HAL
open science

La prosodie au carrefour de la phonétique, de la phonologie et de l'articulation formes-fonctions

Albert Di Cristo

To cite this version:

Albert Di Cristo. La prosodie au carrefour de la phonétique, de la phonologie et de l'articulation formes-fonctions. Travaux interdisciplinaires du Laboratoire Parole et Langage, 2004, 23, pp.67-211. hal-00285554

HAL Id: hal-00285554

https://hal.science/hal-00285554v1

Submitted on 5 Jun 2008

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

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

LA PROSODIE AU CARREFOUR DE LA PHONÉTIQUE, DE LA PHONOLOGIE ET DE L'ARTICULATION FORMES-FONCTIONS

Albert Di Cristo

à la mémoire de Georges Faure,
fondateur de l'Institut de Phonétique d'Aix
et pionnier en la matière.

Résumé

Cet article traite des aspects de la prosodie qui concernent l'analyse linguistique et à la cognition. Il vise à présenter un survol des problématiques de l'interprétation des éléments prosodiques qui se rapportent à leur matérialité, à leur structure formelle et aux fonctions linguistiques et paralinguistiques qu'ils assument dans l'usage du langage et la communication. La première partie (arrière-plan) expose les raisons du succès actuel des recherches prosodiques. La seconde présente d'abord l'architecture formelle de la prosodie, sa substance, sa place dans le système linguistique et des considérations terminologiques. Elle discute ensuite les difficultés inhérentes à l'analyse prosodique en relation avec les articulations substance/forme et substance/forme/fonction. La troisième partie commente les principales approches linguistiques de la prosodie (linéaires traditionnelles, multilinéaires, métriques et autosegmentales). Elle discute enfin les concepts de constituance et de domaines prosodiques et s'interroge sur leur validité psychologique à la lumière des recherches récentes de la psycholinguistique sur les processus de production et de perception.

Mots-clés : accent, intonation, métrique, phonétique, phonologie, prosodie.

Abstract

This papers deals with those aspects of prosody which relate both to linguistic analysis and to cognition. It aims at presenting a comprehensive synthesis of issues linked to the interpretation of prosodic phenomena, in relation with their substance, their formal architecture and the linguistic and paralinguistic functions which they fulfil in language use and communication. The first part of this paper, under the heading « background », explicits the reasons for the popular role currently played by prosody within speech sciences and related domains. We then propose a set of remarks about the status of phonology, its relations with phonetics, and the consequences of these relations on prosodic analysis. The second part falls into two phases: first, we focus on the formal architecture of prosody, its substance, the place it holds within the linguistic system, and terminological considerations; we then emphasize the inherent difficulty of prosodic analysis, in relation with the articulation not only of substance and form, but also of substance, form and function. The third part presents current linguistic approaches to prosody (traditional linear, multilinear, metrical and autosegmental). Finally, this presentation discusses the concepts of constituency and prosodic domains, together with their psychological validity, in the light of recent psycholinguistic research on production and perception processes.

Keywords : accent, intonation, metrics, phonetics, phonology, prosody, stress.

DI CRISTO, Albert (2004), La prosodie au carrefour de la phonétique, de la phonologie et de l'articulation formes-fonctions, *Travaux Interdisciplinaires du Laboratoire Parole et Langage,* vol. 23, p. 67-211.

Introduction

Cet article traite principalement des aspects phonétiques, phonologiques et fonctionnels de la prosodie. L'emploi du mot *aspect* veut signifier que nous proposons d'aborder sous des éclairages divers des problématiques essentielles se rapportant à la phonétique, à la phonologie et à la fonctionnalité prosodiques, sans prétendre les développer dans le détail ni les résoudre. Dans cette perspective, les limites de notre contribution sont imposées à la fois par l'ampleur de la question à traiter (qui justifierait l'écriture d'un ouvrage) et par l'instabilité des acquis qui s'y rapportent. Il est manifeste que des sujets tels que : la place des faits prosodiques dans la description et la représentation des systèmes phonologiques, la définition des concepts et des termes prosodiques, la spécification des niveaux d'analyse de la prosodie et de leur articulation, la nature des relations que la composante prosodique entretient avec les autres composantes de la grammaire (tels que la syntaxe et la sémantique) et les phénomènes prétendus extra-grammaticaux (ou pragmatiques), constituent des thématiques particulièrement denses qui ont déjà animé de nombreux débats. De surcroît, comme les discussions de ces thèmes ont mobilisé de façon croissante l'intérêt des chercheurs, il en est résulté un foisonnement de propositions théoriques et méthodologiques, cumulatives ou antagoniques, qui s'avère aujourd'hui particulièrement enrichissant et stimulant, mais qui éveille conjointement la perplexité en attirant l'attention du chercheur sur la multiplicité des problèmes qui restent à clarifier. En dépit de sa longueur, le texte de cet article n'est qu'un survol dont l'objectif est d'une part de situer les diverses problématiques se rapportant à l'analyse, à la représentation et à l'interprétation de la prosodie et, d'autre part, d'attirer l'attention du lecteur sur des thèmes de discussion qu'il pourra prolonger en se rapportant aux références que nous avons rassemblées à cet effet dans le corps de cette étude.

Au cours des dix dernières années, les études prosodiques ont suscité un engouement hors du commun qui a gagné non seulement les domaines couverts par les sciences du langage (phonologie, syntaxe, sémantique, pragmatique), mais également des disciplines apparentées comme la psycholinguistique, la neurolinguistique et le traitement automatique des langues. Cette propagation soudaine de l'intérêt pour la *chose prosodique*, pourtant si souvent négligée naguère (au point qu'on a pu la qualifier de « *cendrillon de la communication* », *cf.* Bolinger, 1986 ; Fónagy, 1989), pourrait passer pour une simple épidémie de *prosodimania* (ce qui est partiellement vrai), si elle n'était en réalité porteuse d'un esprit d'ouverture salutaire qui favorise un décloisonnement des disciplines et qui attise ainsi la volonté d'appréhender les phénomènes linguistiques dans l'intégralité de leur espace cognitif.

Compte tenu de cette évolution significative, il est indispensable de s'interroger sur la situation qu'occupe actuellement la prosodie dans l'étude linguistique de la phonie des langues (que partagent la phonétique et la phonologie) et sur les particularités des diverses mouvances qui participent à cette entreprise. Cet article s'attache à l'examen de ces questions en adoptant une composition en trois grandes parties. La première partie, intitulée « arrière-plan », est formée de deux sous-sections qui concernent respectivement : les raisons de l'émergence de l'intérêt envers la prosodie, des considérations générales sur la phonétique et la phonologie, ainsi que leurs retombées sur l'analyse prosodique. La seconde partie traite de la quiddité de la prosodie (entendons par là : de son essence et de sa matérialité) et de la problématique des relations entre formes et fonctions qu'elle soulève. Enfin, la troisième partie constitue une présentation illustrée des différentes approches linéaires et plurilinéaires de la prosodie. Elle accorde une place importante à l'exposé des approches métrique et autosegmentale qui représentent le cadre théorique dominant de la *phonologie prosodique* actuelle (cette expression étant prise ici dans son acception la plus large ; *cf.* 3.2.3.1. pour une autre interprétation), ainsi qu'à l'examen de la question des domaines et/ou des constituants, qui a conduit notamment à s'interroger, de façon récurrente, sur la nature des relations de la prosodie à la syntaxe. Nous tentons de montrer en dernier lieu comment cette question s'intègre aux débats actuels de la psycholinguistique se rapportant à la production et à la compréhension de la parole.

1. Arrière-plan

1.1. Les principales raisons de l'émergence de la prosodie

« *La prosodie organise la parole* » ; « *la psycholinguistique ne peut échapper à la prosodie* » ; « *la prosodie est trop importante pour être confiée uniquement aux linguistes* ». Ces formulations pour le moins percutantes, que nous devons à d'éminentes figures de notre communauté scientifique (Mary Beckman, Janet Fodor et Dwight Bolinger, respectivement) suffiraient, si besoin était, à confirmer la prise de conscience de l'impact qu'exerce la prosodie sur l'ensemble des études qui se rapportent au langage et à la parole. Nous évoquons brièvement dans les lignes qui suivent les principales raisons qui sont à l'origine de l'effloraison récente de la prosodie dans ces travaux. L'une de ces raisons concerne à l'évidence l'émergence, dans les années 70-80, de la *phonologie non-linéaire* ou multidimensionnelle (Durand, 1990, 1993) qui, en promouvant *l'autonomie* de la prosodie, a permis ainsi l'élaboration d'outils méthodologiques et théoriques mieux adaptés à sa description et à sa modélisation. Une autre raison tient certainement à l'évolution des recherches en linguistique, qui ont progressivement déplacé leur centre d'intérêt de l'étude du langage-objet vers celle de ses usages (sans négliger la première pour autant). Il est manifeste, en effet, que la « *perspective pragmatique* »

(Verschueren, 1994) a conduit à revisiter la *pluralité des fonctions* de la prosodie dans l'exercice du dialogue et de la conversation quotidienne, et non plus seulement dans le contexte appauvri de la parole dite « *de laboratoire* » et des exemples construits (*cf.* Rossi, 1987 ; Couper-Kuhlen, 1993 ; Couper-Kuhlen & Selting, 1996 ; Swerts & Hirschberg, 1998 ; Wichmann, 2000 ; Di Cristo, 2000b ; Simon, 2003 ; Di Cristo *et al.*, 2004 ; Portes, 2004). La perspective pragmatique a ouvert à la prosodie des horizons que la mouvance « *pan-syntaxiste* » (ou « *syntactocentriste* », selon l'expression de Hagoort *et al.*, 1999) d'obédience générativiste avait plutôt eu l'effet d'occulter. Du reste, le règne de cette mouvance, jusqu'à l'avènement du courant « *néo-générativiste* » de la phonologie *plurilinéaire*, n'a pas donné lieu à des travaux marquants en prosodie, au point qu'il est difficile de citer des références notoires autres que celles de Bierwish (1966) et de Stockwell (1960, 1972) ou encore les études centrées sur la problématique de l'accentuation nucléaire en anglais (Chomsky & Halle, 1968 ; Berman & Szamosi, 1972 ; Bresnan, 1971, 1972 ; Lakof, 1972).

Il faut également voir dans la découverte récente, par les psycholinguistes, du rôle décisif de la prosodie dans l'acquisition du langage (notamment, de sa fonction « *d'amorçage* » de l'acquisition de la grammaire) un facteur de motivation important pour la recherche prosodique (de Boysson-Bardies, 1996 ; Morgan & Demuth, 1996 ; Jusczyk, 1997, Nazzi *et al.*, 1998 ; Ramus *et al.*, 1999 ; Christophe *et al.*, 2003). Il en est de même des travaux en neurolinguistique qui ont incité les chercheurs à réévaluer, dans l'optique stimulante de l'élaboration d'une « nouvelle anatomie fonctionelle du langage » (Poeppel & Hichkok, 2004), la place occupée par la prosodie dans l'architecture neurale et fonctionnelle du cerveau/esprit, en particulier au regard de la récurrente problématique de la *latéralisation hémisphérique* des facultés langagières (Baum & Pell, 1997 ; Dogil et al., 2002 ; Baum et Dwivedi, 2003 ; Astesano *et al.*, 2004). Enfin, il faut rendre aussi justice aux progrès considérables qui ont été réalisés dans la conception de logiciels spécialisés qui se sont montrés très utiles pour l'analyse de la prosodie et son traitement statistique à partir de vastes bases de données qui se sont avérées de plus en plus facilement accessibles (Sagisaka *et al.*, 1997).

1.2. Phonétique et phonologie

1.2.1. Généralités

La tradition admet que la phonétique et la phonologie se partagent les tâches qui incombent à l'étude scientifique de la face sonore (ou de la phonie) des langues. Dans cette optique, la phonologie vise plus particulièrement à construire une représentation de la *connaissance implicite* de la structure des formes sonores contrastives dont disposent les locuteurs d'une langue pour transmettre du sens. De son côté, la phonétique s'attache davantage à décrire et à représenter une *cartographie* de l'espace

articulatoire et acoustique que ces locuteurs utilisent pour procéder à l'encodage et au décodage de cette structure. La phonologie peut être définie de façon concise comme « *l'étude des systèmes de sons linguistiquement signifiants* » (Kaye ; 1989). Cette définition et les commentaires qui l'accompagnent dans l'ouvrage rédigé par cet auteur mettent en avant l'idée que la phonologie n'étudie pas tant les langues que les systèmes qui sous-tendent leur fonctionnement. En effet, dans son orientation moderne, la phonologie s'intéresse particulièrement aux *processus cognitifs* qui supervisent une certaine compétence grammaticale capable d'exploiter des variables abstraites telles que les catégories : P (phonème), C (consonne), V (voyelle), S (syllabe), μ (more), T (ton), etc.

La productivité générative de la phonologie (à l'instar de celle de la syntaxe) plaide en faveur d'une théorie qui soutient que la phonologie d'une langue a le caractère d'une *grammaire formelle*. Cette référence à la grammaire dénote que la théorie propose d'établir l'inventaire des catégories abstraites et d'énoncer les principes qui précisent les conditions de bonne formation de l'agencement de ces catégories. En conséquence, la grammaire phonologique doit inclure une *architecture structurelle* et une *phonotactique* (cette dernière devant spécifier notamment comment l'architecture s'adapte à la contrainte de linéarité du langage proféré). D'autre part, la confrontation des systèmes phonologiques des langues a permis de mettre en lumière de *nombreuses parentés*, tant en ce qui concerne l'inventaire des catégories que les principes qui régulent leur combinatoire, ce qui suggère l'existence d'une grammaire universelle de la structure sonore (Pierrehumbert, 2001). Cette grammaire contient alors à la fois la spécification de l'ensemble des principes qui rendent compte des similitudes entre les langues et celle des paramètres qui attestent de la variation qui les différencie.

La grammaire phonologique, telle que la conçoit le modèle de la phonologie générative classique, opère exclusivement sur des *entités catégorielles*, des suites bien formées, et écarte ainsi l'idée que la phonologie puisse impliquer une quelconque *connaissance statistique*. Ce point de vue a été récemment débattu par Pierrehumbert (2001, 2002) qui invoque notamment une différence fondamentale entre les modèles de traitement des données variables mis en œuvre par le chercheur et par un *quidam*. Alors que le premier abstrait de ces données un modèle purement catégoriel, le second développe un modèle cognitif dans lequel la fréquence d'occurrence des variables joue un rôle central. Bien que le système cognitif demeure fondamentalement *grammatical* (dans le sens où il contrôle la bonne formation des formes complexes à partir de leurs sous-parties), il possède la capacité de construire une *grammaire probabiliste* en vertu du fait que les aspects fréquentiels et distributionnels ont un rôle déterminant dans l'acquisition de la compétence phonologique, de la compétence phonétique, dans la production et dans la perception de la parole, ainsi que dans la construction des représentations mentales à long terme. L'un des aspects porteurs du développement de la phonologie probabiliste est qu'elle incite à réexaminer les *relations de la phonétique à la phonologie*, car de nombreux travaux attestent

également d'une connaissance probabiliste de l'espace phonétique. Selon Pierrehumbert (2001), l'interface entre les représentations phonologiques et les variables phonétiques devrait être modélisé en recourant à une distribution probabiliste de l'espace phonétique. Cependant, il n'existe pas actuellement une théorie susceptible de proposer un traitement intégré de tous les effets probabilistes en phonétique et en phonologie, bien que l'étude impressionnante de Boersma (1998), semble approcher cet objectif. Pour ce qui concerne la prosodie, on pourra se reporter en particulier aux travaux récents de Dainora (2001a, 2004).

Le paradigme de la phonologie classique admet que la phonétique et la phonologie ont en commun la prononciation des langues (au sens large du terme) comme objet d'analyse. Du point de vue de la production, les dispositifs phonologiques sont responsables d'un plan de prononciation qu'il incombe aux dispositifs phonétiques d'implémenter et d'exécuter. Il est évident que la construction mentale de ce plan (ou encodage phonologique) est motivée au niveau le plus profond par la signification de l'énoncé à produire (Levelt, 1989). Ces remarques d'ordre général appellent plusieurs commentaires.

Le modèle de la phonologie segmentale traditionnelle considère que la construction du plan s'effectue sur la base d'une sélection des entités catégorielles contrastives appropriées à l'élaboration d'un message particulier dans une langue donnée. Ces entités catégorielles contrastives représentent les « sons linguistiquement signifiants » dans la définition proposée par Kaye (1989). Toutefois, l'expression peut s'avérer ambiguë si l'on a à l'esprit que les unités minimales de la phonologie segmentale, les phonèmes, ne sont pas signifiantes par elles-mêmes en ce sens qu'elles ne sont pas porteuses de sens. Il est trivial d'affirmer qu'il en va de même pour les unités phonétiques que sont les phones. Mais nous verrons que le problème se pose autrement pour la prosodie.

On peut concevoir l'exécution du plan phonologique que réalisent les dispositifs phonétiques, comme la transmutation de catégories discrètes abstraites (domaine de la phonologie) en variables quantitatives interprétables à la fois en termes de catégories (moins abstraites que les précédentes) et de détails phoniques (domaine de la phonétique), ces derniers pouvant différer d'une langue à l'autre, ce qui dénote qu'ils ont fait l'objet d'un apprentissage. D'après Pierrehumbert (2001), il existe donc également une connaissance implicite des variables quantitatives, connaissance qui, à ce titre, fait également partie de la compétence linguistique des locuteurs. Cette compétence s'avère particulièrement productive, dans la mesure où elle développe une aptitude précoce à prêter attention au signal de parole. La modélisation de cette connaissance nécessite, selon l'auteur cité, le recours à une cartographie cognitive (un analogue de la réalité physique) et un jeu de symboles. À titre d'exemple, une partie de cette cartographie cognitive, qui représente le niveau le plus

superficiel d'encodage de la structure sonore, est reflétée dans l'espace formantique des voyelles. Quant aux symboles, il importe qu'ils soient différents de ceux qui sont utilisés pour représenter les catégories abstraites de la grammaire phonologique.

