phonétique de l'intonème dit de 'parenthèse basse' (Delattre, 1966). Sans trop nous étendre sur la complexité du dispositif cognitif mis en œuvre pour construire la forme propositionnelle d'un énoncé de ce type (*Cf.* Sperber et Wilson, 1989: 290–301), nous supposons que la structure prosodique de la réplique, dont la forme propositionnelle ne répond pas directement à la question posée, permet simultanément à l'auteur de la question d'accéder rapidement à l'information de premier plan qui est que la personne interrogée n'est pas allée voir le film dont il est question et de reconstituer la prémisse que ' 'Les anges de la nuit' est un film débile', celle-ci étant traitée ici comme une information de second plan. Une réponse du type:

↘

(6) *Je ne vais jamais voir les films débiles*

avec l'accent nucléaire assigné par défaut à la dernière syllabe de l'énoncé permettrait certainement d'accomplir les mêmes visées informationnelles, mais avec un temps de traitement accru pour l'allocutaire, car il lui faudrait alors attendre la fin de l'énoncé pour accéder à l'information de premier plan et reconstituer l'échelle focale.

## 4.2 *Présentation du modèle*

### 4.2.1 *Les bases antérieures du modèle*

Le modèle d'accentuation dont nous exposons les fondements au terme de cet article constitue en vérité un prolongement du modèle paramétrique *la* prosodie du français qui a fait l'objet d'une première présentation dans Hirst et Di Cristo (1984) et d'une révision dans Di Cristo et Hirst (1993). Dans la première version de ce modèle, qui s'applique à décrire le *système intonatif*



https://doi.org/10.1017/S0959269500000120 Published online by Cambridge University Press

*noyau* du français dans le cadre d'une approche autosegmentale, nous proposons d'associer la couche des segments phonémiques à celle des segments tonals par la médiation, au niveau sous-jacent, d'une structure intonative hiérarchique formée de deux types de constituants: les Unités Intonatives (UI) et les Unités Tonales (UT). À la différence de la théorie autosegmentale standard (Beckman et Pierrehumbert, 1986) les segments tonals B(as) et H(aut) qui constituent les *primitives* de la description ne sont pas assignés dans le modèle à des syllabes particulières ou à des accents, mais aux unités dont il vient d'être question. Ce choix conduit à la construction des gabarits prosodiques de l'UI et de l'UT qui sont représentés dans la figure 7.

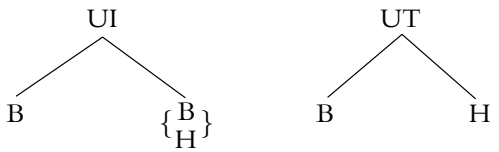



Figure 7. Gabarits tonals des UI (B-B et B-H) et de l'UT (BH) constitutifs du système – noyau de l'intonation du français (d'après Hirst and Di Cristo, 1984)

Nous avons postulé (Hirst et Di Cristo, 1986) l'existence de schèmes tonals similaires (B-B, B-H) dans la représentation sous-jacente des UI de base (terminale/non-terminale) du français et de l'anglais. Nous prétendons, en revanche, en prenant comme critère typologique le paramètre 'de localisation de la tête', que le schème de l'UT est du type (HB) en anglais et se trouve donc être inversé par rapport au français.

Cette dernière remarque permet de préciser la nature des liens que l'on peut établir entre la représentation tonale et la représentation métrique des énoncés. Dans l'approche paramétrique exposée en 1984, il est clair que l'unité tonale correspond à l'unité accentuelle minimale, c'est-à-dire au pied métrique, tel que l'a défini Abercrombie (1964). Ce lien était du reste implicite dans la formulation des règles de construction des UT en français que nous proposions alors. À ce propos, nous rappelons que les deux règles de base, dénommées: 'règle de proéminence initiale' et 'règle de proéminence finale' spécifient que l'on doit former une UT 'à partir de chaque syllabe jusqu'à la première syllabe incluse d'un mot lexical' et 'à partir de chaque syllabe jusqu'à la dernière syllabe incluse d'un mot lexical', respectivement. Une troisième règle, dite de 'réajustement', qui s'apparente à la règle rythmique de 'collision accentuelle' (« clash rule ») rend compte de la partition prosodique dans le cas d'une succession d'UT monosyllabiques. Cette troisième règle spécifie que si une UT monosyllabique est à la fois précédée et suivie d'une UT dans la même UI, les deux dernières UT de la séquence fusionnent pour former une seule



https://doi.org/10.1017/S0959269500000120 Published online by Cambridge University Press

UT. À titre d'exemple, une double application de la règle de réajustement permet de dériver (7b) de (7a):
(7a) (Un jo) (li) (ba) (teau)
(7b) (Un jo) (li bateau).