Enfin, si l'on s'inscrit dans une démarche de recherche empirique (de type ascendante, par exemple), il s'avère difficile de décider de prime abord si un phénomène de prononciation observé doit être classé comme un fait phonologique ou un fait phonétique. Cela dépend en partie de la conception que l'on se fait de la phonologie et de la profondeur de son niveau d'abstraction. Mais, en règle générale, la difficulté majeure tient à l'opacité des observables qui résulte en partie du fait que les catégories phonétiques incluent des informations redondantes et d'autre part au fait qu'elles ne sont pas contextuellement « égalables ». Le point positif est que chaque langue tend à exploiter de façon préférentielle l'espace acoustique défini par le jeu des paramètres acoustiques qui sont par ailleurs soumis à des contraintes universelles (*cf.* par exemple les contraintes exercées sur la fonction de transfert du conduit vocal), ce qui limite positivement la variabilité. L'un des aspects centraux de la démarche méthodologique qui vise à dissocier les faits phonétiques des faits phonologiques concerne l'identification respective de la variation contextuelle contrainte linguistiquement (contraintes phonotactiques) et de la variation contextuelle engendrée par d'autres facteurs. Un autre aspect important, plus général et peut-être plus complexe pour la phonologie dite « *suprasegmentale* » que pour la phonologie segmentale (*cf. infra*), se rapporte à l'interprétation de l'interface forme/fonction.

1.2.2. Les relations de la phonétique à la phonologie

Avant la rupture épistémologique entre la phonétique et la phonologie officialisée par Troubetzkoy (1933), la phonologie était implicitement (Ohala, 1991) une composante de la phonétique classique et l'on était davantage concerné par la phonologie lorsqu'on s'intéressait plus précisément à la fonctionnalité des sons dans les langues (un point de vue qui a perduré comme en témoigne le titre de l'un des ouvrages de Martinet (1949) : *Phonology as a functional phonetics).* Après Troubetzkoy, les relations entre la phonétique et la phonologie ont été pour le moins distendues, voire rompues, en particulier avec l'affirmation de prises de position extrêmes revendiquées par une phonologie autonomiste qui désirait s'affranchir de la matérialité des observables. Avec le recul, il apparaît que la rupture épistémologique, qui consistait notamment à écarter la phonétique de la linguistique en la considérant comme une science naturelle (une idée vivace, formulée aujourd'hui encore dans *L'Homme de Paroles* de Claude Hagège, 1985 : 101, 130-131), est fondée sur un double malentendu. Le premier concerne une conception erronée de la substance phonique (domaine de recherche

traditionnellement concédé au phonéticien) et le second, la prétendue vacuité théorique de la phonétique (qui fait partie de l'implicite de certains phonologues).

La conception de la substance amorphe, héritée de Saussure a été contestée par de nombreux chercheurs qui ont fait valoir que la substance phonique de la parole se distingue de tout autre type de matière sonore car elle concrétise la projection des codes linguistiques d'une langue, et notamment celle de son code phonologique. En ce sens, la substance phonique de la parole n'est pas amorphe, mais « *préformée* » ou « *pré-systématique* ». Comme l'a bien souligné Copceag (1965 : 51) la substance peut être définie comme « *la totalité des éléments (ou phénomènes) articulatoires et acoustiques physiquement possibles, sélectionnés, groupés et arrangés de diverses manières dans les diverses langues.* » (c'est nous qui soulignons). Si l'on adopte ce point de vue, il apparaît que le rôle de la phonétique (notamment de la phonétique descriptive, nantie de ses outils expérimentaux et de son « appareil notationnel ») ne se borne pas à enregistrer des représentations analogues de la réalité physique. Il consiste également, comme l'illustre l'histoire de la discipline, à recenser les régularités superficielles qui émanent à la fois de l'incidence des contraintes universelles inhérentes à l'exercice de la parole et de la projection des représentations symboliques du système structurant qui sous-tend ces dernières. En d'autres termes, les recherches phonétiques contribuent à mettre en lumière des régularités du langage qui pourraient échapper, en tout bien tout honneur, à la connaissance introspective du phonologue.

Si l'on convient de cette mission de la phonétique, il devient difficile de nier le caractère incontournable de sa relation à la phonologie. Cependant, cette relation a parfois été interprétée d'une façon telle que la phonétique s'est vue trop souvent dotée d'une fonction mineure ou subsidiaire, comme c'est notamment le cas dans le modèle standard de la phonologie générative (Chomsky & Halle, 1968). En effet, bien que les auteurs du *Sound Pattern* accordent à la phonétique - notamment par le biais de la théorie du marquage (*markedness*) - une place qu'elle n'avait pas auparavant, il n'en demeure pas moins qu'ils la regardent comme une composante extra-grammaticale, universelle et automatique, ce qui limite singulièrement sa portée linguistique. Il semble que cette conception, qui se trouve être réaffirmée dans le Programme Minimaliste (Chomsky, 1995 ; Chomsky & Lasnik, 1995), à travers le concept de « *niveau de représentation de la Forme Phonétique* » (*Phonetic-Form*), ne soit pas viable en l'état (Keating, 1985 ; Pierrehumbert, 2002). En effet, les processus phonétiques, même les plus largement attestés, ne sont pas inexorablement universels (*cf. supra* 1.2.1). Les bases de données qui ont pu se constituer grâce aux technologies modernes témoignent de l'existence de patrons et de principes d'implémentation phonétiques spécifiques à la langue. Les processus phonétiques ne peuvent, de ce fait, être réduits à des automatismes générés par une « *machine physique parlante* ».

La dichotomie simpliste proposée par Cutler & Ladd (1983) entre les « mesureurs » et les « faiseurs de modèles » tend à conforter l'idée obsolète que la phonétique, perçue essentiellement dans sa dimension expérimentale, est une discipline auxiliaire dont la fonction principale serait d'illustrer les représentations structurelles découvertes par les phonologues qui, en définitive, seraient les seules personnes capables d'expliquer le fonctionnement du langage. À l'encontre de cette idée, il serait facile, et probablement utile, de démontrer que le champ couvert par les travaux des phonéticiens n'est pas théoriquement vide. Faute de place, nous nous limiterons à citer quelques-unes des théories marquantes élaborées par des phonéticiens (parfois associés à des psycholinguistes), telles que la théorie quantique de Stevens (1989), la théorie motrice de Liberman *et al.* (1963), Liberman & Mattingly (1985), la théorie de la dispersion vocalique de Liljencrants & Lindblom (1972), la théorie Hyper & Hypo (H & H) de Lindblom (1990), etc. Ces théories, et bien d'autres encore, ont le principal avantage de mettre en lumière la valeur explicative de la phonétique au regard de la phonologie et attestent ainsi de sa place significative dans l'étude du langage. En outre, certaines de ces théories, notamment celles qui se rapportent à la modélisation des aspects non-linéaires de l'articulation et de l'acoustique de la parole, comme c'est le cas de la théorie de Stevens par exemple, pour ce qui concerne l'acoustique, ou de la théorie de Browman & Goldstein (1990), pour les aspects articulatoires, montrent que les données phonétiques sont déjà *quasi-catégorielles* (Pierrehumbert *et al.*, 2000) et donc à « mi-chemin » entre les faits continus de la matérialité physique et les objets phonologiques proprement dits.

De nos jours, l'évolution des mentalités tend à favoriser un rapprochement de la phonétique et de la phonologie qui paraît répondre aux vœux des représentants des deux disciplines, mais qui est cependant loin d'être définitivement acquis, comme le soulignent ces remarques formulées récemment par J. Goldsmith :

> *« Dans l'état actuel des choses , il y a peu de communication entre la phonologie et les recherches sur la parole. Des deux côtés de la frontière qui sépare ces champs de recherche, nous étudions les mêmes phénomènes, on dirait, à savoir : les propriétés objectives des sons linguistiques. Il est à mon avis essentiel d'éliminer la barrière qui sépare l'étude de la phonologie et de la parole. »* (Goldsmith, 1995b : 197).

Une des manifestations visibles du rapprochement qui vient d'être évoqué est représentée par la mouvance de la « phonologie-ancrée-substance » qui propose d'évaluer l'incidence des interactions sensori-motrices sur la mise en forme des systèmes phonologiques ». (Il est intéressant de constater que cette orientation nouvelle, qui constitue une sorte d'antidote aux excès de la

phonologie autonomiste, paraît s'imposer de nos jours comme une absolue nécessité aux yeux de certains phonologues :

> *« Ancrer la réalité des systèmes phonologiques dans la substance apparaît maintenant comme un moyen indispensable pour penser la phonologie comme lieu d'interface entre objets physiques et entités phonologiques. »* (Laks & Plénat, 1993 : 7).

Une autre manifestation en faveur du rapprochement des deux disciplines tend à gagner progressivement un statut institutionnel. Il s'agit du développement de la « Phonologie de Laboratoire » auquel les « *prosodistes* » ont apporté une contribution déterminante (Kingston & Beckman, 1990 ; Pierrehumbert *et al.*, 2000).

Le rapprochement de la phonétique et de la phonologie a pour effet de mettre au premier plan la question cruciale de leur articulation. Dans cette optique, plusieurs linguistes se disent favorables à une fusion des deux disciplines et à la promotion d'une « *conception intégrative* » (Ohala, 1991), ce qui est aussi une manière de contourner la problématique de l'existence d'une interface entre la phonétique et la phonologie (Ohala, 1990), interface sur la nature de laquelle de nombreux chercheurs continuent cependant de s'interroger. Nous estimons pour notre part que le nœud de la problématique des relations entre la phonétique et la phonologie se rapporte davantage à une définition explicite des *niveaux d'analyse du matériau sonore* qu'au questionnement des différentes modalités de communication plausibles entre des disciplines séparées. Ce problème avait été perçu par B. Malmberg, il y a près d'un demi-siècle, lorsqu'il écrivait notamment :

> *« ... la différence entre une description de la forme - ou description structurale - et une description de la substance n'est pas une simple opposition entre deux procédés descriptifs, et, bien entendu, entre deux aspects de la même réalité linguistique. Nous sommes au contraire en face de toute une série de niveaux ou de degrés d'abstraction, choisis arbitrairement par le chercheur dans le but que celui-ci s'est proposé. »* (Malmberg, 1966 : 100).

Il est pour le moins significatif que l'on retrouve une formulation similaire dans des propos tenus par J. Pierrehumbert en 2002 :

> *« Dans les théories viables des relations entre phonétique et phonologie, il y a une échelle d'abstraction dans laquelle chaque niveau possède son propre appareil représentationnel. La théorie dans son ensemble doit donc spécifier à la fois la représentation disponible à chacun des niveaux et les principes qui relient un niveau à l'autre. »* (Pierrehumbert, 2002 : 2).

1.2.3. Quelques retombées sur la prosodie

Nous avons présenté, dans les sections précédentes, un ensemble de remarques générales concernant la nature de la phonétique et de la phonologie, ainsi que la problématique de leurs

relations. Cette brève mise au point a été conçue dans le but de dégager quelques concepts fondamentaux et d'évaluer par la suite (c'est-à-dire dès à présent) leurs principales retombées sur l'analyse prosodique. Plus généralement, nous pensons que ces remarques liminaires seront utiles à la compréhension de cette étude consacrée essentiellement à la quiddité phonétique et phonologique de la prosodie.

Si la question du statut linguistique de la prosodie ne se pose plus, il subsiste en revanche de nombreuses interrogations concernant le transfert des principes et des méthodes de la phonologie segmentale à la prosodie, dans la mesure où la première est souvent perçue comme un parangon théorique et méthodologique. Ces interrogations s'avèrent d'autant plus insistantes que les éléments prosodiques ne sont plus considérés comme des phénomènes linguistiques secondaires (comme c'est le cas dans l'approche du structuralisme américain, telle qu'elle est représentée par Bloomfield et Trager, par exemple, ou dans l'approche fonctionnaliste de Martinet), mais comme des entités représentatives d'une *structure phonologique autonome* (*cf. infra*). La problématique des retombées de la phonologie sur le traitement de la prosodie s'inscrit dans trois grands thèmes qui concernent respectivement les propriétés de la (des) grammaire(s) phonologique(s) de la prosodie, la place de la phonétique vis-à-vis de cette (ces) grammaires(s) et les relations de la composante prosodique avec les autres composantes de la grammaire générale, comme la syntaxe et la sémantique, ainsi qu'avec cette composante externe à la grammaire que constitue la pragmatique dans le modèle génératif classique. Il est à noter que le modèle de la Grammaire Générative n'intègre pas une composante (ou un module) prosodique propre. Il est possible, cependant, d'admettre l'existence d'une telle composante, ou de considérer qu'il s'agit d'une sous-composante du module phonologique pouvant être interfacé avec les autres composantes de la Grammaire, comme la syntaxe et la sémantique.

Les aspects grammaticaux concernent au premier chef *l'analyse formelle* de la prosodie, dans les trois perspectives de la Grammaire Universelle, des grammaires spécifiques et de la grammaire probabiliste. Pour chacune de ces orientations, le problème des relations entre les représentations phonologiques et les représentations phonétiques (quels que soient par ailleurs les niveaux d'abstraction postulés, *cf. infra*) est difficilement contournable et peut être envisagé sous différents angles.

La Grammaire Universelle de la compétence linguistique comporte un ensemble de *principes* et de *paramètres* qui sous-tendent, respectivement, le fonctionnement général de la composante phonologique et la variabilité phonologique inter-langues. Dans le cours d'une évolution récente, ces primitives théoriques ont été réinterprétées à la lumière d'une nouvelle approche : la Théorie de l'Optimalité : OT (Prince & Smolensky, 1993) qui postule un ensemble de contraintes

universelles, interactives (mutuellement transgressantes) et hiérarchiquement ordonnées de façon différente selon la langue (ou les familles de langues).

La perspective de la Grammaire Universelle classique conduit à s'interroger sur l'identité des principes et des paramètres qui régissent la part prosodique de la composante phonologique. La difficulté principale provient ici du fait que cette partie prosodique est constituée, comme on le verra par la suite (section 2), de *plusieurs systèmes* (notamment, pour simplifier, d'un système accentuel et d'un système intonatif). Il importe, en conséquence, de rendre explicite la nature des principes généraux et des paramètres propres à chaque système (*cf. infra*). Il en va de même en ce qui concerne la spécification des *contraintes prosodiques universelles* et celle de la variation inter-langue de leur agencement hiérarchique, dans le cadre de l'approche OT. Mais cette dernière n'est pas encore appliquée de façon systématique à l'analyse formelle de la prosodie. Cependant, on pourra se reporter à Post (2000) et à Selkirk (2000) pour des applications récentes de OT à l'interprétation de certains faits de phrasage prosodique en français et en anglais. (Nous utilisons dans cet article le terme de *phrasage* au lieu de *phraséologie*, comme un équivalent de *phrasing,* pour désigner le dispositif de fragmentation d'un texte oral en *unités* ou en *domaines* de rang inférieur).

Par grammaire phonologique spécifique, nous entendons le système prosodique d'une langue particulière. En règle générale, la description et la représentation de ce système en appelle à des connaissances qui font référence à son *architecture structurelle* et à des *contraintes phonotactiques* séquentielles qui régulent le déploiement temporel de cette architecture. L'architecture structurelle met en jeu un ensemble de primitives et de constructions. Etant admis qu'un système prosodique est constitué de plusieurs sous-systèmes (Crystal, 1969), il est nécessaire de s'interroger sur la nature des *primitives* et des *constructions* afférentes à chacun de ces systèmes. Quelles sont les primitives et les constructions du système métrique et du système intonatif d'une langue donnée ? Quelles relations ces entités structurelles entretiennent-elles dans l'architecture générale du système prosodique ? Quels sont les principes qui gouvernent la bonne formation de ces constructions ?

Si on laisse de côté l'option de la stricte *autonomie*, il apparaît que l'étude de la grammaire prosodique d'une langue doit également s'interroger sur la nature des relations entre les entités structurelles abstraites qui constituent le noyau de cette grammaire et leurs corrélats phoniques concrets qui relèvent en principe de l'analyse phonétique. Dans l'optique integrative de la phonétique et de la phonologie (*cf. supra*) il est logique que l'étude de ces relations soit une partie intégrante de la grammaire même. Dans une autre perspective, plus traditionnelle, il s'agit d'analyser la texture des liens établis entre des entités internes à la grammaire et des objets physiques externes à celle-ci (ce qui revient à dissocier nettement phonétique et phonologie). Enfin, il reste la solution intermédiaire de l'interfaçage (ou du chevauchement) entre la phonétique

et la phonologie, solution qui représente un type d'association différent des deux précédents. Dans tous les cas se pose en vérité un seul et même problème crucial: celui de la définition explicite des différents niveaux d'abstraction de l'analyse prosodique et de l'étagement de ces niveaux au regard du corps de la grammaire et des représentations cognitives. Il est prématuré de traiter dès à présent cette question sur laquelle nous reviendrons ultérieurement, lorsque les principaux concepts qui se rapportent à l'analyse prosodique auront été clarifiés.

La conception de la *grammaire probabiliste* exposée récemment par Pierrehumbert (2001, 2002), que nous avons évoquée succinctement dans une section précédente, est particulièrement attrayante, car elle défend l'idée que la phonologie implique à la fois une « *connaissance formelle* » des entités catégorielles et une « *connaissance statistique* ». En se référant à des données empiriques, cette approche soutient également que les représentations cognitives des structures sonores sont probabilistes, cette remarque concernant à la fois la phonétique et la phonologie, dans la mesure où des effets probabilistes ont été établis aux divers niveaux d'abstraction qui se rapportent aux deux champs d'étude. La décision d'introduire dans un modèle des aspects probabilistes, sans renoncer pour autant à la dimension catégorielle classique, ouvre la voie à l'élaboration d'outils particulièrement bien adaptés à la maîtrise de la variabilité des données et des sorties graduelles qu'elles engendrent. Ces outils peuvent s'avérer très utiles à l'analyse prosodique qui se trouve être davantage confrontée à un foisonnement de variables pour lesquelles il est difficile de décider au départ si elles relèvent de la compétence linguistique ou des avatars de la performance.

L'usage d'une grammaire prosodique stochastique, notamment d'une grammaire prosodique *à états-finis*, a donné de bons résultats dans des applications dédiées aux technologies de la parole et en particulier à la synthèse vocale. Il reste cependant à élargir les bases de connaissances sur les détails d'implémentation phonétique propres à une langue particulière. Il est certain que le développement de ces connaissances dans le domaine de la phonétique prosodique serait très profitable à la synthèse de la parole. Il permettrait également de former une image plus précise de la façon dont le système cognitif exploite dans chaque langue les ressources acoustiques universelles (une idée qui va à l'encontre de celle qui stipule des catégories phonétiques universelles, selon Pierrehumbert, 2000) et d'ancrer ainsi plus fermement les relations de la phonétique et de la phonologie dans l'édification d'un modèle cognitif de la phonie linguistique.

Pour des raisons pratiques et afin d'éviter des redites inutiles, la problématique des niveaux d'analyse et de représentation de l'espace phonético-phonologique, ainsi que celle des relations de la grammaire phonologique avec les autres composantes grammaticales, comme la syntaxe et la sémantique, n'ont pas été développées dans cette sous-section à caractère très général. Elles seront reprises à la place qui nous a semblé leur convenir au mieux dans les autres parties de cet article.

2. De la quiddité de la prosodie

Le concept de prosodie n'est pas univoque et se prête à plusieurs interprétations comme l'atteste la diversité des définitions qui parsèment la littérature. D'autre part, l'analyse prosodique dispose d'une terminologie prolifique qui s'avère souvent confondante par le fait que les termes utilisés ne font pas toujours référence de façon explicite au cadre théorique ou au niveau d'analyse auxquels ils se rapportent. Enfin, l'un des aspects essentiels de l'analyse prosodique concerne sa fonctionnalité et, corrélativement, la question de la relation forme/sens, qui se pose avec plus d'acuité que pour la phonologie segmentale eu égard à la plurifonctionnalité de la prosodie. L'objectif de cette section est d'apporter quelques éclaircissements sur les points qui viennent d'être soulevés.

2.1. Les systèmes prosodiques et leurs domaines d'application (prosodie lexicale et prosodie post-lexicale)

Dans le passé, le concept de prosodie n'a cessé de fluctuer (comme l'attestent les définitions données dans les dictionnaires usuels et spécialisés) entre deux pôles représentés respectivement par les termes de métrique et d'intonation et il est courant de constater que les termes de prosodie et d'intonation sont pris pour des synonymes. Selon notre point de vue, le système prosodique est constitué de plusieurs *sous-systèmes* dont le rôle consiste à contrôler un jeu de *paramètres abstraits* et de *paramètres physiques*, afin d'assumer un ensemble de *fonctions diversifiées* se rapportant à différentes unités de la chaîne linguistique (mot, morphème, syntagme, énoncé, etc.). Nous proposons (Di Cristo, 2000c) d'appréhender ces sous-systèmes comme trois ordres structuraux relativement autonomes, mais néanmoins interactifs : l'ordre de structuration métrique , l'ordre de structuration tonale et l'ordre de structuration temporelle.

Ces propositions peuvent être développées en précisant que chacun de ces ordres structurels peut s'appliquer soit à la mise en forme des unités du *lexique*, soit à celle des *unités de rang supérieur* (comme la phrase, l'énoncé et le discours). En conséquence, il est nécessaire de distinguer entre *prosodie lexicale* et *prosodie supra-lexicale* ou post-lexicale. C'est ainsi que l'ordre de structuration métrique appliqué au lexique concerne la gestion de *l'accentuation lexicale*. Appliqués à ce même domaine, l'ordre de structuration tonale et l'ordre de structuration temporelle ont en charge de gérer, respectivement, les *oppositions de tons* (dans les langues dites à tons) et les *oppositions de quantité*. Au niveau post-lexical, l'ordre de structuration tonale est responsable de la gestion de *l'intonation*, tandis que l'ordre de structuration métrique est chargé de la gestion du *rythme*. Enfin, l'ordre de structuration temporelle prend en charge, au niveau post-lexical, la gestion des *pauses linguistiques* (ou structurales), des *effets d'allongement* (comme le phénomène d'allongement final : *final lengthening*) et des changements de *tempo*

(variations de la vitesse d'élocution et du débit de parole). Il est désormais admis que l'organisation temporelle est le reflet d'une *structure complexe* (Campbell, 1992, 1993 ; Ladd, 1993) et que les effets de variation de la durée sont intimement liés à *l'organisation* des énoncés et du discours (Duez, 1987, 1997 ; White, 2002 ; Kohler, 2003). Plus généralement, l'orchestration temporelle des activités qui président à la production et à la perception de la parole constitue un *dispositif central* du processus par lequel la structure linguistique est oralement communiquée. Selon Schirmer (2004), l'information qui peut être dérivée de l'organisation temporelle supervise une large gamme de processus linguistiques qui vont de la perception des segments individuels (voyelles ou consonnes) au traitement des ambiguïtés syntaxiques. Il apparaît en définitive que l'étude systématique de la temporalité de la parole représente un vecteur de recherche particulièrement important pour la compréhension de la nature des relations qui peuvent s'établir entre la gestion des éléments prosodiques (dans leurs relations à la syntaxe, à la sémantique et à la pragmatique) et celle des éléments segmentaux (Warren, 1985 ; Byrd & Salstman, 2003).

La tripartition des systèmes prosodiques que nous venons de proposer sied particulièrement à une *approche modulaire* de la structure prosodique. De ce point de vue, il est possible que la mise en œuvre des différents systèmes prosodiques s'adresse à des dispositifs de contrôle de haut niveau distincts et distribués différemment au sein de l'architecture cérébrale. Cette hypothèse reste à vérifier, bien que des travaux en neurolinguistique semblent effectivement attester une latéralisation différenciée du traitement des paramètres de mélodie et de durée, par exemple (van Lancker & Sidtis, 1992 ; Zatorre *et al.*, 1992). Par ailleurs, il est intéressant de constater que deux des trois sous-systèmes cités précédemment ont fait l'objet de traitements phonologiques relativement indépendants dans les cadres respectifs de la *théorie métrique* et de la théorie de la *phonologie intonative* (dont nous reparlerons ultérieurement). Il n'en demeure pas moins que les trois systèmes sont interalliés à plus d'un titre. D'abord, par le fait qu'ils collaborent à l'accomplissement de certaines *fonctions grammaticales* (comme le signalement des différents types de démarcation des unités linguistiques) et *extra-grammaticales* (comme l'expression des attitudes et de l'affect en général). Ensuite, par l'instauration de relations de préséance. En effet, plusieurs chercheurs estiment que l'attribution du *patron intonatif* spécifique à un énoncé nécessite la spécification préalable de son *squelette métrique*. Cette relation est sensible dans les modèles qui se recommandent de l'approche métrique-autosegmentale (*cf. infra*, dernière partie), modèles dans lesquels les *segments tonals* qui forment la structure intonative sont ancrés à ces entités constitutives de l'organisation métrique que sont les *accents* (Pierrehumbert, 1980) et les bornes de constituants prosodiques (Hirst & Di Cristo, 1984). Enfin il apparaît clairement, comme nous le verrons par la suite, qu'il est loisible aux entités abstraites de l'organisation prosodique, comme les accents, les tons, les

domaines, de sélectionner les paramètres prosodiques (mélodie, durée, intensité) qui s'avèrent les plus appropriés à l'exercice des fonctions que ces entités assument dans une langue particulière.

2.2. Matérialité des systèmes prosodiques

La matérialité des systèmes prosodiques est de nature *pluriparamétrique*, car elle se rapporte aux variations concomitantes des trois *paramètres physiques* fondamentaux que sont la fréquence fondamentale (F0), la durée et l'intensité (il est parfois considéré que la *structure spectrale* du son qui détermine son *timbre*, doit également être prise en compte. En ce qui nous concerne, nous avons choisi de qualifier ce paramètre de *paraprosodique* : Di Cristo, 2000b). Plus précisément, la matérialité des systémes prosodiques fait référence à un ensemble de phénomènes concrets qui concernent à la fois la *production*, la *transmission acoustique* et la *perception* de la prosodie dans les diverses manifestations de la parole. La figure 1 résume les correspondances élémentaires que l'on peut établir entre ces *faits concrets* et les *entités abstraites* qui participent à la construction des structures prosodiques. Dès à présent se pose donc le problème des *niveaux d'analyse* de la prosodie et, par voie de conséquence, celui de la *cohérence terminologique* relative à ces niveaux, problèmes qui seront débattus dans les sections suivantes.

Caractéristiques abstraites		**Caractéristiques concrètes: mesurables (aspects physiques et sensoriels)**		
Lexicales	***supra-lexicales***	***Production*** (physiologie)	***Transmission*** (Acoustique)	***Perception*** (Audition)
Quantité	Tempo, Pauses	chronométrage des gestes articulatoires	durée des événements et des silences	longueur
Accent lexical	Accents supra-lexicaux. Rythme	effort articulatoire	F0, amplitude spectre, durée	mélodie, intensité timbre, longueur
Ton	Intonation	activité musculaire activité respiratoire	fréquence fondamentale (F0)	mélodie

Figure 1
Les entités abstraites des systèmes prosodiques et leur matérialité

2.3. Niveaux d'analyse

Si l'on se place dans la perspective d'une démarche ascendante (*Bottom-up*) qui procède d'une analyse des phénomènes les plus concrets vers celle des entités les plus abstraites, il est évident que les niveaux physiques de la production et de la représentation acoustique constituent les niveaux les plus concrets de l'étude de la matérialité prosodique. Les analyses expérimentales permettent en effet d'interpréter directement ces niveaux en termes de *mesures* (par ex. mesure de la pression sous-glottique et de l'activité électrique des muscles du larynx, pour ce qui concerne la *physiologie de la production* ; mesure des variations de la fréquence fondamentale (F0) et de l'intensité physique, pour *l'acoustique*, mesure de la durée physique des unités de la chaîne linguistique : phonèmes, syllabes, mots, syntagmes, etc.)

Un niveau d'abstraction est franchi lorsque le chercheur utilise son oreille ou celle de ses semblables pour décrire la prosodie. Malgré les performances élevées des sujets entraînés pour l'accomplissement de cette tâche, les résultats de l'analyse auditive restent très en deçà de ceux de l'analyse objective. Cette granularité plus grossière peut cependant être vue comme une contrainte bénéfique, dans la mesure où l'on suppose qu'elle permet d'extraire du signal physique les informations les plus pertinentes en vue de l'interprétation linguistique. C'est sur la base d'une hypothèse de ce type que se sont développées des approches (dites de « *stylisation prosodique* ») qui visent à convertir manuellement ou automatiquement des données *physiques* en données *subjectives*, à partir de la prise en considération de contraintes de perception expérimentalement validées (Di Cristo, 1978 ; Rossi *et al.*, 1981 ; Mertens 2004 ; Auran & Di Cristo, 2004). Les approches automatiques ou semi-automatiques du transcodage des données acoustiques en données perceptives présentent un avantage indéniable sur les analyses auditives traditionnelles qui sont empreintes d'une inévitable subjectivité dont on connaît depuis longtemps les avatars (Lieberman, 1965). Cependant, ces approches sont difficiles à mettre en œuvre et requièrent des méthodes de calcul sophistiquées.

Le niveau d'abstraction le plus élevé concerne évidemment l'interprétation des éléments prosodiques comme des *pures formes* et des *structures*, représentatives des catégories mentales que gère la compétence linguistique des locuteurs. On peut donc tirer des remarques précédentes la conclusion qu'il existe au moins *trois niveaux* d'analyse fondamentaux de la prosodie ('t Hart & Collier, 1975) dont les extrêmes sont représentés par le niveau acoustique et le niveau phonologique profond. Le niveau de l'analyse auditive constitue un niveau intermédiaire, une sorte d'interface qui permet de relier la substance physique de la prosodie aux représentations abstraites de la phonologie prosodique. Ces trois niveaux sont-ils suffisants pour instruire l'analyse

prosodique et comment sont-ils reliés les uns aux autres ? Nous tenterons d'apporter des éléments de réponse à ces questions dans la dernière partie de l'article, qui est consacrée à la présentation des modèles et des théories.

2.4. Situation de la prosodie dans l'étude de la langue

Selon le schéma développé par Hjelmslev (1953), la langue s'articule selon deux plans : le *plan de l'expression* et le *plan du contenu*, qui se rapportent, respectivement, aux propriétés phoniques des langues et aux significations qu'elles véhiculent. Chacun de ces plans se subdivise à son tour en deux parties dénommées *forme* et *substance*. La tradition considère que l'étude de la forme sonore des langues constitue le domaine de la phonologie et celle de la substance, celui de la phonétique. De même, cette tradition perpétue l'idée que l'étude de la forme du contenu est dévolue à la syntaxe et à la sémantique, tandis que celle de la substance du contenu revient à la pragmatique.

Il est généralement admis de nos jours que le système phonologique d'une langue est constituée de deux ensembles (figure 2) que circonscrivent, respectivement, le champ de la *phonologie segmentale* (ou phonématique) et celui de la *phonologie suprasegmentale* (ou prosodique). Néanmoins, cette séparation n'est pas entièrement étanche comme le montrent des travaux récents dont l'intérêt se porte sur l'interface phonologie/phonétique et qui, de ce point de vue, mettent en évidence l'influence de la structure prosodique sur le niveau segmental (Fant *et al.,* 1987 ; Beckman & Edwards, 1992 ; Fougeron, 1998 ; 1999 ; Keating *et al.*, 2003 ; Meynadier, 2003 ; Chao, 2005). Les deux ensembles sont cependant analysés le plus souvent comme des systèmes indépendants.

L'analyse *formelle* de la prosodie (que nous qualifions de *prosodématique*) s'attache à la description et à la représentation des primitives et des constructions abstraites (ou mentales) des systèmes prosodiques, tandis que l'analyse de la substance s'applique à décrire les corrélats acoustiques et auditifs de ces entités abstraites, en termes de variations paramétriques concrètes.

Nous avons avancé l'idée (1.2.2.) que *l'analyse phonétique* ne se réduit pas à enregistrer mécaniquement des faits de substance, mais qu'elle procède à une *interprétation* de ces derniers en y relevant (par la description, la schématisation et la transcription) des régularités superficielles qui actualisent la projection des codes phonologiques sous-jacents. C'est la raison pour laquelle nous utilisons l'expression de « *phonétique interprétative* » pour désigner cette branche de la phonétique distincte de la *phonétique paramétrique* ou quantitative, laquelle s'attache pour sa part aux mesures et aux représentations analytiques des paramètres. Dans cet esprit, nous proposons de regarder la *phonétique interprétative* comme une interface (*cf.* Hirst & Di Cristo, 1998) entre les représentations phonologiques et la substance phonique. Notre conception diffère ainsi de celle de Laks & Plénat (1993) qui envisagent la phonologie comme lieu d'interface entre les objets physiques et les entités

phonologiques (*cf. supra* : 1.2.2.). Dans le cadre théorique que nous proposons, la discipline prosodique (ou *prosodologie*) comporte deux branches : la *prosodématique* et la *phonoprosodie*. Cette dernière intègre alors la *phonétique paramétrique* (ou phonétique physique) et la *phonétique interprétative*, qui opère une première interprétation de l'évolution des paramètres physiques, sans préjuger toutefois de leur valeur phonologique dans le système de la langue. Selon cette conception, la phonologie interprète les données phonétiques quasi-catégorielles, alors que la phonétique interprète pour sa part les événements continus du monde physique. Conçue comme une interface entre la matière et la forme, la phonétique interprétative est à la phonologie ce que la pragmatique est à la sémantique (Hirst & Di Cristo, 1998). Cette conception nous amène à proposer une représentation interactive des relations de la Grammaire aux composantes extra-grammaticales qui est illustrée par la vignette incluse dans la figure 2. Pour une discussion récente et élargie concernant *l'interprétation phonétique*, on se reportera à l'ouvrage collectif édité par Local *et al.* (2003), notamment au chapitre de Beckman & Pierrehumbert, dans lequel la phonétique est alors définie comme la « *sémantique de la phonologie* » et la phonologie, comme la « *syntaxe de la phonétique* ».

En ce qui concerne l'articulation du plan de l'expression et du plan du contenu (en relation de présupposition réciproque) qui relève, pour le sujet traité ici, d'une *sémioprosodie*, il apparaît que les éléments prosodiques assument, à la différence des éléments phonématiques, de *multiples fonctions* qui concernent à la fois la morphologie, la syntaxe, la sémantique et la pragmatique. Une telle pluralité fonctionnelle est rendue possible par le fait que les contrastes prosodiques véhiculent plusieurs types de significations (*linguistiques* et *paralinguistiques*). Cette observation a conduit les linguistes à souligner que, à la différence des phonèmes, les éléments prosodiques possédent à la fois une forme (ou un signifiant) et un contenu (ou un signifié). C'est la raison pour laquelle les unités prosodiques sont conçues, dans le cadre du modèles hjemslévien, comme des unités *biplanes* ou *hétéroplanes*, par opposition aux unités *monoplanes* que sont les phonèmes (Greimas & Courtès, 1979). Toutefois, une telle conception n'est pas généralisable et doit être reconsidérée à la lumière d'une définition explicite de la notion *d'unité prosodique* (*cf.* ci-après la discussion sur les primitives et les constructions des systèmes prosodiques).

2.5. Problèmes terminologiques

La question des niveaux d'analyse et de représentation de la prosodie est inséparable de celle de la terminologie, dans la mesure où les termes utilisés ne doivent pas transgresser ces niveaux sous peine d'être à l'origine de confusions regrettables. Nous présentons quelques exemples pour illustrer ces propos.

En règle générale, la cohérence voudrait que s'applique à la prosodie la distinction *–étique/-émique* consacrée par l'étude phonétique et phonologique (phonémique) des éléments segmentaux. Le terme de prosodème devrait ainsi s'imposer pour la désignation de toute unité prosodique du système phonologique. Il conviendrait de même d'utiliser systématiquement les termes de *tonème*, pour faire référence aux unités phonologiques des systèmes tonals, ceux *d'intonème* et de *contourème*, pour identifier des unités du système intonatif, et ceux *d'accentème* et de *chronème* pour nommer les unités du système accentuel et celles du système de l'organisation temporelle, respectivement. Enfin, et eu égard à l'étendue du champ couvert par la prosodologie, les termes tonologie, intonologie, accentologie et chronématologie seraient appropriés pour désigner les diverses approches sous-disciplinaires qui se rapportent à son étude. Des propositions dans ce sens ont été faites dans le passé. Toutefois, la plupart des termes que l'on vient de mentionner sont devenus obsolètes.