De même, la mise en œuvre des règles de proéminence initiale et finale dans Hirst et Di Cristo, 1984, qui constitue à notre connaissance la première tentative de traitement phonologique de l'accent initial en français, conduit à dériver l'organisation prosodique de l'exemple proposé dans (8) où les symboles ( ) indiquent les limites des UT:
(8) [(Il a ren) (contré) (les é) (crivains) (de la con) (testation)]

Cet exemple, dont nous supposons qu'il constitue une seule UI (comme dans: [Il a rencontré les écrivains de la contestation] [lorsqu'il a fait son reportage]) illustre de façon indirecte l'isomorphisme de l'organisation métrique et de l'organisation tonale qui constitue un parti pris de la première version du modèle paramétrique, car les pieds métriques et les UT sont présentés comme des unités coextensives. Nous avons précisé dans ce sens (Hirst et Di Cristo, 1986) que nous concevions le rythme et la mélodie comme l'actualisation de deux projections phonétiques d'une même représentation phonologique sous-jacente. Ce point de vue a fait l'objet d'une révision dans Di Cristo et Hirst (1993) où nous proposons d'introduire entre l'UI et l'UT, une unité prosodique de rang intermédiaire, l'Unité Rythmique (*UR*), pour rendre compte de la distinction entre accentuation primaire et accentuation secondaire. L'intégration de cette unité intermédiaire a pour conséquence que l'organisation prosodique de l'exemple donné en (8) est encodée sous la forme de l'exemple (9) où les symboles: < et > signalent les frontières gauche et droite des UR:
(9) [<(il a ren) (contré)> <(les é) (crivains)> <(de la con) (testation)>]

Il est aisé de déduire de cet exemple que si les UT correspondent aux pieds métriques, les UR coïncident en revanche avec ce qu'il a été convenu plus haut d'appeler les mots prosodiques.

Nous devons préciser que cette modélisation de la phonologie prosodique du français ne fait pas l'unanimité. C'est ainsi que nos propositions ont été récemment discutées par Jun et Fougeron (1995), qui considèrent que l'unité prosodique minimale définissable sur la base de l'organisation tonale est le 'syntagme accentuel' (accentual phrase), caractérisé par un patron mélodique du type B H B H. Selon ces auteurs, l'exemple (9) serait donc analysable comme une séquence de trois syntagmes accentuels (correspondant à nos mots prosodiques), ce qui exclut la prise en compte des UT, telles que nous les concevons. Sans vouloir entrer ici dans un débat de fond qui prendrait trop de place, nous rappellerons que nous avons présenté à notre tour un certain



https://doi.org/10.1017/S0959269500000120 Published online by Cambridge University Press

nombre d'arguments théoriques et des données empiriques qui nous incitent à maintenir notre position (Hirst et Di Cristo, 1996).

4.2.2 *Nouvelles perspectives*

Notre essai de l'accentuation est placé sous le signe de la *bipolarisation prosodique* qui procède de la mise en œuvre d'un principe de 'promotion des extrémités' applicable à l'ensemble des unités linguistiques (du mot au tour de parole). Selon ce principe, le mot, qui constitue l'unité accentuelle virtuelle minimale en français, est doté dans la représentation sous-jacente d'une proéminence initiale et d'une proéminence finale. Cette représentation est d'autre part assujettie au *principe de dominance finale*, qui attribue à l'accent final du mot la valeur potentielle d'accentuation primaire et à l'accent initial, celle d'accentuation secondaire. Ce choix est déterminé par le caractère facultatif de l'accent initial qui paraît être plus soumis à des contraintes phonotactiques (Pasdeloup, 1990) que ne l'est l'accent final du mot prosodique. Conformément aux deux principes susnommés, qui s'inscrivent dans le cadre théorique résumé dans la section précédente, la représentation prosodique sous-jacente du mot peut être spécifiée au moyen de la grille présentée dans la figure 8 qui définit son gabarit métrique:

| | | | |
|---|---|---|---|
| | | | x |
| x | | | x |
| x | x | x | x |
| Fé | li | ci | té |