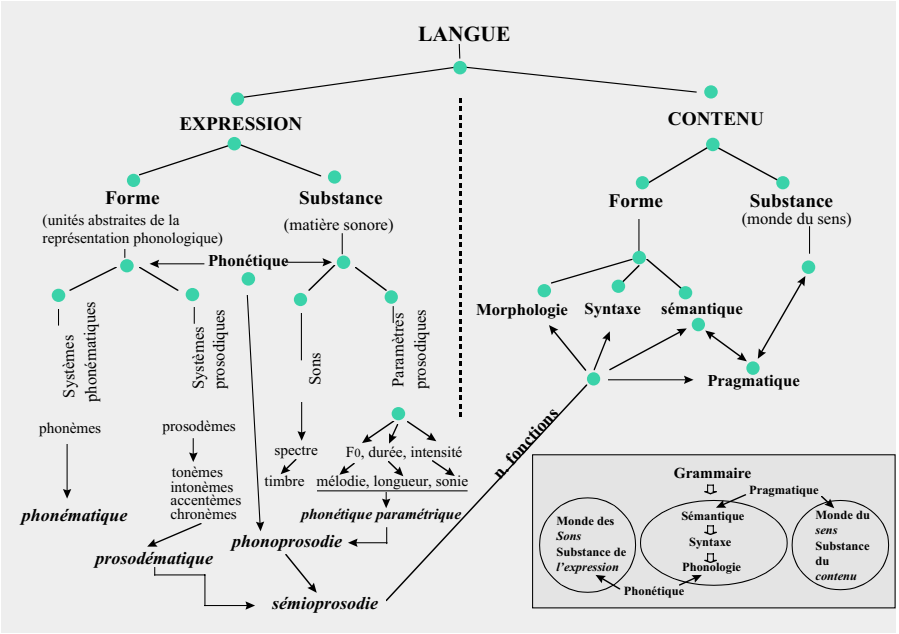

Figure 2

Situation de la prosodie dans l'étude de la langue (une extension personnelle du modèle classique proposé par Hjemslev). La vignette illustre notre conception relative à la nature des liens entre les composantes de la Grammaire (selon la conception classique du modèle génératif) et les phénomènes extra-grammaticaux. Selon cette conception, la phonétique est à la phonologie ce que la pragmatique est à la sémantique (Hirst & Di Cristo, 1998).

Contrairement à l'usage qui en est fait, les termes de *fréquence fondamentale* (F0), de *mélodie* et *d'intonation* ne doivent pas être considérés comme des équivalents sémantiques. En effet le premier fait référence à une substance physique, le second à une substance auditive et le dernier à une catégorie formelle. De ce point de vue, le terme de « courbe mélodique » n'est pas approprié pour décrire une variation de (F0), pas plus que ne l'est celui de « courbe d'intonation » pour faire référence à la variation d'un paramètre physique ou à la perception de cette variation. En revanche, des expressions comme « schéma intonatif » ou « contour intonatif » sont utilisés à bon escient s'ils servent à désigner des entités abstraites de la représentation phonologique.

Un autre problème de dénotation terminologique, auquel il est courant de se heurter dans l'analyse prosodique, concerne les termes de *proéminence* et *d'accent*. Le premier est ambigu car il fait à la fois référence à une entité abstraite de l'organisation métrique et à la manifestation concrète de cette entité. Seule une notification explicite du niveau d'analyse permet dans ce cas de résoudre l'ambiguïté. Il en va de même du terme accent qui peut désigner soit une *catégorie phonologique* (comme l'accent lexical, par exemple), soit une *catégorie phonétique*, auquel cas il est souvent assorti d'une précision sur sa nature concrète : « accent de hauteur », « de durée » « d'intensité », etc.

À la différence du français, l'anglais utilise deux termes pour marquer ce type de différence : le terme « *stress* », qui désigne l'accent lexical ou l'accent sous-jacent et le terme « *accent* », qui renvoie à la réalisation d'une proéminence accentuelle. Etant donné l'importance du paramètre mélodique dans la manifestation de l'accentuation de l'anglais, l'expression « *pitch accent* » (*accent mélodique*) tend à supplanter systématiquement le mot « accent » dans cette langue. Le terme « *stress* » sert alors à désigner soit l'accent lexical, soit un accent dont la manifestation ne s'accompagne pas d'une variation significative de la mélodie (comme l'accent d'intensité, par exemple). Beckman (1986) utilise la distinction stress/accent à des fins de classification typologique. Elle précise à cet effet que l'expression « *stress-accent* » sert à désigner une proéminence accentuelle non marquée principalement par la mélodie, comme c'est le cas pour le « *non-stress accent* » ou le « *pitch accent* ». Le terme de « *désaccentuation* » (*deaccenting*) fait référence soit à l'absence de l'accent mélodique (qui n'exclut pas pour autant la présence d'un accent métrique non mélodique), soit à la réduction du niveau de proéminence au sein d'une structure prosodique. Sans la connaissance préalable de ces subtiles distinctions terminologiques, il peut s'avérer difficile de comprendre le titre d'un ouvrage de Beckman (1986) : « *Stress and non-stress accent* », ou la signification d'une expression comme : « *syllables are desaccented but not destressed* » (*cf.* Di Cristo & Jankowski, 1999).

L'expression « *accent nucléaire* » (*nuclear stress, accent*) fait habituellement référence à l'accent qui est investi du niveau de proéminence le plus élevé dans un *domaine donné* (Cruttenden, 1997). C'est ainsi qu'il permet de désigner soit l'accent principal (« *primary stress* » : *accent primaire*) de l'Unité

Intonative auquel est associé un ton particulier appelé *ton nucléaire* (*nucleus*, *nuclear tone*), soit un *accent de phrase* (*sentence stress*) dont la position est principalement motivée par la structure informationnelle et le marquage du focus. La notion d'accent nucléaire se situe, du point de vue conceptuel, à l'interface de l'accentuation et de l'intonation, ce qui soulève des difficultés d'interprétation qui ne sont pas encore entièrement résolues.

Les niveaux de proéminence retenus par les linguistes de la mouvance américaine pour être intégrés dans les représentations linguistiques ressortissent à trois catégories :

- « *accented* ». Le terme peut désigner notamment : *l'accent nucléaire* « nuclear stress », *l'accent mélodique* : « pitch accent » de Bolinger, *l'accent primaire* : « primary stress » de Trager & Smith, ou « *l'accent1* » : « one-stress » de Chomsky & Halle. Dans la proposition très influente en son temps de Vanderslice & Ladefoged (1971), la catégorie « accented » est définie par les traits : [+Intonation, + Accent, + Heavy] (soit : + Intonation, + Accent, + Lourd).

- « *stressed* ». Le terme fait référence à la syllabe *accentuée non-nucléaire*, à *l'accent morphologique* : « morphological stress » de Bolinger, à *l'accent secondaire* : « secondary stress » de Trager & Smith, ou à *l'accent 2 ou à l'accent 3* de Chomsky & Halle : « two- stress, three-stress ». Pour Vanderslice & Ladefoged, la catégorie « stressed » est définie par les traits : [- Intonation, + Accent, + Heavy].

- « *unstressed* ». Le terme désigne le niveau d'accentuation le plus faible (sans réduction cependant du timbre vocalique « vowel reduction ») pour la majorité des auteurs ; ce qui correspond à *l'accent tertiaire* : « tertiary stress » de Trager & Smith, ainsi qu'à la matrice de traits [-Intonation, - Accent, + Heavy] chez Vanderslice & Ladefoged.

Nous prendrons pour dernier exemple la distinction nécessaire entre intensité *objective* (mesurable en *décibels*) et intensité *subjective* (mesurable en *sones*). L'anglais dispose à cet effet des deux termes : *« intensity* » et « *loudness* », que nous proposons de traduire respectivement par « *intensité physique* » et par « *sonie* », pour éviter une confusion similaire à celle qui consiste à utiliser de façon interchangeable les termes de fréquence fondamentale (F0) et de mélodie.

2.6. Eléments de définition

Il est difficile, voire aventureux, de débattre de prosodologie sans s'appuyer sur quelques éléments de définition permettant de fixer les références conceptuelles fondamentales qui sont attachées à ce domaine complexe. Cette démarche s'impose d'autant plus que les définitions proposées dans la littérature et les dictionnaires spécialisés s'avèrent parfois opaques ou conflictuelles.

Prise dans sa *signification générique* moderne, le terme de prosodie désigne une branche de la linguistique consacrée à la description factuelle (aspects phonétiques) et à l'analyse formelle (aspects phonologiques) des éléments systématiques de l'expression phonique non-coextensifs aux phonèmes,

tels que les accents, les tons, l'intonation et la quantité, dont les manifestations concrètes, dans la production de la parole, sont associées aux variations des paramètres physiques de la fréquence fondamentale (F0), de la durée et de l'intensité qui représentent les paramètres objectifs de la prosodie. Ces variations sont perçues par l'auditeur comme des changements de hauteur (ou de mélodie), de longueur et de sonie, qui constituent les paramètres subjectifs de la prosodie. Les éléments prosodiques exercent au niveau du mot (prosodie lexicale) et au-delà de ce dernier (prosodie supra-lexicale ou post-lexicale) un faisceau de fonctions grammaticales, para-grammaticales et extra-grammaticales, qui se rapportent à « ce qui est dit », à « la façon dont cela est dit », ainsi qu'à « l'identité du sujet parlant, et qui s'avèrent déterminantes pour signaler la structure des énoncés et du discours et pour procéder à leur interprétation sémantique et pragmatique.

Le ton est une entité primitive de l'organisation mélodique du niveau lexical dans les langues dites « *à tons* », c'est-à-dire dans les langues où les changements de la fréquence fondamentale (et, partant, pour l'auditeur, de la mélodie) assument une fonction distinctive pour des unités plus petites que le mot ou équivalentes à ce dernier. Dans la mouvance de la phonologie autosegmentale (*cf. infra*), le ton est également considéré comme une primitive (un *autosegment)* de l'organisation phonologique de l'intonation. Il en découle que les langues à tons et les langues à intonation peuvent être décrites au moyen d'un même formalisme qui spécifie l'alignement des segments tonals (notamment des tons Haut et Bas) avec le matériau lexical et supralexical.

Il est possible, comme l'atteste la littérature, de définir le terme *intonation* de deux façons différentes, selon que l'on se réfère ou non à son statut phonologique. Dans le premier cas, il convient de réserver ce terme (*cf. supra*) pour faire référence au système des unités mélodiques contrastives du niveau supra-lexical (quel que soit par ailleurs le statut que la théorie attribue à ces unités distinctives : tons ; contours, mouvements, configurations, etc.). Dans le second, cas le terme intonation est pris dans le sens plus général de prosodie : « *L'intonation, telle que nous la concevons, fait référence aux tons, à l'accent, aussi bien qu'à l'intonation de phrase. Elle recouvre tous les phénomènes linguistiques perçus comme une modulation au sens large du continuum de parole.* » (Rossi, 1977 : 56). Cette conception « intégrative » de l'intonation (qui s'avère être la plus communément partagée), relève davantage d'une interprétation phonétique que d'une interprétation phonologique, dans la mesure où elle évoque plus un continuum phonique qu'une forme discontinue, constituée d'unités discrètes. Conformément à notre conception de la phonétique interprétative comme une interface entre la substance physique et la phonologie (*cf. supra*), le terme intonation, pris dans cette seconde signification, est interprété comme une construction phonique qui relie les primitives prosodiques du niveau lexical et du niveau supra-lexical (quel que soit le

choix de leur représentation formelle) aux paramètres prosodiques acoustiques (Hirst & Di Cristo, 1998 : 12).

Le mot accent désigne un phénomène de proéminence sous-jacent ou avéré, que la phonétique interprète comme le renforcement d'une syllabe (actualisé par le jeu d'un ensemble de paramètres acoustiques , tels que la F0, la durée, l'intensité et le spectre) ayant pour effet de la détacher de son contexte phonique. Du point de vue phonologique, l'accent est conçu comme une entité primitive de la structure métrique (ou rythmique). On peut considérer qu'il existe trois catégories fonctionnelles fondamentales d'accentuation : lexicale, métrique et de focalisation. Le rythme linguistique est interprété comme la récurrence de structures ou de groupements constitués sur la base d'une organisation hiérarchique des niveaux de proéminence.

2.7. De la polysémie du terme suprasegmental

Les travaux sur la prosodie font un usage très fréquent du terme *suprasegmental* pour désigner une propriété commune à l'ensemble des éléments qui relèvent de l'analyse prosodique. Etant donné que ce terme revêt diverses significations selon les mouvances théoriques (*cf.* Rossi *et al.*, 1981 ; Rossi, 1999, pour un exposé général du sujet), il nous a paru utile d'en identifier les plus courantes, afin de donner un meilleur éclairage au contenu de cet article.

En règle générale, le terme suprasegmental employé dans une acception concrète, signifie que les variations des paramètres prosodiques (durée, mélodie, intensité) qui actualisent les entités de la structure prosodique (accent, ton, intonation, etc.) s'appliquent à des tronçons de parole *plus grands que le segment* (voyelle ou consonne). Cette référence à *l'empan du segment* ne permet pas toutefois de circonscrire sélectivement les faits prosodiques, car d'autres traits phoniques comme la nasalité ou l'harmonie vocalique possèdent également cette faculté et méritent donc aussi d'être qualifiés de suprasegmentaux (Firth, 1948 ; Lehiste, 1970). La prosodie ne constitue plus alors qu'un sous-ensemble du vaste domaine des faits suprasegmentaux. Il en va de même dans l'approche de Crystal (1969) où le terme suprasegmental fait référence à la fois aux faits prosodiques et à des phénomènes de *qualité vocale* ponctuels ou permanents (raucité, voix soufflée, voix grincée, etc.). L'usage du terme suprasegmental pris dans sa signification première est problématique si l'on tient compte du fait que dans certaines langues les variations signifiantes des paramètres prosodiques peuvent être assignées à des tronçons *plus petits que le segment*. C'est notamment le cas du japonais où la planification prosodique peut s'effectuer en termes de *mores* et où une voyelle peut alors être constituée de deux ou trois mores successives (en japonais, par exemple, la voyelle [i] de *biru* (bâtiment) porte deux mores, par opposition à celle de *biiru* (bière), qui en porte trois. On pourrait interpréter de la même manière la distinction entre « *Sahra* » et « *Sahara* », en français).

Enfin, dans le cours du débat qui s'est récemment instauré sur la nature du *domaine prosodique minimal*, des arguments empiriquement fondés ont été avancés (van Heuven, 1994) pour soutenir l'idée que ce domaine irréductible serait le segment, et non la syllabe, comme l'affirme la tradition). Si une telle proposition était retenue (ce qui est loin d'être le cas), la précédente signification du terme suprasegmental n'aurait alors plus lieu d'être.

En dépit des réserves formulées précédemment, la conception du terme suprasegmental que nous venons de commenter semble être celle qui a été la plus communément acceptée par les linguistes, notamment par les structuralistes américains (Hockett, 1942 ; Trager & Smith, 1951). Pour ces chercheurs, elle s'accompagne cependant le plus souvent de l'idée que les éléments suprasegmentaux sont *en marge* de l'analyse phonologique : « *Les phénomènes qui transcendent les segments sont déclarés métalinguistiques, relèvent de l'étude du style et non de l'analyse microlinguistique. C'est donc ici que s'arrête la phonologie.* » (Trager & Smith, 1951. C'est nous qui soulignons).

Une attitude similaire envers la prosodie s'exprime sous la forme d'une argumentation différente dans le *Dictionnaire de la Linguistique* de Mounin (1974), où il est déclaré que les éléments suprasegmentaux sont des éléments *insegmentables* qui ne peuvent, de ce fait, être analysés en termes *d'unités discrètes* et isolables. Il est évident que cette seconde acception du terme suprasegmental est devenue aujourd'hui totalement obsolète.

Les linguistes ont également attiré l'attention sur le fait que la caractérisation des éléments suprasegmentaux nécessite de recourir à la notion de *contraste syntagmatique* et à la *dimension temporelle* de la chaîne linguistique, par opposition aux éléments segmentaux qui peuvent être définis sur la base de leurs *propriétés phoniques locales intrinsèques* (Lehiste, 1970 ; Nespor & Vogel, 1983). Par exemple, le trait d'arrondissement d'une voyelle peut être établi sans qu'il soit nécessaire de faire référence aux sons adjacents dans la chaîne. En revanche, l'accentuation d'une voyelle ne peut être établie sans la comparaison avec les autres voyelles de la séquence (ce qui a fait dire à Hjelmslev que l'accent n'est pas un prosodème syllabique !). Que doit-on penser alors des tons (des langues à tons) dont le fonctionnement paradigmatique est similaire à celui du phonème ? On entrevoit, ici encore, la difficulté qu'il y a de vouloir trop généraliser la signification du terme suprasegmental.

Il existe cependant une acception de ce terme qui s'avère moins problématique : celle qui consiste à identifier comme suprasegmental ce qui peut « *figurer au-dessus des segments* ». Cette conception, qui s'inspire des remarques prémonitoires de Hockett (1947) sur la non-linéarité des représentations phonologiques et de la métaphore du « *modèle orchestral* » introduite par le même auteur (la transcription d'un énoncé doit être comparable à une partition d'orchestre et donc comprendre plusieurs lignes), est celle qui s'apparente le mieux à la conception moderne de la phonologie, dans

la mesure où les représentations des éléments segmentaux et des éléments suprasegmentaux sont conçues comme des *chaînes autonomes superposables*.

Nous souhaitons soulever un dernier point avant de conclure l'exégèse du terme suprasegmental. Dans la mesure où la phonologie actuelle considère les entités de la représentation phonologique comme des *autosegments*, et que cette qualification s'applique aussi bien aux phonèmes qu'à la prosodie, il pourrait se révéler embarrassant d'en arriver à recourir à des expressions comme « segments suprasegmentaux » ou « autosegments suprasegmentaux » pour en référer à des unités prosodiques.

2.8. Des difficultés de l'analyse prosodique

L'étude prosodique envisagée sous l'angle des relations de la phonétique à la phonologie soulève plusieurs problèmes qui se rapportent à deux thèmes fondamentaux : celui qui concerne *l'analyse physico-formelle* de la prosodie (qui pose la question des niveaux d'analyse et de représentation de la prosodie : de l'acoustique à la phonologie) et celui qui se rapporte à l'étude des *relations entre l'analyse physico-formelle et le sens* (qui touche au problème de l'interprétation des significations transmises par la prosodie).

2.8.1. Les sources de difficulté de l'analyse physico-formelle

Le succès rencontré aujourd'hui par la *phonologie-ancrée-substance* et par le développement de la *phonologie de laboratoire* (Pierrehumbert *et al.*, 2000) attire plus que jamais l'attention des chercheurs sur la nécessité d'établir des liens rigides entre les théories phonologiques de la prosodie et la matérialité des faits prosodiques. Cette matérialité peut être appréhendée à la fois par l'approche objective (qui se rapporte à l'étude acoustique de la prosodie et aux représentations analytiques qui s'y rattachent) et par l'approche subjective (qui concerne l'étude auditive de la prosodie et les notations impressives qui en découlent). Il a été démontré de façon convaincante que l'étude auditive est trop empreinte de subjectivité pour constituer une approche scientifique fiable (Lieberman, 1965). L'analyse objective, qui est servie par le développement d'outils de plus en plus conviviaux et performants, tels que : PRAAT (*www.fon.hum.uva.nl/praat/*), MES (« *Motif Environment for Speech* » : *www.lpl.univ-aix.fr*) ou WINPITCH (*www.winpitch.com/detail/new.html*), s'avère en revanche incontournable pour établir une première classification scientifique des observables.

Les démarches ascendantes (procédant du signal) ou descendantes (procédant de la théorie) qui cherchent à confronter les observables (interprétés en termes de variations paramétriques physiques) et les représentations phonologiques (construites sur des hypothèses linguistiques) rencontrent une difficulté majeure, qui concerne les ruptures de correspondances entre ces niveaux d'analyse et d'interprétation. Ce problème n'est pas propre à la prosodie. Il est cependant particulièrement

complexe pour cette dernière et il nécessite, de ce fait, des éclaircissements, sans lesquels il s'avère difficile de différencier, de prime abord, les variations paramétriques qui sont motivées par des *instructions linguistiques* de celles qui reflètent l'influence de *contraintes dites de bas niveau*.

2.8.1.1. Variabilité des paramètres prosodiques et contraintes de production

Nous examinerons principalement le cas de F0, qui est le paramètre prosodique physique auquel les analyses se réfèrent le plus fréquemment. Si nous tentons d'établir, en premier lieu, la nature des relations qui lient les variations d'un paramètre physique comme la fréquence fondamentale (F0) à l'actualisation des entités phonologiques, lexicale et supralexicale que représentent le ton, l'accent et l'intonation, il s'avère utile, voire indispensable, de connaître la nature et l'importance des contraintes de production qui affectent l'évolution de ce paramètre (Ohala, 1978 ; Xu & Xuejing, 2002). Ces contraintes, qui sont associées à plusieurs facteurs, engendrent des *effets universels* (quoique quantitativement variables d'une langue à l'autre) que nous proposons de décrire brièvement.

2.8.1.1.1. Les contraintes idiosyncrasiques

Il est clairement établi que les variations de F0 sont en partie déterminées par les caractéristiques physiologiques des locuteurs (notamment par la masse volumique des cordes vocales) et que par conséquent la tessiture tonale peut varier d'un sujet à l'autre, et varie de toute évidence systématiquement entre les voix d'homme, de femme et d'enfant. Cette source de variation non-linguistique étant reconnue, il importe de la neutraliser avant toute interprétation linguistique, ce qui peut être réalisé par la mise en œuvre de procédures de normalisation. Ces dernières consistent principalement à convertir, à l'aide de formules appropriées, les valeurs absolues de F0 mesurées en Hertz (Hz) en valeurs relatives (ou en valeurs logarithmiques). Plusieurs échelles sont alors disponibles à cet effet, telles que l'échelle des demi-tons ou l'échelle ERB (Hermes & van Gestel, 1991), qui présentent également l'avantage de correspondre à des échelles auditives (pour obtenir plus de détails, lancer la recherche : « *auditory scales of pitch perception* » sur le web).

2.8.1.1.2. Les contraintes inhérentes à la gestion des variations de F0

Une autre contrainte de production qui affecte également les variations de Fo dans la parole concerne la vitesse avec laquelle celle-ci peut croître ou décroître dans un temps imparti, défini, par exemple, par les limites temporelles qu'impose la prononciation d'une syllabe (en admettant que cette dernière constitue une unité de programmation des variations de F0, comme cela a été observé par Di Cristo, 1978). On peut s'attendre ainsi à ce qu'une importante variation de Fo, activée par une *instruction linguistique* particulière, ne puisse atteindre la *cible acoustique* planifiée à la demande de cette instruction, si la variation est associée à une syllabe particulièrement brève

(intrinsèquement brève ou abrégée par suite d'une accélération du débit de parole). L'effet de troncation des variations de F0, auquel nous venons de faire allusion, est attesté par *l'analyse expérimentale* (Di Cristo & Chafcouloff, 1981). Le recours à des expériences de perception permet également de constater que le système cognitif « *connaît* » cette contrainte et peut la surmonter en reconstituant subjectivement la partie tronquée et, du même coup, la cible planifiée par le locuteur (Rossi, 1981). À la lumière de ces remarques, la prise en considération des contraintes de production qui affectent les trajectoires de F0 et leurs alignements avec le matériau segmental (Xu, 2002, 2004) paraît donc s'imposer comme un préalable à l'interprétation des variations de ce paramètre prosodique, en vue d'établir des liens entre ces variations et les représentations phonologiques qu'elles actualisent.

La figure 3 est une illustration (parmi d'autres) du problème posé par l'interprétation directe de F0. Bien que les deux énoncés procèdent d'une même représentation phonologique sous-jacente de la structure prosodique, les syllabes [fij] et [mi], qui délimitent la fin de la première Unité Intonative de ces énoncés, exhibent, en raison de l'influence du contexte consonantique, des configurations de F0 très différentes. Une interprétation « naïve » de ce document pousserait à conclure que la syllabe [fij] est nantie d'un ton statique haut, et la syllabe [mi], d'un ton montant.

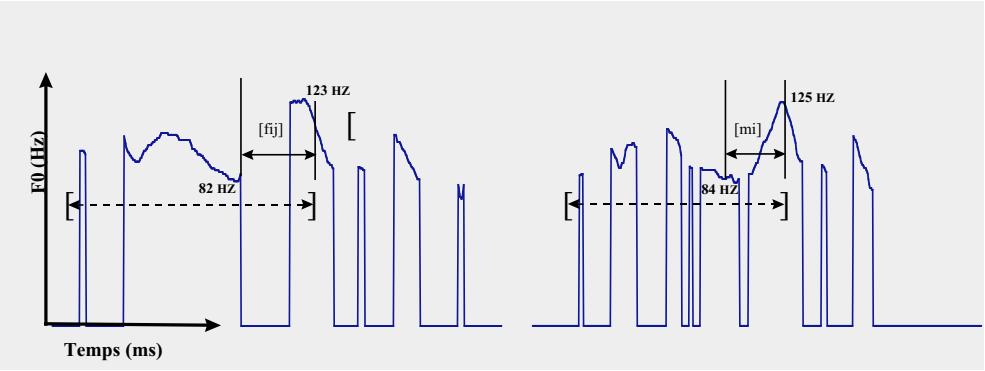

Figure 3

Représentations analytiques des courbes de F0 des séquences [parce que la fille de mon ami] et [parce que l'ami de ma fille], extraites des énoncés : « parce que la fille/l'ami de mon ami/de ma fille n'a pas pensé à lui téléphoner ».

2.8.1.1.3. Les phénomènes dits « d'abaissement » (actualisation de plusieurs niveaux de contraintes)

Outre l'influence des contraintes propres à la gestion des variations de F0 dans les dimensions fréquentielle et temporelle, il convient de mentionner les effets qui résultent de l'incidence d'un

autre paramètre de production. Il s'agit en l'occurrence des variations de la *pression sous-glottique*. En raison de la baisse régulière du volume d'air pulmonaire dans le cours de l'énoncé, la pression sous-glottique (PS) tend à diminuer aussi de façon graduelle. Etant donné la relation physique qui lie les paramètres F0 et PS (Titze, 1989), il est attendu que F0 décroisse pareillement de façon progressive (toutefois, cette décroissance graduelle n'est pas due uniquement à l'influence du paramètre PS, mais aussi à des ajustements laryngés comme le déplacement vertical du larynx : Honda, 2004). Cet effet d'abaissement graduel est couramment appelé : *déclinaison*. Comme il affecte à la fois les minima et les maxima de la courbe de F0 (que l'on peut relier par deux lignes distinctes), il est d'usage de distinguer entre la ligne de déclinaison basse (*Baseline*) et la ligne de déclinaison haute (*Topline*) (Vaissière, 1983).

Dans la perspective de l'interprétation des relations de la phonétique à la phonologie, il est nécessaire de distinguer les effets d'abaissement qui sont dus à l'influence de contrainte de *bas niveau* (ou de production) comme la déclinaison, que l'on vient d'évoquer (pour une plus ample discussion du phénomène, *cf.* Cohen *et al.*, 1982 ; Ladd, 1984 et Gussenhoven & Rietveld, 1988) et ceux qui sont liés à des contraintes linguistiques (syntaxiques, morphologiques ou tonales) comme les phénomènes de *downdrift* et/ou de *downstep*, bien connus de la tradition des recherches en tonologie. Le premier terme (ou, plus simplement, celui de « *drift* ») a souvent été employé dans le passé comme synonyme de déclinaison avant la lettre pour désigner une baisse graduelle de la F0 du début à la fin de l'énoncé (Pike, 1945 ; Maeda, 1976 ; Ohala, 1978). Dans l'étude des langues à tons, ce sont en revanche les termes de *downdrift* et de *downstep* qui sont utilisés, souvent l'un pour l'autre. Dans ce cas, ils se rapportent au fait qu'un ton haut est abaissé sous l'influence du ton bas précédent, ce dernier pouvant disparaître de la prononciation, sans que l'empreinte de son influence sur le ton haut ne soit annihilée. D'autre part, il a été proposé (Laniran, 1992) de réserver le terme de *downdrift* pour désigner un effet de déclinaison qui ne s'applique qu'à une chaîne de tons de « même nature » (des tons hauts, par exemple) et non à l'ensemble des tons hauts et bas d'une séquence donnée. À la différence de la déclinaison, le *downstep* est appréhendé comme un phénomène local et itératif, qui peut se réaliser de façon récurrente dans une séquence où alternent les ton phonologiques hauts et les tons bas (entraînant ainsi une configuration tonale « en terrasses »), qui s'applique à un domaine plus restreint que celui couvert par la déclinaison (un syntagme prosodique au lieu de l'énoncé) et qui aurait la faculté d'assumer une fonction démarcative ou distinctive. En réalité, le statut linguistique du downstep ne semble pas clairement établi en dehors des langues à tons, bien qu'il soit considéré comme un élément important de la théorie de l'intonation de l'anglo-américain (Pierrehumbert, 1980) où il bénéficie d'un codage symbolique (!H). Toutefois, d'importantes analyses statistiques effectuées sur cette langue par

Dainora (2001b) ne permettent pas de justifier une distinction entre les tons étiquetés comme des downsteps et ceux qui ne le sont pas. À la lumière de ces résultats, l'auteur propose donc d'éliminer le downstep de l'inventaire des entités phonologiques de l'anglais. L'effet dit *d'abaissement final* (« *final lowering* ») est également décrit comme un *phénomène local* qui concerne la baisse de la hauteur des tons au terme d'un énoncé déclaratif (Liberman & Pierrehumbert, 1984).

Étant donné que le phénomène de déclinaison se rapporte à des *contraintes de production*, il est normal que sa présence soit attestée dans l'ensemble des langues naturelles (langues à accent et langues à tons), aussi bien pour la lecture que pour la parole spontanée. En ce qui concerne les langues à accent, il est particulièrement intéressant de vérifier si le phénomène de déclinaison affecte à la fois la réalisation des pics de F0 associés aux accents (plus précisément, l'échelle des valeurs de la F0 des pics : « *peak scaling* » : *cf.* Prieto *et al.*, 1996) et la perception de ces pics. La *réalité auditive* de la déclinaison a été testée expérimentalement, pour la première fois à notre connaissance, par Breckenridge (1977), alias Pierrehumbert (1979). Ces travaux montrent notamment que le phénomène de déclinaison a une incidence sur la perception de la hauteur relative des pics mélodiques, dans la mesure où, d'une part, deux syllabes accentuées successives sont perçues à la même hauteur lorsque la seconde possède une valeur de F0 plus basse que la première et où, d'autre part, la seconde syllabe est perçue plus haute lorsqu'elle possède la même valeur de F0 que la précédente. Cet effet suggère la mise en œuvre d'un *dispositif de compensation* qui serait informé par un « *cadre de référence* » où la déclinaison jouerait un rôle décisif. Il reste à vérifier si ce dispositif est opératoire dans toutes les langues naturelles, y compris dans les langues à tons.

Une mise au point utile sur la signification des termes relatifs aux phénomènes d'abaissement a été proposée récemment par Connell & Ladd (1990) et par Laniran & Clements (2003). On trouvera également un historique de la notion de downstep particulièrement éclairant dans l'étude de Rialland (1997).

2.8.1.1.4. Les contraintes interactives : effets microprosodiques

Les variations de F0, en particulier, et celles des paramètres prosodiques physiques, en général, sont soumises à des contraintes de production « *interactives* », qui résultent dans ce cas d'une interaction entre ces paramètres et la prononciation du matériau segmental. Les phénomènes engendrés par ce type de contrainte sont appelés *microprosodiques*, car leur empan est très limité et n'excède pas la taille du segment phonémique. Il existe deux sortes de phénomènes microprosodiques, qui sont qualifiés par les termes : *intrinsèques* et *co-intrinsèques*, respectivement (Di Cristo, 1978). La figure 3 résume la teneur de ces phénomènes et fournit une explication lapidaire de leur origine.

Les *effets intrinsèques* sont dénommés ainsi parce qu'ils concernent l'influence locale que peut exercer la production de certains segments (voyelles ou consonnes) sur la configuration et les valeurs d'un paramètre prosodique physique donné. C'est ainsi que les voyelles hautes (ou fermées) ont tendance à accroître systématiquement les valeurs de F0 qui sont associées à leur production (y compris dans les langues à tons, *cf.* Connell, 2002) ; que l'émission d'une *voyelle nasale* ou d'une voyelle basse (ou ouverte) s'accompagne d'une durée plus grande que celle d'une voyelle haute ; que dans un même contexte, la voyelle [a] exhibe une intensité objective supérieure à celle de [i], etc. Il est loisible d'expliquer le premier phénomène par un effet de couplage physiologique entre le soulèvement de la langue et l'attraction qu'elle exerce sur le larynx, ce qui conduit à accroître la tension des cordes vocales et, par voie de conséquence, la valeur de F0 ; le second, par des contraintes articulatoires concernant la vitesse de déplacement et l'inertie relative des organes articulateurs : la réalisation d'un [i], qui ne nécessite ni ouverture du voile du palais (comme pour les nasales), ni abaissement de la mandibule (comme pour les voyelles ouvertes) prend moins de temps ; le troisième, par une contrainte perceptive qui interfère avec la production de la voyelle en question (Rossi, 1971b).

Bien qu'ils ne soient pas consignés sur la figure 3, les effets intrinsèques concernent également les consonnes. Il s'agit notamment de l'importante baisse de F0 qui accompagne l'émission des obstruantes voisées, alors que les consonnes sonorantes (les nasales, les liquides et les glides) exhibent un effet moins marqué, voire quasi inexistant (Di Cristo, 1978). C'est pourquoi, les méthodes de réitération, qui souhaitent s'affranchir de ces effets, utilisent parfois des séquences de syllabes [la] ou [ma] (Liberman & Streeter, 1978).

Les *phénomènes co-intrinsèques* doivent leur appellation au fait qu'ils relèvent de la coarticulation, en l'occurrence de l'influence des consonnes sur les caractéristiques prosodiques des voyelles adjacentes. L'effet dit de « *pitch skip* » (*saut de hauteur*) se rapporte au fait que la valeur de F0 à l'attaque d'une voyelle est nettement plus élevée lorsque cette dernière est précédée d'une *consonne non-voisée*. La valeur-cible de cette voyelle est également soumise à la même influence, mais dans une proportion moindre. En revanche, la présence d'une *consonne obstruante voisée* a pour effet d'abaisser à la fois la valeur initiale et la valeur-cible de la voyelle subséquente. Les consonnes sonorantes exercent peu d'effets sur la F0 des voyelles adjacentes. Il est également établi que les consonnes influencent à des degrés divers la durée des voyelles qui font partie de la même syllabe. C'est ainsi que dans une séquence : [C1 V C2], la présence d'une [C2] non-voisée abrège considérablement la durée de la voyelle [V], alors que celle d'une consonne constrictive voisée l'allonge tout aussi considérablement (Di Cristo, 1978).