Figure 8. Représentation de la forme prosodique sous-jacente du mot ''Félicité'' par une grille métrique simple

On pourraît voir dans cette façon d'aborder le problème un ralliement à l'hypothèse générale de Hyman (1977) d'après laquelle l'accentuation constituerait une forme de 'grammaticalisation de l'intonation' en vertu du fait que les traits suprasegmentaux qui caractérisent le mot isolé sont encodés avec ce dernier lorsqu'il est intégré à un énoncé. En fait, pour séduisante que soit cette hypothèse, elle ne permet pas de rendre compte des diverses stratégies accentuelles attestées dans les structures dérivées de surface dont il sera question par la suite.

Outre le mot, l'Unité Intonative constitue également une entité fondamentale de la représentation sous-jacente eu égard à sa planification précoce qui précède la mise en forme du matériau lexical par la syntaxe (Rossi et Peter-Defare, 1998). De ce point de vue, il semble qu'il existe une forte similitude entre l'émergence de cette unité et la genèse de certains schèmes gestuels, comme les battements, qui coïncideraient avec le moment où les idées prennent forme dans le cerveau (McClave, 1994). Pour cette raison, nous considérons que la métricité des Unités Intonatives (UI) est motivée en



https://doi.org/10.1017/S0959269500000120 Published online by Cambridge University Press

premier lieu par des contraintes sémantico-pragmatiques, ce qui n'exclut pas pour autant l'influence de la syntaxe qui est à même d'exercer ses propres contraintes à un niveau moins profond. Comme nous l'avons déjà évoqué, il existe un lien rigide, entre la planification des (UI) et le focus d'information en ce sens que le choix de ce dernier conditionne l'affectation dans la composante phonologique de l'accent nucléaire (ou tête nucléaire, sic!) et de son corrélat intonatif: l'intonème conclusif. L'accent nucléaire constitue donc dans notre approche, *l'élément terminal désigné* d'une suite de mots appartenant au constituant focal. La représentation métrique de ce dernier peut être illustrée par la figure 9, dans laquelle nous avons repris le formalisme de la grille parenthésée exposé plus haut (*Cf.* supra, 2.2). S'agissant d'une illustration théorique, nous n'avons pas jugé utile de prendre ici pour exemple un énoncé complet. Toutefois, pour faciliter la compréhension on pourra imaginer qu'il s'agit de la représentation sous-jacente d'un énoncé comme 'Il a *rencontré* les *écrivains* de la *contestation*' qui comprend les items accentogènes soulignés (A, B, C) et qui est supposée faire l'objet d'une focalisation large. Par exemple, la réponse à une question du type: 'Qu'est-ce qu'il a fait?'

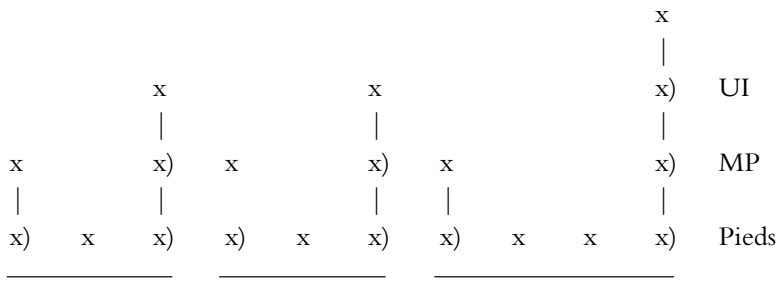



Figure 9. Représentation par une grille parenthésée de la structure métrique d'une UI correspondant à un constituant focal formé de trois mots accentogènes A, B, C