Variations microprosodiques: les voyelles	F0	Durée	Intensité
Intrinsèques	**Effet**: F0 des voyelles *hautes* supérieure à celle des voyelles *basses*. **Explication**:couplage physiologique: Théorie de «l'attraction linguale»	**Effet** voyelles *nasales* > voyelles orales *basses* > voyelles orales *hautes*. **Explication**: vélocité des effecteurs articulatoires	**Effet**: [a] >[i] **Explication**: contrainte perceptive (*bandes critiques*)
Co-intrinsèques	**Effet**: après consonne *non-voisée*: - saut de F0 (effet à court-terme). - augmentation de F0 (effet global) **Explication**: facteurs aérodynamiques	**Effet**: dans: [C1VC2]: [C1] et [C2] *non-voisées* (principalement [C2]) abrègent [V] **Explication**: timing articulatoire et facteurs aérodynamiques	**Effet**: intensité de [V] supérieure après consonne *voisée* **Explication**: facteurs aérodynamiques: pression sous-glottique

Figure 4
Synthèse des principaux effets intrinsèques et co-intrinsèques associés à la production de F0

Il serait fastidieux de recenser dans cette section tous les effets microprosodiques et d'en expliquer les causes. La figure 4 n'est proposée qu'à titre indicatif et nous engageons le lecteur à se reporter à Di Cristo (1978) pour une présentation détaillée de tous ces phénomènes. Cependant, nous souhaitons attirer l'attention sur le caractère non trivial de la problématique que suscite leur omniprésence. Imaginons un instant un système de traitement automatique des observables aveugle à la connaissance des effets intrinsèques et co-intrinsèques. Un système caricatural de ce type sera « tenté » de décider que les voyelles hautes reçoivent un ton haut et les voyelles basses, un ton bas ; que seules les voyelles ouvertes et les nasales sont linguistiquement longues ; que les tons mélodiques sont l'apanage des voyelles entourées de consonnes voisées et les tons statiques, celui des voyelles entourées de consonnes non-voisées, etc. C'est afin d'obvier à ces erreurs d'interprétation que nous avons conçu des traitements destinés à l'effacement des variations microprosodiques intrinsèques et co-intrinsèques (Di Cristo, 1978 ; Rossi & Di Cristo, 1980) et une méthode objective conduisant à dissocier le *profil segmental* (microprosodique) du *profil suprasegmental* (macroprosodique) d'une courbe de F0 (Di Cristo & Hirst, 1986), étant entendu que cette dernière procédure s'appuie sur l'hypothèse que seul le profil suprasegmental a la faculté d'actualiser la projection des instructions linguistiques de haut niveau.

2.8.1.2. Variabilité des paramètres prosodiques et contraintes de perception

Les retombées acoustiques des effets intrinsèques et co-intrinsèques ne constituent pas l'unique source de difficulté de l'analyse prosodique. Une autre problématique émerge, en effet, de l'observation des ruptures de correspondances entre les *signaux physiques et leur perception.*

Ainsi que l'ont montré des expériences de psycho-acoustique, la plupart des variations microprosodiques sont *perceptibles*. Il ne faut pas perdre de vue, à ce propos, que la *tonogenèse* (l'émergence d'un système phonologique tonal) est liée à l'incidence de contraintes microprosodiques (Hombert, 1976). De même, il a été démontré que les variations microprosodiques contribuent à l'identification des classes de segments phonémiques (Di Cristo, 1978). Il n'en demeure pas moins que les variations microprosodiques ne sont pas *perçues* (dans le sens de « prises en compte ») quand il s'agit de traiter *l'information prosodique suprasegmentale*, ce qui suggère que les systèmes de traitement perceptif du segmental et du prosodique sont indépendants.

Même si l'on écarte de l'analyse les variations *microprosodiques*, il reste que les variations *macroprosodiques* ne sont pas perçues *en l'état*. À titre d'exemple, le terme d'une variation positive ou négative intra-syllabique de F0 n'est pas intégré perceptivement (en raison d'une contrainte inhérente au système auditif) et, de ce fait, la cible auditive perçue ne correspond pas à la cible acoustique (*cf.* par exemple, les travaux fondateurs sur la perception du « *glissando* » de Rossi (1971b), repris de nos jours par d'Alessandro & Mertens (1995). Il en va de même des variations de l'intensité et de la rupture que cela entraîne entre l'interprétation de l'intensité objective et de l'intensité subjective (Rossi, 1971b).

Un autre effet de rupture de correspondance entre le signal prosodique et sa perception a pour origine des contraintes de « *perception différentielle* ». Afin d'être perceptible, il est nécessaire, en effet, que la différence de F0, de durée, ou d'intensité entre deux syllabes consécutives soit supraliminaire (Rossi *et al.*, 1980).

Enfin, les ruptures de correspondance sont liées à un phénomène qui n'a pas donné lieu, jusqu'à présent, à des recherches approfondies et que nous abordons, par conséquent, sous la forme d'une question. Etant donné le caractère *pluriparamétrique* de la prosodie, comment les différents paramètres interfèrent-ils pour construire les *percepts prosodiques* qui constituent les points d'ancrage des structures phonologiques, comme la construction du percept de proéminence, par exemple (car la variation d'un seul paramètre - quand bien même il s'agirait de F0, dont on a pu établir dans le passé le rôle prééminent - peut s'avérer insuffisante pour former un percept prosodique signifiant : Neuhoff, 2003). À vrai dire, les données actuellement disponibles sur les interactions des paramètres sont encore fragmentaires. Nous savons, par exemple, que la baisse simultanée de F0 et de l'intensité interfèrent dans la perception de la chute mélodique finale d'un énoncé (Rossi, 1978). De même, il est établi que la durée physique d'une syllabe et la variation de F0 qui

accompagne son émission interfèrent dans la perception de la longueur de cette unité ; ou que la durée de la syllabe a une influence sur la perception de sa sonie (Di Cristo, 1978 ; Rossi *et al.*, 1980). L'interaction entre les variations de F0 et le spectre doit également être considérée, dans la mesure où il apparaît que la perception des variations de la hauteur est sensible à la stabilité spectrale (House, 1990). Une découverte importante dans ce domaine concerne l'incidence du spectre sur la sonie, notamment le fait que l'intensité perceptivement efficace d'une syllabe (dans la perception de l'accentuation, par exemple) n'est pas corrélée à la mesure de l'intensité globale RMS (*Root Mean Square*) qui constitue la procédure habituelle, mais à l'intensité comprise dans une zone particulière du spectre, supérieure au premier formant. Ce phénomène est identifié dans la littérature sous la référence de « *balance spectrale* », ou de « *tilt spectral* » (Sluijter & van Heuven, 1996).

En l'absence d'un *modèle intégratif global* de la perception des faits prosodiques, capable de gérer une transmutation automatique des données physiques en données perçues, de nombreuses tentatives ponctuelles se sont appliquées à mettre au point des méthodes de *stylisation prosodique* centrées sur des paramètres particuliers, notamment la F0 et la durée. Des recherches de ce type ont été particulièrement développées par l'école d'Eindhoven (*cf.* 't Hart *et al.*, 1990, pour une synthèse) et l'école d'Aix (*cf.* Rossi *et al.*, 1980, pour une synthèse). Malgré ces travaux, nous ne disposons pas encore d'un cadre théorique unificateur sur le rôle de la prosodie dans la perception de la parole. Cependant, un article de synthèse récent fait bien le point sur la question (Vaissière, 2004).

Si la phonologie actuelle ambitionne, comme elle le prétend depuis peu, de *s'ancrer à la substance*, il est clair que la phonologie prosodique ne peut ignorer les phénomènes que nous venons de mentionner. C'est ici que la *phonétique interprétative*, conçue comme une interface entre les faits physiques et la représentation phonologique, a probablement un rôle décisif à jouer.

2.8.2. Les difficultés liées à l'interprétation de la relation des objets physiques et des objets formels au sens

2.8.2.1. Remarques préliminaires

L'approche linguistique actuelle de la prosodie tend à concevoir cette dernière comme un champ de recherche cohérent dont l'objectif global est d'établir l'inventaire des relations entre les formes et les fonctions qu'entretiennent les éléments de l'expression phonique appartenant au domaine des phénomènes dits « *suprasegmentaux* ». La perspective fonctionnelle est inséparable du recours au sens et à l'usage du langage, que les mouvances du structuralisme, de la grammaire formelle et de la phonologie autonomiste ont tenu à l'écart (Fuchs & Le Goffic, 1992). Le danger d'une telle ségrégation a été signalé très tôt par Paul Garde, lorsqu'il affirmait notamment : « *On peut (donc) concevoir la linguistique à tous les niveaux de l'analyse comme l'étude des variations concomitantes du signifiant et*

du signifié. Toute méthode qui ignore la nécessité de ce double point d'appui réduit la linguistique à une simple énumération. » (Garde, 1966 : 1). L'idée que la structure des langues se ramène à un pur jeu de formes exploitables de façon indépendante semble faire place, de nos jours, à une autre approche cognitive de la linguistique pour laquelle l'exploitation des données sur *l'usage du langage* constitue une source d'information non négligeable pour comprendre et tenter de modéliser ce qui se passe dans la tête des locuteurs.

La prise en compte de la signification est cruciale en ce qui concerne la prosodie, car il est notoire que la façon dont un énoncé est prononcé exprime beaucoup plus que l'information lexicalement encodée par les mots dont il est formé, ou, en d'autres termes, que l'information transmise par son sens littéral. S'il en est ainsi, c'est parce que les systèmes prosodiques recèlent une plus grande capacité à véhiculer des distinctions de sens que les autres systèmes. Ils disposent pour ce faire d'une organisation phonique et d'une architecture formelle bien plus complexes, dont la linguistique traditionnelle n'a pas réussi à capturer toute la subtilité.

Trois questions principales se posent au sujet de l'articulation des formes et des fonctions en prosodie : Quelles sont les fonctions générales assumées par la prosodie ? Quelles sont les unités prosodiques signifiantes ? Quels types particuliers de significations sont associés à ces unités ?

2.8.2.2. Fonctionnalité, signification et articulation expression/contenu

La prosodie se caractérise avant tout par *l'hétérogénéité de sa fonctionnalité* et, partant, par son aptitude à véhiculer conjointement des significations très diverses, notamment dans l'exercice du discours et de la conversation (Bertrand, 1998). Cette richesse fonctionnelle représente une première source de difficulté au regard de l'interface formes/fonctions ou expression/contenu. La seconde difficulté tient au fait que l'interprétation de la prosodie participe d'une double sémiotique (Di Cristo, 2000a) :

- d'une *sémiotique syncrétique* : pour deux raisons : les signaux prosodiques ne sont pas toujours signifiants par eux-mêmes et se combinent à d'autres signaux (verbaux et para-verbaux) pour transmettre du sens. D'autre part les signaux prosodiques véhiculent en parallèle plusieurs types d'informations linguistiques et idiosyncrasiques ;
- d'une *sémiotique à géométrie variable*, dans la mesure où un même signal prosodique peut recevoir des interprétations différentes en fonction du *contexte* (ce terme étant pris dans sa signification la plus large, *cf.* Akman & Bazzanella, 2003).

Pour ces raisons et malgré les nombreuses correspondances observées entre fonctions et formes prosodiques, ces relations sont loins d'être systématiquement bi-univoques (*cf.* Dik, 1989 ; Lambrecht, 1994 ; Givón, 1995 ; Smith, 2000 pour une interprétation des ruptures de correspondances entre formes et fonctions, en général). L'hétérogénéité fonctionnelle de la prosodie se manifeste par l'expression concomitante de plusieurs classes de fonctions, dont les

mieux connues sont : les fonctions « *délimitative* » et « *intégrative* » qui réalisent l'empaquetage du matériau verbal et sa mise en forme (aux niveaux du lexique, de l'énoncé et du discours, en relation avec la phonologie, la lexicologie, la morphologie, la syntaxe et la macrosyntaxe), la fonction de « *mise en perspective de l'information* » (signalement du : connu/nouveau, topique/commentaire, thème/rhème, des diverses formes de focalisation et de la structure rhétorique du discours), la fonction « *expressive* » (expression des attitudes, de l'affect et des émotions), la fonction « *interactionnelle* » (diverses marques de négociations interpersonnelles), et la fonction « *identificatrice* » (caractéristiques idiolectales, dialectales, sociolectales des individus et des styles de discours qu'ils pratiquent). Le problème est de savoir sur quels critères il est permis de décider si une catégorie fonctionnelle donnée relève ou pas de l'analyse linguistique proprement dite. Cette distinction est d'importance dans la perspective d'un « *mapping* » entre l'organisation *physico-formelle/structurelle* de la prosodie et *l'interprétation sémantique et pragmatique* de cette organisation.

Pour amorcer la discussion de cette problématique, il nous a paru opportun de rapporter ici la définition de l'intonation proposée par Ladd :

> *« J'utilise le terme intonation pour en référer à l'usage que l'on fait des éléments phonétiques suprasegmentaux pour véhiculer au niveau postlexical, ou de la phrase, des significations pragmatiques d'une manière linguistiquement structurée. »* (Ladd, 1996 : 6 ; traduction par nos soins).

Il est clair que l'évocation de la structuration linguistique du matériau phonétique suprasegmental fait ici référence à la phonologie et que cette définition aborde donc la problématique de l'analyse de l'intonation comme celle de *l'interface phonologie/pragmatique.* Cette conception a le mérite d'ouvrir une perspective plus large que ne le font les approches linguistiques classiques en termes d'interfaces phonologie/syntaxe ou syntaxe/sémantique. Nous irons encore plus loin en suggérant que cette perspective conduit logiquement, si on l'adopte, à chercher à établir des liens entre les *propriétés de la grammaire* (dont la composante interprétative s'appuie traditionnellement sur la sémantique véri-conditionnelle) et les *principes gricéens* de l'analyse conversationnelle (qui intègrent la référence au contexte). La démarche est fondée, si l'on admet que les *principes de la grammaire* et les *principes pragmatiques* interagissent dans la détermination du contenu propositionnel des énoncés (Kempson, 1988). De surcroît, elle s'inscrit logiquement dans le paradigme de la recherche des liens entre la prosodie et la cognition, dans la mesure où il est affirmé que la pragmatique relève d'une théorie générale de cette dernière (Sperber & Wilson (1986).

La conception de l'interfaçage de la grammaire et de la pragmatique et l'interprétation fonctionnelle de cette relation nécessitent que l'on apporte des précisions sur la *gamme des significations* véhiculées par la prosodie et sur les *ressources phonologiques et phonétiques* qui sont mises en

œuvre pour traduire ces significations, bien qu'il n'existe pas encore, à notre connaissance, une théorie générale capable de rendre compte de l'intégration de l'ensemble de ces trois phénomènes.

Les questions soulevées par le problème de l'interface grammaire/pragmatique se rapportent à l'interprétation d'un ensemble de termes, tels que : linguistique, paralinguistique, pragmatique, pragmalinguistique, phonologie, phonétique, système, contraste et discrétion. Le point de départ du débat concerne la dichotomie : *linguistique/paralinguistique*, qui repose à la fois, comme le montre la littérature, sur des considérations *phonético-formelles* et *fonctionnelles*.

L'ancien usage du terme suprasegmental fait référence à l'ensemble des phénomènes appartenant à la « *vocalité* ». Ce domaine est subdivisé en deux branches qui regroupent respectivement, d'après Crystal, les *systèmes prosodiques* et les *systèmes paralinguistiques* (Crystal, 1969). Les éléments constitutifs des premiers sont la mélodie (direction et gamme), la pause, l'intensité, le tempo et la rythmicité, alors que les seconds se rapportent essentiellement à des phénomènes liés à la « *qualité de la voix* » (chuchotement, voix soufflée, voix de fausset, raucité, etc.). Des variations fines de la qualité de la voix ou du mode phonatoire peuvent être porteuses d'informations qui renseignent, à l'instar des faits prosodiques, sur l'humeur et l'attitude du locuteur et, plus généralement sur le style langagier (Laver, 1980 ; Gobl & Ni Chasaide, 1992). Ces variations peuvent également participer à l'identification des proéminences actuelles et des frontières de constituants. Il n'est donc pas surprenant que des travaux récents suggèrent de ne plus opposer, comme le fait la tradition, prosodie et qualité de la voix (Ni Chasaide & Gobl, 2004 ; Auran, 2004).

Il est clairement établi que les éléments appartenant aux deux types de systèmes véhiculent des informations qui contribuent, à des degrés divers, à *l'interprétation pragmatique* des messages proférés par les locuteurs. Selon Crystal (1975), s'il est loisible de démontrer qu'un effet vocal est *systématique*, *contrastif* (dans la mesure où il permet de juger qu'un énoncé est *sémantiquement différent* d'un autre) et qu'il assume une *fonction communicative* que partagent les membres d'une communauté donnée, on devrait en conclure que cet effet appartient au *système sonore global* d'une langue et qu'il relève donc, à ce titre, de *l'analyse linguistique*. Donc, l'analyse linguistique doit intégrer des faits linguistiques et des faits para-linguistiques, contrairement à ce que cette dichotomie terminologique donnerait à penser. Elle n'intègre pas, cependant, la totalité des faits paralinguistiques, lorsque ces derniers se rapportent aux phénomènes kinésiques non-verbaux (activités mimo-gestuelles) dont l'étude conjointe à celle de la vocalité relève davantage de l'analyse de la *communication* que de la *linguistique* proprement dite.

Dans la mesure où *l'interprétation pragmatique*, dans le contexte de la « *cognition située* » (Lave, 1988), prend en compte, outre l'information verbale, à la fois des informations prosodiques et des informations paralinguistiques, on doit convenir soit qu'elle est une *partie intégrante de l'analyse*

linguistique (ce dont conviennent de nombreux linguistes), soit qu'elle instaure des *liens rigides* avec cette dernière. Mais quelle est alors la relation de la pragmatique à la grammaire susceptible de concerner la phonologie ?

Une bonne partie de la linguistique formelle actuelle semble rester traditionnellement attachée à une conception tranchée de la distinction linguistique/paralinguistique, telle qu'elle a été formulée à l'origine par Trager (1958, 1964). Pour ce dernier, en effet, la linguistique au sens propre (*linguistic proper* ou *microlinguistique*) ne concerne que l'analyse des *systèmes phonologiques et morphologiques* et, pour ce qui est de la prosodie, celle des traits distinctifs de *jointure*, de *hauteur* et *d'accentuation*. Pour de nombreux auteurs (*cf.* Crystal, 1975, pour une synthèse de la question), cette manière de concevoir la prosodie est arbitraire, car elle procède d'une simple *transposition* des principes et des critères de l'analyse phonémique traditionnelle à l'étude de la prosodie, sans que la question préalable du bien fondé de cette transposition ne soit véritablement posée :

> *« The valuable discreteness of phonemic and morphemic definition, with their amenability to clear-cut substitutability tests of an either/or character, usally free from overlap, does not exclude the fact that there are parts of language which are not amenable to such treatment, but are more accurately and reastically covered by setting up scales of contrastivity, which are of a more/less character... There seems to be no real reason why these [prosodic and paralinguistic] features should be excluded of the field of linguistic proper, and why they should not be taken as wholly intralinguistic items, but of different kind from phonemic and morphemic units, as there are normally understood. »* (Crystal, 1969 : 190 ; c'est nous qui soulignons).

Les précédentes remarques mettent en relief une question « *brûlante* » qu'avait déjà formulée Bolinger (1961) et suivant laquelle, à côté des *contrastes catégoriels*, il existe aussi des *contrastes graduels* dont l'occultation conduit à ignorer une grande partie des significations véhiculées par la prosodie. Cependant, comme l'avait déjà souligné Bolinger en son temps, l'important n'est pas de débattre des bienfaits respectifs du catégoriel et de la gradation, puisqu'ils *coexistent*, mais de savoir *comment traiter* cette dernière dans une approche linguistique non ségrégative.

Les propositions que nous soumettons visent à fournir des éléments de réponse à cette question non triviale. Si l'on tient à rester fidèle à l'esprit de la Grammaire telle que la conçoit le modèle génératif, il est clair que celle-ci ne s'intéresse qu'aux *contrastes catégoriels discrets*, constitutifs du système phonologique. En acceptant ce point de vue, il est loisible de montrer que des contrastes prosodiques de ce type participent effectivement à l'interprétation pragmatique des énoncés (notamment à celle de la contextualisation informationnelle, *cf.* von Heusinger, 2002), ce qui est compatible avec la définition de l'intonation de Ladd (*cf. supra*) et avec notre façon de concevoir l'articulation de la pragmatique à la grammaire (*cf. supra*, figure 2). Toutefois, une grande partie de l'interprétation sémantico-pragmatique (notamment, celle qui se rapporte au « *contenu propositionnel enrichi* » : Sperber & Wilson, 1986) peut reposer sur des *contrastes prosodiques non-discrets*, mais

néanmoins *conventionnels* et *systématiques (cf.* Watt, 1994). Ces derniers, que nous proposons d'appeler « *contrastes prosodiques paraphonologiques* », pourraient être considérés, selon nous, comme des entités constitutives d'une grammaire *non-discrète* (*cf.* le concept de « *non-discrete grammar* » de J.R. Ross, 1970, cité par Crystal, 1991, p. 235). dans la mesure où ils ne valent pas par *leur présence ou leur absence,* mais par des *distinctions de degrés.* Cela ne signifie pas pour autant qu'ils sont inexorablement assimilables à un *pur continuum*, comme le sous-entend par exemple l'hypothèse de la « *variabilité graduelle libre* » (*Free Gradient Variability Hypothesis*), qui est adoptée dans le modèle d'intonation développé par Pierrehumbert (1980) et par Beckman & Pierrehumbert (1986) afin d'écarter le problème posé par certaines variations prosodiques liées à la réalisation des proéminences, au motif qu'elles sont de nature *paralinguistique* (*cf.* les critiques de Ladd, 1994, à ce sujet). Il est en effet envisageable de traiter, au moins partiellement, la composante dite graduelle de la grammaire non-discrète en termes de *distinctions scalaires*, en ayant recours, par exemple, à un modèle de quantification approprié (*cf.* Moxey & Sanford, 1993).

Il n'existe pas encore, à notre connaissance, un modèle susceptible de déboucher sur un cadre représentationnel explicite de la variabilité prosodique. Cependant, nous avons mis en œuvre, dans notre programme de recherche sur les relations de la prosodie au discours (Di Cristo *et al.,* 2004) une démarche qui s'efforce d'apporter une contribution à cette entreprise. Nous proposons en effet de noter les faits prosodiques en termes de distinctions catégorielles discrètes binaires, telles que : présence *vs* absence d'une marque de frontière, de proéminence, d'allongement, etc., tout en en spécifiant conjointement des degrés de frontière, de proéminence, d'allongement, etc. Les hypothèses qui sous-tendent cette procédure sont que les degrés ainsi encodés (qui doivent être validés expérimentalement par des tests de perception), sont culturellement signifiants et constituent un inventaire fermé. Il est possible que la quantification scalaire ne permette pas de résorber toutes les manifestations de la variabilité prosodique graduelle signifiante, notamment celles qui s'attachent à l'expression de l'affect. Cette variabilité, qui se rapporte à l'aspect « gestuel » de la prosodie, est évoquée par la « *métaphore de la bête* » (Hayes, 1994), qui est une réminiscence de l'idée chère à Bolinger selon laquelle la prosodie est incrustée dans une matrice de réactions instinctives. Toutefois, comme le fait observer Hayes, la bête « *connaît la grammaire* », dans la mesure où la variabilité gestuelle s'ancre sur les contrastes phonologiques et tend à les renforcer. Une autre hypothèse, qui est également compatible avec la formulation précédente, mais qui reste à vérifier, serait celle de la *distribution complémentaire,* hypothèse selon laquelle la variabilité gestuelle occuperait les espaces laissés libres par l'expression des contrastes phonologiques proprement dits (Gussenhoven, 1999).

L'expression de la prosodie procède de trois *codes biologiques* (Gussenhoven, 2003) : le *code de fréquence* (fréquence haute : petitesse, soumission/fréquence basse : grosseur, autorité), le *code d'effort*

(proportionnalité de l'effort vocal au degré d'emphase) et le *code de production* (marquage des débuts et des fins des événements). La mise en œuvre permanente de ces codes, qu'actualise l'implémentation phonétique, dénote la valeur éminemment *iconique* de la prosodie. Il n'en demeure pas moins qu'une grande partie de la variabilité iconique fait l'objet d'une *grammaticalisation* (pour le code de fréquence, la distinction question/assertion ; pour le code d'effort, le signalement des proéminences linguistiques ; pour le code de production le balisage des domaines prosodiques). Cette grammaticalisation participe ainsi d'un encodage en termes de contrastes discrets et/ou scalaires qui diffèrent selon la langue. Malgré tout, une part résiduelle de la variabilité demeure graduelle et difficile à traduire en catégories, au dire des chercheurs. Elle ne doit pas être exclue pour autant de *l'analyse linguistique* au sens large (qui *intègre* pour nous la phonétique et la phonologie, d'une part et la pragmatique, de l'autre), car elle est une des *propriétés inhérente* de l'usage des langues. Il importerait dès lors d'élaborer un *modèle interprétatif des contrastes graduels* dont l'absence fait obstacle à l'avancement de la recherche prosodique (Crystal, 1975).

Les remarques précédentes concernant la « *métaphore de la bête* » (Hayes, 1994) et les codes biologiques nous ramènent à la question centrale qui concerne l'articulation de la phonétique et de la phonologie prosodiques et sa relation au sens. Nous formulerons trois conclusions partielles sur ce point.

La première, qui s'efforce de reprendre le contenu de l'ensemble de cette section, résume la complexité de la problématique et peut se formuler de la sorte : toutes les variations prosodiques mesurables ne sont pas forcément perceptibles et toutes les différences perceptibles ne sont pas nécessairement perçues ; quand elles le sont, elles ne sont pas nécessairement signifiantes, et quand elles sont signifiantes elles n'expriment pas que des contrastes phonologiques.

La seconde conclusion concerne plus précisément la nature des relations entre expression et contenu. De ce point de vue, il est permis d'identifier deux types de liens : un *lien direct*, traduisant des aspects iconiques conventionnels, mais non grammaticalisés et donc enclins à la *variabilité graduelle* ; et un *lien indirect* grâce auquel les relations entre les événements physiques et le sens sont *médiatisées* par des représentations phonologiques (contrastes discrets) et paraphonologiques (contrastes scalaires). Certains auteurs considèrent que la *relation directe* est le propre du fait paralinguistique (Ladd, 1996 ; von Eusinger, 2003), ce qui est une autre manière (*cf. supra*) de concevoir le paralinguistique : dans ce cas, le terme *linguistique* ne fait plus référence qu'au système de contrastes discrets et celui de *paralinguistique*, à la totalité des effets de sens qui découlent d'une relation directe entre un phénomène vocal concret (prosodique ou non, selon la distinction présentée plus haut) et un contenu particulier. L'opposition langage/paralangage est donc fondée ici sur *la façon dont la relation son - sens est structurée*. Le danger d'une telle conception est qu'elle peut conduire, dans sa version la plus radicale, à soutenir l'idée que, appliquée à la face sonore des

langues, l'étiquette « *linguistique* » ne doit être attribuée qu'aux éléments qui entretiennent des contrastes catégoriels. Personnellement, nous préférons maintenir l'affirmation de l'existence conjuguée d'une « *linguistique de la langue* » et d'une « *linguistique de la parole* » (même si ces expressions ne sont pas prises dans la même acception - comme il ressort des commentaires qui précèdent et qui suivent - que celle qui leur a été donnée à l'origine par F. de Saussure).

En vérité, nous ne souhaitons pas retomber dans le piège d'une distinction dichotomique préétablie. Nous préférons plutôt adopter, en l'état des connaissances, le profil d'une heuristique et nous rallier, dans cette perspective, à l'hypothèse de « *l'échelle de linguisticité* » proposée par Crystal (1969, 1975). Selon cette échelle, la *polarité positive* correspondrait aux phénomènes que l'on peut identifier comme appartenant à des systèmes hiérarchiques minimaux de contrastes discrets, qui attestent d'une relation indirecte entre les événements physiques et le contenu, que l'on peut décrire concrètement en se référant à des variables paramétriques permanentes (F0, durée, intensité, timbre) et qui entretiennent des relations avec d'autres composantes constitutives de la structure de la langue, comme la syntaxe, par exemple. À *l'autre extrémité* de l'échelle se situeraient les phénomènes non discrets, non systématiques et non intégrés à la structure de la langue, comme les effets vocaux de « raucité », ou de « voix grincée », qui transmettent cependant des informations prises en compte lors de *l'interprétation pragmatique des messages contextualisés*. Il est intéressant de constater que cette conception se trouve être assez proche de celle qu'ont récemment défendue des chercheurs influents et représentatifs de la mouvance de la phonologie de laboratoire : « *In short, knowledge of sound structure appears to be spread along a continuum* » (Pierrehumbert *et al.,* 2000, p.10). La troisième conclusion est plus générale. La conception de la *phonétique* et de la *pragmatique* comme des interfaces entre la matière (physique et conceptuelle, respectivement) et la grammaire, conduit à une révision de la nature des relations entre expression et contenu qui, eu égard à la pluralité fonctionnelle des usages de la prosodie, doit aller bien au-delà de ce que propose le modèle classique inspiré de la phonologie segmentale. L'un des aspects cruciaux de cette révision concerne l'interprétation de la dualité des relations (directes/indirectes) exposées précédemment. Afin de rendre compte de la gamme des significations véhiculées par la prosodie dans le cadre d'une approche linguistique élargie, doit-on inclure dans la grammaire formelle des contrastes scalaires ? (*cf.* figure 5). Comment rendre compte de l'interaction entre les régularités phoniques qui relèvent du paralangage et le système linguistique noyau construit sur des contrastes catégoriels ? À ces deux questions ouvertes, il convient d'ajouter une troisième interrogation, qui se rapporte à une problématique évoquée dans la première partie de cet article, mais dont nous n'avons pas parlé dans la précédente section, faute de connaissances suffisantes : comment intégrer au tableau précédent le concept de grammaire probabiliste et quelles informations cette grammaire est-elle susceptible

d'apporter à l'interprétation des « *représentations intégrées* » qui émanent de l'interfaçage entre la pragmatique et les représentations physico-formelles de la prosodie ?

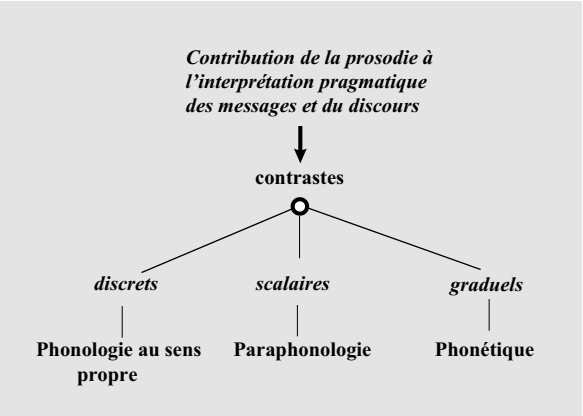

Figure 5
Proposition d'une typologie des contrastes linguistiques impliqués dans l'interprétation sémantico-pragmatique des énoncés et du discours

2.8.2.3. De la nature des unités prosodiques signifiantes et des significations qu'elles véhiculent

2.8.2.3.1. Nature des unités prosodiques signifiantes

La longue discussion qui précède est centrée sur la difficulté majeure que constitue l'articulation forme/sens dans l'analyse de la prosodie. Sans abandonner cette problématique, nous nous intéressons plus précisément, dans cette nouvelle section, à l'identité des unités prosodiques signifiantes et à l'analyse des informations qu'elles contribuent à transmettre. Avant toute chose, il convient de préciser que, mises à part des considérations à caractère général, il n'existe pas de consensus véritable sur les types *précis* de signification que la prosodie est susceptible de véhiculer. Cela tient d'une part à des divergences théoriques accusées et, d'autre part, au caractère insuffisamment développé du cadre sémantico-pragmatique unifié qui serait nécessaire pour procéder à l'interprétation de ces types de signification (von Eusinger, 2003).

2.8.2.3.2. La conception holistique (ou gestaltiste)

En ce qui concerne la nature des unités signifiantes, il est permis de relever une opposition traditionnelle entre deux conceptions que l'on peut qualifier respectivement par les termes de *holistique* et de *compositionnelle*. Si nous considérons, par exemple, le cas de l'intonation, il apparaît, au regard de la conception holistique radicale (d'inspiration gestaltiste), que le *contour intonatif global,* qui

correspond généralement à ce que l'on appelle Unité Intonative, Groupe Intonatif ou Syntagme Intonatif, exprime une signification qui ne peut faire l'objet d'une décomposition en éléments signifiants plus petits. Le contour possède ainsi un statut similaire à celui d'un *morphème*, ce qui explique le choix du terme de « morphème intonatif » que l'on utilise pour nommer cette entité signifiante. Il est alors usuel de parler de morphème conclusif, continuatif, de question, etc. (Delattre, 1966 ; Rossi, 1999). Pour les tenants de l'approche holistique, il existe deux manières de concevoir la relation entre l'expression et le contenu. La première considère que le morphème est caractérisé formellement par des traits de niveau et/ou des traits de configuration distinctifs qui le définissent globalement (par exemple, le morphème de *continuation majeure* /CM/ est défini par Delattre par les traits de niveau 2-4 et le trait de configuration convexe, par comparaison à l'intonème de *question* /Q/ qui possède les mêmes traits de niveau, mais qui se distingue de /CM/ par le trait de configuration concave). Selon la seconde conception, le morphème est constitué, au plan formel, d'unités plus petites : *mouvements*, pour l'école britannique et l'école hollandaise ; *niveaux* pour l'école américaine structuraliste. Les niveaux des structuralistes sont assimilables à des *phonèmes de hauteur* (« *pitch phonems* », *cf.* Wells, 1945). Il est tentant de parler ici d'une *double articulation prosodique*, tout à fait comparable à celle du *matériau verbal* en phonèmes/morphèmes. L'approche holistique est particulièrement à l'œuvre dans l'étude de la fonction *attitudinelle* de la prosodie (Wichmann, 2002) où l'on considère que la signification de l'attitude (doute, surprise, etc.) est véhiculée par l'ensemble du contour que l'on assimile alors à un lexème. C'est également le cas dans l'approche originelle de la théorie métrique de Liberman (1975), où il est stipulé que la forme intonative de l'énoncé est choisie dans un *lexique de contours* (*tunes*) nantis d'une signification et d'une structure accentuelle propres.

La particularité de l'approche holistique est que la signification du *contour global* ne peut en principe être inférée de la signification de ses *parties*. Cette conception doit être amendée si l'on admet que certaines parties du contour sont plus informatives que d'autres. C'est notamment le cas de *l'attaque*, de la *prétonique* et de la *tonique* en français (qui correspondent, respectivement, à la syllabe initiale, à la syllabe pénultième et à la syllabe finale du contour), comme l'ont montré les expériences réalisées par Autesserre & Di Cristo (1972). Puisque la signification du contour est induite d'une relation entre ces parties, et non de la globalité de sa configuration, il serait plus juste de parler dans ce cas de *morphème discontinu* (Rossi *et al.*, 1980).

2.8.2.3.3. La conception compositionnelle

L'approche compositionnelle (*compositionality of tune meaning*), telle qu'elle est illustrée par les travaux de l'école américaine « auto-segmentaliste » (Pierrehumbert, 1980 ; Pierrehumbert & Hirschberg,

1990 ; Hayes & Lahiri, 1991), considère que les entités constitutives des contours (*tunes*), telles que les accents mélodiques (*pitch accents*) et les tons de frontière (*boundary tones*) véhiculent chacune des significations particulières qui contribuent à construire la *signification générale* portée par le contour qui les intègre. Dans cette approche, qui s'inspire des conceptions exposées antérieurement par Bolinger (notamment en ce qui concerne l'interprétation sémantique des configurations accentuelles qualifiées de « *pitch obtrusions* » : Bolinger : 1957 ; 1958), les accents mélodiques et les tons de frontière sont analysés comme des morphèmes à part entière, c'est-à-dire comme des *signes prosodiques* (dotés d'un signifiant et d'un signifié) qui sont enchâssés dans des unités signifiantes plus larges. Faute de données empiriques suffisantes, il est prématuré de discuter le bien-fondé de l'approche compositionnelle. On retiendra cependant que la validité de cette approche a été récemment remise en question par Dainora (2002).