Le fait qu'aux deux premiers niveaux de la grille figurent parfois deux marques de proéminence contiguës n'est pas entièrement attribuable à la seule représentation des mots accentogènes (privés de leurs clitiques). Il veut signaler que cette contiguïté est acceptable dans la représentation profonde où ne se pose pas encore le problème de la prononçabilité de la séquence. Mais il apparaît, de surcroît, à la lumière de travaux récents et de nos propres observations, que la succession de deux accents est également attestée en surface en français, même si ces derniers ne sont pas séparés par une frontière d'UI, comme le suggère Pasdeloup (1990). Pour que la contiguïté soit acceptable, il suffit, comme le fait remarquer Mertens (1992), que les deux accents consécutifs soient de nature différente. Ceci rend compte par exemple de l'accentuation possible des syllabes adjacentes 'die' et 'or' dans 'La



https://doi.org/10.1017/S0959269500000120 Published online by Cambridge University Press

PROsoDIE ORgaNISE la paROLE'. Outre que cette observation est de nature à conforter le bien-fondé de notre hypothèse initiale de la bipolarisation lexicale, elle ébranle sérieusement la *règle rythmique de collision* ('clash rule') prétendument universelle qui, même dans le cas de l'anglais, ne semble pas mieux résister à l'épreuve des faits (Cooper et Eady, 1986).

Les modalités de cohabitation des contraintes métriques, dont les premières versions de notre modèle donnent un aperçu, et des contraintes sémantico-pragmatiques qui pèsent sur l'organisation prosodique des énoncés en usage, constituent le thème central de notre approche, dans la mesure où elles permettent de rendre compte de la diversité des motifs rythmiques attestés dans les formes de surface (*Cf.* nos remarques liminaires sur accentuation, mètre et rythme).

Les processus de traitement cognitif mis en œuvre dans l'exercice du langage impliquent, au niveau des représentations mentales, une mise en relation des objectifs fonctionnels du locuteur et des formes sonores sélectionnées pour atteindre ces objectifs. Nous considérons, pour simplifier, que l'acte de communication vise à atteindre deux objectifs majeurs: faciliter la compréhension du message et rendre manifeste l'intention de pertinence. Dans cette double perspective, parmi les multiples fonctions que l'on attribue généralement à la prosodie, deux s'imposent au regard de la généralisation (Bruce, 1985): la fonction de hiérarchisation (*Cf.* supra, la distinction: premier plan/second plan, relative à la focalisation) et la fonction de groupement (indicatrice des frontières et de la cohérence interne des unités ainsi délimitées). Afin de rendre compte des principales fonctions prosodiques qui relèvent de ces catégories, il est possible de les répertorier, ainsi que l'a proposé Hirst (1977), par un ensemble de traits binaires, tels que: [± accent], [± frontière], [± terminal], [± focus], [± emphase]. La discussion des principes qui sous-tendent la sélection de ces traits et leur mise en relation avec le principe fédérateur de bipolarisation constituera la matière de la dernière partie de cet article.

Nous rappelons qu'au niveau sous-jacent l'unité intonative constitue, dans notre approche, le domaine privilégié de l'organisation prosodique au sein duquel les mots accentogènes sont représentés par un gabarit métrique hiérarchique à deux têtes (*Cf.* figure 7), ce qui signifie que le trait [accent] est assignable à la syllabe initiale et à la syllabe finale du mot. L'actualisation de ces proéminences virtuelles, qui procède de la mise en œuvre des diverses stratégies d'encodage des structures rythmiques de surface (*Cf.* supra, section 1.1.) est assujettie aux objectifs précisés ci-dessus et à l'exécution des fonctions correspondantes.

L'eurythmie constitue probablement un élément important de l'organisation prosodique des énoncés, notamment en français où il est possible de combiner la signalisation des frontières des mots avec la recherche de l'équilibre rythmique, ce dernier pouvant être estimé, soit sur la base du nombre minimal de syllabes qui sépare deux accents consécutifs, soit sur celle de l'empan



https://doi.org/10.1017/S0959269500000120 Published online by Cambridge University Press

critique de la fenêtre temporelle inter-accentuelle (Di Cristo et Hirst, 1997). À ce sujet, il nous paraît édifiant de souligner que si l'information prosodique relative à la structure rythmique de la langue est bien utilisée de façon ascendante pour contraindre le processus de segmentation lexicale, ce rôle est davantage assumé, contre toute attente, par les accents non-lexicaux (comme c'est le cas du français) que par les accents proprement lexicaux (McQueen et Cutler, 1997). La visée eurythmique, qui tend à éviter (sans toutefois l'exclure, *Cf.* supra) la contiguïté des proéminences, devrait se traduire en surface par une alternance relativement régulière des accents initiaux et finaux au sein de l'UI. Dans cette perspective, l'actualisation de l'accent initial assume une double fonction, car il participe au rythme et il signale le début de l'unité lexicale à laquelle il est affecté. L'accent final participe également à la construction du rythme et démarque la fin de l'unité lexicale qui le reçoit. Le principe de bipolarisation lexicale peut toutefois être annihilé par la mise en action d'une contrainte pragmatique qui réclame une accélération du tempo (Uhmann, 1992), comme c'est le cas notamment pour l'incise. On obtient alors une réduction drastique du nombre des groupes accentuels (Horne, 1978; Fougeron et Jun, 1996) qui est rendue possible par l'inhibition de l'accent initial et le groupement de plusieurs items lexicaux en une seule unité accentuelle.

D'un autre point de vue interprétatif, nous considérons qu'il existe une dissymétrie fonctionnelle potentielle entre l'accent initial et l'accent final, dans la mesure où le premier est, en règle générale, générateur d'emphase et le second de frontière. Comme on l'a noté précédemment, l'accent d'insistance (marqueur du focus d'intensification) a tendance à frapper systématiquement la première syllabe du mot, quel que soit par ailleurs son statut grammatical, si l'on en juge par le fait que les clitiques monosyllabiques sont également 'emphatogènes'. À ce propos, il est à remarquer que les clitiques ainsi focalisés (comme dans, par exemple '*IL ne l'a pas fait exprès*', où les capitales indiquent la syllabe accentuée) n'engendrent pas la promotion d'une frontière (Di Cristo et Chafcouloff, 1975), ce qui constitue également un critère susceptible d'accréditer notre distinction.

Ces remarques nous conduisent à étayer l'hypothèse, déjà formulée dans cette étude, qu'il est possible de projeter à la surface deux types d'accent initial: l'accent initial rythmique (dont la projection procède uniquement de l'attribution du trait [accent]) et l'accent initial d'emphase (dont la projection résulte de l'affectation des traits [accent] et [emphase]). Quels sont les critères qui président à la sélection de l'un ou de l'autre de ces accents? Pour fournir des éléments de réponse à cette question, nous faisons référence à la 'Théorie Adaptative' de Lindblom (1990), qui stipule que le locuteur sélectionne une expression Hyper-articulée ou Hypo-articulée, en fonction de la demande de l'allocutaire ou de l'auditoire. Nous considérons, au regard de cette théorie, que l'accent emphatique représente la forme de surface marquée (ou Hyper) de l'accent initial (ce qui vient d'être confirmé expérimentalement par une



https://doi.org/10.1017/S0959269500000120 Published online by Cambridge University Press

étude récente (Astesano *et al.*, 1997). Il n'est pas étonnant, dans cette perspective, que son usage systématique soit associé à certaines formes de discours pour lesquels l'efficacité perlocutoire de l'acte de langage est cruciale, comme le discours didactique, le discours politique ou le style publicitaire (Türk, 1994). De même, il est normal que son usage soit moins fréquent dans la conversation où il n'est pas nécessaire que la fonction impressive s'exerce de façon continue. En effet, l'accent initial d'emphase peut être utilisé ponctuellement pour focaliser un item considéré comme particulièrement important dans la lecture ou dans la conversation (focalisation étroite). Il peut être également employé de façon itérative, soit à court terme, traduisant ainsi une forme de contextualisation discursive qui implique un engagement plus marqué du locuteur envers une partie de son message, soit à long terme où il dénote le choix d'un phonostyle imposé par la situation et la visée communicative. Ceci revient à dire que nous considérons l'accent dit 'didactique' comme une variante de l'accent emphatique dont la nature graduelle permet également de signaler divers degrés d'insistance.

L'accent final peut être considéré principalement comme un attracteur de frontière en ce sens qu'il sous-tend l'attribution des traits [frontière] et [focus], ce dernier lui conférant alors le statut d'accent nucléaire (AN). AN constitue, comme on l'a vu, la tête du constituant focal (focus large) dont la sélection entraîne automatiquement l'affectation de cet accent à la dernière syllabe de l'item lexical appartenant à cette unité. Nous rappelons à ce propos que le déplacement de AN dans l'énoncé est régi par plusieurs facteurs parmi lesquels l'intention du locuteur et le contexte occupent une place de premier plan. Une focalisation large peut avoir pour domaine une tranche de discours qui est segmentable en plusieurs UI non-terminales afin de satisfaire à des contraintes d'ordre sémantico-pragmatique et/ou syntaxique. Dans ce cas, le trait [frontière] est assigné à chacune des syllabes finales des items lexicaux qui se trouvent à la périphérie de la proposition ou du constituant syntaxique qui doit faire l'objet d'une délimitation. Il va de soi que la double distinction accent initial/accent final et accent initial rythmique/accent initial emphatique que nous venons de poser demande à être validée empiriquement. Les données expérimentales restreintes dont nous disposons actuellement (Astesano, Di Cristo et Hirst, 1995) accréditent l'idée avancée par plusieurs auteurs que l'accent initial est un accent mélodique, alors que l'accent final est doublement marqué par la mélodie et par la durée, ce dernier paramètre s'avérant prééminent. Plusieurs paramètres concourent à la manifestation de l'accent initial emphatique et permettent de le différencier de l'accent initial rythmique. C'est ainsi que, outre l'importance de l'amplitude de la variation tonale (Dahan et Bernard, 1996; Astesano *et al.*, 1997), on note la présence d'une pause devant la syllabe accentuée, le renforcement de la consonne initiale, et l'incrémentation significative de l'intensité du noyau vocalique. D'autres effets plus subtils, que nous n'avons pas la place de décrire ici,



https://doi.org/10.1017/S0959269500000120 Published online by Cambridge University Press

concernent l'organisation temporelle des éléments constitutifs de la syllabe (attaque, noyau, coda) qui semble différer d'une catégorie accentuelle à l'autre.

Bien que les contraintes rythmiques aient un rôle déterminant à jouer dans la mise en forme prosodique du discours, elles ne suffisent pas à expliquer tous les dispositifs accentuels qu'une analyse approfondie de divers phonostyles conduit à mettre en lumière. Le premier exemple qui peut faire l'objet d'une discussion sur ce point concerne la formation des 'arcs accentuels'. La contrainte rythmique de non-contiguïté peut rendre compte sans difficulté du schème accentuel de 'SOUliers NOIRS'. De même, il est possible d'interpréter des exemples comme 'BEAUjolais nouVEAU, 'DIScussion stéRILE', 'CENT grammes de BEURRE', 'REdécouverte de L'AUtre' en invoquant une règle spécifiant que deux accents successifs doivent être séparés par plus d'une syllabe inaccentuée et que si cette condition n'est pas remplie, un accent final de mot doit être déplacé sur la syllabe accentuable la plus proche à sa gauche, soit sur la syllabe initiale. Mais que dire alors d'un exemple comme 'THEorie de la littéraTURE', qui est également attesté dans la conversation courante et qui participe du même principe de bipolarisation accentuelle. Notre interprétation est que la formation de ces schèmes n'est pas motivée exclusivement par le rythme, mais aussi par la *visée* pragmatique de mettre en relief un syntagme au sein duquel les éléments sont fortement liés par le sens et de participer ainsi à la construction de l'échelle focale.

L'option de l'arc accentuel a pour conséquence la mise en œuvre d'un principe d'inhibition qui efface toutes les proéminences comprises entre les deux accents périphériques, ce qui permet d'expliquer la survivance d'ambiguïtés comme dans la suite 'SInistre industriEL' (exemple authentique entendu à la radio) à partir de laquelle il n'est pas possible de dire si le substantif précède l'adjectif ou inversement. Sans recourir à la formation d'un arc, l'accent initial par sa seule présence peut également constituer une source d'ambiguïté comme c'est le cas dans 'L'AnaLYSE' versu 'L'Ane à LISE' (exemple fabriqué pour les besoins de la démonstration!) où la distinction de sens peut être neutralisée.

À l'instar de l'accent emphatique, le dispositif de l'arc accentuel (qui n'a pas nécessairement une valeur emphatique) permet d'insérer un (ou plusieurs) élément focalisé au sein d'un focus plus large que nous avons appelé focus d'information, ou d'un segment de discours qui ne participe pas de cette fonction. À cet égard, il est remarquable d'observer que ce procédé est fréquemment utilisé dans la conversation pour signaler ce focus informatif et contrebalancer ainsi l'effet de la contrainte universelle de positionnement linéaire (Schmerling, 1976) qui, en déplaçant l'accent nucléaire à la position droite extrême du constituant focal, ne permet d'effectuer qu'un bornage unilatéral de ce constituant, rendant ainsi ambiguë la portée du focus (Ladd, 1996). Nous considérons que le principe de bipolarisation accentuelle s'applique également au focus informatif qui se trouve ainsi délimité à sa droite par l'accent nucléaire et à sa gauche par un accent que nous appellerons



https://doi.org/10.1017/S0959269500000120 Published online by Cambridge University Press

pour l'instant, faute de mieux, 'accent initial de focalisation', pour le différencier de l'accent initial rythmique. Un exemple emprunté à notre description de l'intonation du français (Di Cristo, 1998) permettra d'illustrer les emplois de ce dispositif de bipolarisation à des fins pragmatiques (noter que les accents sont indiqués par des capitales et les focus d'information par des italiques).

(10)

Question - De quoi avez-vous parlé hier soir avec ces individus?

Réponse (a) - Hier SOIR, avec ces *SInistres industriELS*, nous avons *LONguement parlé de poliTIQUE.*

Réponse (b) - *DE poliTIQUE*, nous avons discuTE hier SOIR.

On remarquera que dans la réponse (a), le premier arc accentuel ('sinistre industriel') est hors de la portée du focus d'information ('longuement parlé de politique') qui est pour sa part repéré par un dispositif du même type. On notera également que dans la réponse (b), l'accent initial de focalisation est affecté à la préposition 'de', ce qui montre que cet accent s'affranchit des contraintes imposées à l'accent initial rythmique qui n'est assignable en principe qu'à la première syllabe des mots accentogènes. En outre, dans cette réponse, la contrainte pragmatique d'effacement des proéminences mélodiques dans le segment post-focal bloque l'attribution d'un accent initial rythmique à l'item 'discuté', sans pour autant interdire celle de l'accent final à ce mot et à 'soir'.

Le dispositif de bipolarisation, qui est au cœur même de notre tentative de modélisation, semble également s'appliquer à la manifestation de la focalisation contrastive (ou focalisation d'information étroite) dont il a été question plus haut. Comme nous le montrons dans Di Cristo (1998), l'accent initial de focalisation bénéficie dans ce cas d'une grande mobilité puisqu'il peut théoriquement être assigné à chacune des syllabes du syntagme en contraste et déroger ainsi à l'application du principe rythmique, comme l'illustrent les exemples suivants:

(11)

Question - C'est la directrice qui a pris la clé?
Réponse (a) - Non, c'est *LE direcTEUR, qui l'a prise.*
Réponse (b) - Non, c'est le *DI recTEUR*, qui l'a prise.
Réponse (c) - Non, c'est le *diRECTEUR* qui l'a prise.
Réponse (d) - Non, c'est le *direc***TEUR**, qui l'a prise!

Dans toutes les réponses possibles à la question posée, la syllabe 'teur' porte l'accent nucléaire. On remarque que l'accent initial de focalisation est assigné successivement à 'le', 'di' et 'rec', ce qui entraîne dans ce dernier cas une contiguïté accentuelle. Nous interprétons la dernière réponse (d) comme l'aboutissement d'une superposition de l'accent nucléaire et de l'accent initial



https://doi.org/10.1017/S0959269500000120 Published online by Cambridge University Press

de focalisation au sein d'une même syllabe, cet amalgame étant selon nous générateur d'une emphase cumulative à valeur expressive. C'est la raison pour laquelle nous avons noté en caractères gras cette syllabe.

Cette dernière remarque nous amène à évoquer de nouveau la problématique de l'emphase. Nous considérons qu'il existe deux types de focalisation emphatique: une focalisation d'emphase étroite, qui concerne un item particulier et qui implique le recours à l'accent initial dit d'insistance (emphase impressive) ou à un procédé de cumul de l'accent nucléaire et de l'accent initial de focalisation dans le cas du contraste (emphase expressive), et une focalisation d'emphase large, qui s'applique à une proposition complète ou à l'ensemble de l'énoncé. Dans ce dernier cas, on relève deux stratégies d'encodage. La première consiste à affecter l'accent emphatique à la dernière syllabe de l'énoncé, ce qui se traduit par un renforcement notable de cette dernière et une augmentation drastique de l'amplitude de la variation tonale montante - descendante qui accompagne son émission (*Cf.* la courbe caractéristique de l'intonème d'implication décrit par Delattre, 1966). Cette forme d'accentuation est à rapprocher du dispositif de 'climax' (ou d'accentuation de cumul) décrit par Bolinger (1985–1986) pour l'anglais. La seconde stratégie se distingue de la précédente par le fait que l'on assigne un accent emphatique supplémentaire à la première syllabe de l'énoncé (ce qui constituerait pour Bolinger un 'anti-climax'). De ce fait, l'énoncé se trouve doublement marqué par l'emphase et acquiert ainsi un profil mélodique global en creux, comme ce serait le cas dans l'exemple (12) où l'analyse phonétique révèle la présence d'un mouvement mélodique descendant de la syllabe initiale à la syllabe 'pas', suivie d'un mouvement ascendant qui culmine avec la syllabe finale.

(12) JE ne te critique pas parce que tu es mon PERE!

Il apparaît donc, à la lumière de ce dernier exemple, que le principe de bipolarisation accentuelle est généralisable et s'applique également à l'emphase d'énoncé. On retiendra donc en définitive que le principe de bipolarisation accentuelle est doublement motivé, par des contraintes métriques qui régissent le jeu de l'accentuation initiale et finale des items lexicaux accentogènes, et par les diverses formes de focalisation, emphatique et non-emphatique. Ces dernières induisent la formation de schèmes accentuels qui prennent généralement appui sur les gabarits métriques de langue. La diversité de ces schèmes dont on commence à percevoir la complexité structurelle constitue probablement l'une des principales sources de la variabilité rythmique attestée dans les formes prosodiques de surface.

## CONCLUSION

Après un long détour que nous espérons utile et qui nous a conduit, d'une part à discuter certains concepts fondamentaux relatifs à la problématique de



https://doi.org/10.1017/S0959269500000120 Published online by Cambridge University Press

l'accentuation en général et, d'autre part, à établir un bilan des connaissances sur l'accentuation du français, nous exposons, dans la dernière partie de cet article l'ébauche d'une modélisation du cadre accentuel de cette langue.

Il va sans dire que cette tentative est incomplète, dans la mesure où nous avons pris le parti de nous consacrer presque exclusivement à des considérations d'ordre fonctionnel, ce qui nous semble fondé dans le cadre d'une démarche qui souhaite prendre en compte les divers aspects cognitifs de la communication. Pour être complet et en conformité avec notre approche en termes de niveaux de représentation, le modèle doit également s'attacher à décrire les relations entre les fonctions et les formes prosodiques de surface qui les rendent tangibles. Il doit également montrer, pour satisfaire au principe d'interprétabilité, comment les représentations phonétiques actualisent ces formes prosodiques et, ultimement, quelles sont les stratégies d'encodage des paramètres physiques qui sous-tendent ces représentations. Ces perspectives de recherche font partie de nos préoccupations actuelles et donneront lieu à des publications ultérieures.

Nous avons surtout voulu montrer dans cette étude que si l'on est en droit de contester l'idée établie de l'excessive simplicité du système accentuel du français, le principe de bipolarisation, qui est à la base de notre essai de modélisation, constitue peut-être l'une des clés permettant d'expliquer sa réelle complexité.

Nous tenons à remercier Jacques Durand, Mario Rossi et Elisabeth Delais-Roussarie qui ont bien voulu lire notre manuscrit et nous prodiguer leurs précieux conseils.

*Adresse de l'auteur:*
*Institut de Phonétique*
*Université de Provence*
*29, avenue R. Schuman*
*13621 Aix-en-Provence cedex 1*
*France*
*e-mail: albert.dicristo@lpl.univ-aix.fr*



https://doi.org/10.1017/S0959269500000120 Published online by Cambridge University Press