Au terme de cette rapide analyse, nous avancerons la conclusion provisoire que les deux approches ne sont pas nécessairement *mutuellement exclusives*. Il est possible d'envisager que certains effets de sens puissent résulter d'une interaction complexe entre des informations relativement indépendantes que peuvent véhiculer des configurations prosodiques *globales* et des entités prosodiques *locales*. Il ne faut également pas perdre de vue que l'interprétation du sens des messages transmis par la parole résulte le plus souvent d'un *autre type d'interaction* entre l'information apportée par le *canal prosodique* (que l'on peut évaluer à l'aide de tests de perceptions appliqués à de la parole filtrée ou à de la parole réitérée en séquences de syllabes « ma », *cf.* Liberman & Streeter, 1978) et les informations véhiculées par les *autres canaux* (verbaux et para-verbaux). Cette interaction est manifestement *fluctuante*, dans la mesure où, par exemple, l'information prosodique peut supplanter dans certains cas celle de la syntaxe et inversement, ce qui tendrait à justifier une approche interactive de la construction du sens en termes de *poids relatifs* apportés par les différentes composantes linguistiques (Blache & Di Cristo, 2002) et paralinguistiques.

2.8.2.3.4. Que signifient les unités prosodiques signifiantes ?

La problématique soulevée par cette question est si vaste et si peu résolue que nous ne pouvons y apporter tout au plus que des réponses fragmentaires. D'autre part, afin d'éviter une trop grande dispersion, nous nous intéresserons principalement aux informations relatives aux significations linguistiques (*linguistic meaning*) qui se rapportent à l'interprétation du contenu propositionnel des messages (y compris le contenu propositionnel enrichi par l'activation des inférences) et à celle de « *l'attitude propositionnelle* » que les locuteurs manifestent envers ces messages. Nous laisserons de côté, par conséquent, les informations extérieures à ces contenus, qui renseignent notamment sur l'identité du locuteur et qui ne possèdent pas, de ce fait, une valeur *proprement contrastive*. Il est

évident, en effet, que les caractéristiques prosodiques qui dénotent l'appartenance d'un individu à une communauté dialectale ou socioculturelle particulière ne sont pas linguistiquement contrastives pour les membres de cette communauté.

En règle générale, l'étude des nombreuses fonctions signifiantes de la prosodie n'est envisageable que dans le cadre de l'analyse du discours oral et de la conversation familière. Le choix de tels observables amène alors alors à constater que la *phrase* traditionnelle, à laquelle se réfèrent la plupart des travaux en prosodie, ne constitue pas un objet d'analyse adéquat (*cf.* Miller & Weinert, 1998) pour entreprendre l'investigation de la diversité des fonctions assumées par la prosodie. À moins d'admettre l'existence d'une « *phrase prosodique* » (*cf.* Blanche-Benveniste, 2002) à laquelle il conviendrait dans ce cas de donner une définition formelle explicite. L'analyse des fonctions de la prosodie dans le cadre du discours et de la conversation, met en avant un phénomène dérangeant au regard des approches « phrastiques » traditionnelles (consacrées souvent à l'analyse des propriétés prosodiques de phrases lues isolées) : l'interprétation de la signification d'un patron prosodique est souvent *dépendante du contexte* (ce terme étant pris dans son acception la plus large, qui inclut les références au co-texte et au contexte situationnel ; *cf.* Aklan & Bazzanella, 2003). En effet, un même patron prosodique peut être employé dans différents contextes pour exprimer des effets de sens diversifiés, de même que différents patrons peuvent être utilisés dans différents contextes pour induire une même signification.

La dépendance du contexte est l'une des raisons (avec la construction pluripartite du sens en termes de poids relatifs évoquée plus haut) pour lesquelles il n'existe pas nécessairement de relation biunivoque entre un signifiant prosodique particulier et une signification spécifique. À titre d'exemple, la mélodie montante peut signaler bien plus que l'expression d'une question ou d'une continuation, de même que l'expression de ces contenus peut faire appel à de multiples ressources (verbales et paraverbales, linguistiques et paralinguistiques). Encore une fois, il n'existe pas de modèle théorique qui possèderait la capacité de traiter explicitement et exhaustivement cette problématique. À cet égard, il est permis de penser qu'une approche apparentée à la construction d'une *grammaire probabiliste* apporterait une contribution significative à l'édification d'un tel modèle (*cf.* la première partie de cet article). Quoi qu'il en soit et malgré les incertitudes et les zones d'ombre, nous devons nous efforcer d'établir ici une synthèse partielle des connaissances acquises sur la nature des significations véhiculées par les unités prosodiques. Afin de préserver la cohérence de l'exposé, nous réutiliserons en filigrane la distinction précédente entre approche *holistique* et *compositionnelle*. Toutefois, nous intégrerons également à la discussion les notions incontournables de domaine, de proéminence et de phrasage qui transcendent cette distinction.

Il convient de préciser que la distinction holistique/compositionnelle ne s'applique pas aux unités relevant de la prosodie lexicale. Dans ce cas particulier, en effet, l'entité prosodique désignée : l'accent ou le ton, appartient à la structure du mot au sein duquel il assume une fonction distinctive comparable à celle du phonème. De ce point de vue, il est établi que, dans une langue dite à *accentuation libre*, l'accent peut assumer cette fonction au niveau du *morphème* dont il constitue alors la marque *d'identification formelle* (Garde, 1968). Par exemple, en italien, l'opposition CONtino/ conTIno signifie que la suite *ino* doit être interprétée comme un diminutif : *le petit conte*, si la pénultième syllabe du mot est accentuée (les diminutifs étant *accentogènes* en italien) et comme une désinence verbale (non-accentogène) : *qu'ils comptent*, si c'est la première syllabe du mot qui bénéficie de cette propriété.

Dans le champ d'étude de la prosodie post-lexicale, Bolinger (1958) a développé une théorie holistique de l'accentuation selon laquelle les configurations des *accents mélodiques* (« *pitch accents* ») sont interprétées comme des *morphèmes*. D'après cet auteur, le système accentuel de l'anglo-américain comporte trois *configurations accentuelles* de base (identifiées par A, B, C). Ces formes accentuelles expriment des fonctions qui font référence à des actes de parole (*speech acts*) ou à des opérations sémantico-pragmatiques réalisées sur des contenus propositionnels larges concernant l'assertion, la contre-assertion, la connexité, l'inachèvement, etc. Pour Bolinger, la fonction de base de l'accentuation est une fonction de *mise en relief* (*hightlighting*) ou de *focalisation*. Selon cet auteur, la prosodie accentuelle (et intonative) est fondamentalement *iconique*, dans la mesure où il existe une corrélation entre la force de la *proéminence phonologique* et celle de la *proéminence informative*. D'autre part, toujours selon Bolinger, la distribution des formes accentuelles liées à la focalisation n'est pas systématiquement contrainte par la syntaxe, contrairement aux prédictions de la *règle d'accentuation nucléaire* (*Nuclear Stress Rule*), formulée par Chomsky & Halle, (1968) : mais dépend en fait des choix du locuteur, qui sont *pragmatiquement motivés*. De ce fait, selon la formule chère à Bolinger : « *l'accent n'est pas prévisible, à moins d'être télépathe* » (« *Accent is predictable if you are a mind-reader* », Bolinger, 1972).

Bien que le concept *d'accent mélodique* (« *pitch accent* ») soit accepté par plusieurs auteurs, ces derniers adoptent cependant des points de vue différents sur *l'interprétation de sa fonctionnalité*. À la différence de Bolinger, Gussenhoven (1984) estime que les opérations sémantiques qu'actualisent les accents mélodiques n'opèrent pas nécessairement sur des contenus propositionnels larges, mais sur des entités plus petites comme l'argument et son référent discursif (*cf.* von Eusinger, 2003).

Mises à part ces divergences théoriques, il existe en revanche une forme de consensus sur la nature des relations de l'accent au *focus* (« *FTA*: *focus to accent theory* » : Ladd, 1996). Par son placement, l'accent ne véhicule pas un contenu propositionnel propre, mais indique une *relation de saillance/arrière-plan* qui contribue à signaler le domaine ou la portée de la focalisation. La

conception traditionnelle distingue habituellement entre la *focalisation large*, qui peut concerner la totalité de l'énoncé et la *focalisation étroite*, qui ne s'applique qu'à une partie de ce dernier. La focalisation étroite est souvent interprétée comme une focalisation de *contraste* dont le domaine minimal est le morphème. La focalisation de contraste peut *activer des inférences* qui peuvent affecter *les implicatures conversationnelles* de l'ensemble de l'énoncé, comme le montrent les exemples suivants :

(1) Je ne critique pas parce que tu es mon PERE.
(2) a - Pourquoi n'invites-tu pas Jean ?
b- Je n'aime PAS les radins.

Dans l'exemple (1), l'accentuation contrastive de « *père* » induit l'interprétation : « *ce n'est pas parce que tu es mon père que je te critique* », alors que l'absence d'accentuation contrastive de cet item lexical favorise l'interprétation : « *tu es mon père et donc je ne te critique pas* ». Dans l'exemple (2), la réponse donnée en (b) ne répond pas directement à la question posée en (a). Cependant, l'accentuation contrastive de « pas » et la réduction prosodique du complément « *les radins* » (réalisé avec une mélodie plate, équivalente à l'intonation dite de *parenthèse basse*) permet à (a) d'inférer d'une part que (b) n'invitera pas Jean et, d'autre part, que ce dernier est avare. Ces exemples sont particulièrement intéressants car ils illustrent une contribution possible de la prosodie à la Théorie de la Pertinence de Sperber & Wilson (1986), alors que ces auteurs affirment que leur théorie n'est pas compatible avec l'analyse linguistique.

L'approche holistique de la signification prosodique « *prosodic meaning* » concerne principalement *l'interprétation de l'intonation* « *intonation meaning* ». Cette approche s'appuie notamment sur l'idée fort ancienne que les *unités intonatives* (formellement identifiables par des marques de frontière et de cohérence mélodique internes) délimitent des *groupes de sens*. Ce point de vue a été réactualisé par Selkirk (1984 : 286) sous la forme de la « *condition de groupement sémantique* » (*sense-unit condition*) qui contraint, selon cet auteur, le *phrasage intonatif*. L'adoption consensuelle de cette condition soulève cependant le problème de la nature de l'information véhiculée par *l'Unité Intonative* (ou le *morphème intonatif* qu'elle représente). La tradition attribue le plus souvent à cette unité à la fois une signification relativement vague, comme le signalement de la continuité ou de la clôture (qui relève d'une forme de ponctuation de l'oral), et la capacité d'exprimer un faisceau de significations qui informent sur la force illocutoire des actes de parole et sur l'attitude des locuteurs qui profèrent ces actes. La complexité des relations qui résultent de la diversité de ces niveaux de signification a conduit certains chercheurs à proposer une *typologie* élémentaire des unités intonatives signifiantes qui est fondée essentiellement sur le critère de la *fréquence d'emploi*. C'est ainsi que la typologie proposée par Delattre (1966) - encore largement utilisée de nos jours pour la description du français - est limitée aux dix

« *intonations de base* » suivantes : Continuité mineure, Continuité Majeure, Finalité, Question, Interrogation, Ordre, Exclamation, Parenthèse haute, Parenthèse basse, Implication.

Il n'est pas évident que ces intonations de base, qualifiées *d'intonèmes* par Delattre, doivent être considérées comme des entités phonologiques à part entière, dans la mesure où la distinction continuation mineure/continuation majeure repose sur des différences de degré ; où la question (totale) et l'interrogation (partielle) possèdent des marques morpho-syntaxiques différentes ; où la parenthèse haute et la parenthèse basse n'apparaissent pas dans un même paradigme et où l'exclamation et l'implication ne sont peut-être que des variantes emphatiques d'une même unité phonologique. Le débat sur ce point reste ouvert. On se reportera aux travaux récents de Post (2000) et de Portes (2004), qui représentent des contributions significatives à cette discussion, grâce à l'apport de données empiriques et de réflexions théoriques nouvelles.

L'Unité Intonative, que l'on suppose être l'objet d'une planification précoce dans les étapes de production de la parole (Boomer, 1965 ; Shattuck-Hufnagel, 2000), peut apporter, par le *balisage de son domaine*, une contribution décisive à l'interprétation sémantique des énoncés, en signalant à elle seule comment les éléments verbaux qui les constituent doivent être regroupés. Cette fonction prééminente de *phrasage* (*cf.* Sanderman, 1996) se manifeste notamment dans la résolution des ambiguïtés d'attachement syntaxique. C'est ainsi que l'interprétation sémantique de l'exemple suivant (extrait de la présentation d'un bulletin d'information radiophonique) :

(3) Deux policiers sont sous les verrous depuis trois mois ils organisaient un réseau de voitures volées,

dépend de la façon dont la limite de l'Unité Intonative indique le rattachement du circonstant « *depuis trois mois* » à la première ou à la seconde proposition. Il serait possible de multiplier les exemples de ce type qui attestent du rôle de la prosodie dans l'indication de la structure syntaxique qui est livrée à l'interprétation sémantique. Dans ce cas de figure, l'Unité Intonative n'est pas signifiante par elle-même, mais dans sa relation à l'organisation syntaxique ou macro-syntaxique Il ne faudrait pas en conclure hâtivement (comme c'est le cas encore dans de nombreux travaux) que la fonction majeure de la prosodie est d'être un indicateur syntaxique. Il est manifeste que la syntaxe exerce des contraintes sur le *phrasage prosodique* (qui constitue, comme on le verra dans la dernière partie de l'article, l'une des facettes de la phonologie prosodique). Cependant, l'étude de l'interface syntaxe/prosodie ne permet de rendre compte que d'une part assez réduite des nombreuses fonctions assumées par la prosodie. Parmi ces dernières, il importe de souligner particulièrement le rôle des unités intonatives (et de l'accentuation) pour le signalement de la structure informationnelle des messages, dans la quadruple perspective de la distinction : connu/nouveau, thème/rhème,

topique commentaire et focus/arrière plan (Rossi, 1999). Quoi qu'il en soit, il est souhaitable d'éviter le piège de la circularité dans lequel sont souvent tombés les structuralistes et les générativistes. Piège qui a consisté, comme le remarque Bolinger (1972), à *fabriquer* des règles prosodiques pour interpréter la syntaxe ou des règles syntaxiques pour interpréter la prosodie, en amalgamant dans les deux cas des domaines qui doivent être traités séparément.

L'interprétation des fonctions de la prosodie dans le cadre de l'analyse du discours fait apparaître que les significations véhiculées par les unités intonatives dérivent également des relations syntagmatiques qu'elles contractent. Cela se vérifie à la fois au niveau de l'énoncé et des complexes d'énoncés. Dans le premier cas, il est loisible de montrer, notamment, que la relation qui s'instaure entre deux unités intonatives adjacentes permet d'établir des liens de causalité réversibles entre les entités verbales qui leurs sont associées dans la construction de l'énoncé. Par exemple, dans l'énoncé :

(4) Marie est infidèle, Pierre boit,

la séquence prosodique: [Implication + Parenthèse basse] permet d'inférer la signification que « *Marie est infidèle parce que Pierre boit* », tandis que la séquence [Parenthèse basse + Implication] est interprétable comme : « *Pierre boit parce que Marie est infidèle* ».

En ce qui concerne les complexes d'énoncés, il s'avère que les suites d'unités intonatives actualisent diverses formes de relations, notamment des *relations rhétoriques*, qui contribuent à l'interprétation du discours (Wichmann, 2000). Cette fonctionnalité plurielle est assurée par la mise en œuvre conjointe de deux catégories de marques physico-formelles : des marques prosodiques *syntagmatiques* de limites et de frontières et des marques *orthogonales* relatives à la variation de la gamme tonale et du registre (Ladd, 1996 ; Portes et Di Cristo, 2003). Le traitement linguistique de ces dernières alimente aujourd'hui de nombreuses discussions (*cf.* Ladd, 1996 et les contributions aux volumes édités par Kingston & Beckman, 1990 et Keating, 1994).

Enfin, envisagée sous l'éclairage de l'analyse du discours, l'approche holistique (qui conçoit, rappelons-le, l'Unité Intonative comme un *gestalt*), conduit à considérer que les unités intonatives se regroupent en unités de rang supérieur, telles que *la période* (Hazaël-Masieux, 1983) et *le paraton* (Fox, 1973 ; Yule, 1980), qui semblent correspondre respectivement à l'énoncé complet et au paragraphe oral, et dont la fonction principale est de participer à la structuration du discours. Il reste à démontrer dans quelle mesure et sous quelle forme ces macro-unités peuvent être intégrées à la représentation des systèmes phonologiques et quelle est leur contribution à l'analyse macro-syntaxique et sémantico-pragmatique du discours (Sabio, 1996 ; Berrendonner, 1991, 2002, Andersen & Nølke, 2002 ; Di Cristo, 2000, 2002 ; Lacheret & Victorri, 2002 ; Delais-Roussarie & Choi-Jonin, 2004).

Les remarques précédentes nous amènent à conclure momentanément que l'Unité Intonative prise *comme un tout* véhicule soit des significations directes, à la manière d'un lexème ou d'un morphème,

significations qui se rapportent à la force illocutoire des actes de paroles et des attitudes manifestées par les locuteurs, soit des significations indirectes relatives à l'empaquetage du matériau verbal qui forme le texte des énoncés et du discours. Dans tous les cas, l'information apportée par les unités intonatives interfère avec celle que véhicule le *matériau verbal*. Un cas limite est toutefois représenté par *l'expression de l'ironie* où l'information prosodique contredit celle qu'exprime le canal verbal (ex. « *ah, elle est belle ta veste !* »). Tous les contrastes manifestés par les unités intonatives ne sont donc pas de même nature, car certains sont plus enclins à véhiculer de l'information sur *l'organisation* des énoncés et l'interprétation de leur *forme logique*, alors que d'autres sont davantage associés à diverses formes d'interprétation *pragmatique* relatives à la *contextualisation* de ces énoncés.

L'approche *compositionnelle*, qui se fonde sur le formalisme du modèle de l'intonation de l'anglais élaboré par Pierrehumbert (1980) - et révisé par Beckman & Pierrehumbert (1986)- stipule que la signification de l'Unité Intonative ou du contour (*tune*) résulte de l'addition des informations apportées par *ses parties constitutives*, en l'occurrence, par les accents mélodiques locaux (*pitch accents*) et les divers tons de frontières (*boundary tones*).

C'est ainsi que dans le cadre de l'approche compositionnelle de la signification des contours mis en jeu dans le discours (« *compositional approach to tune meaning* »), qui est exposée par Pierrehumbert & Hirschberg (1990), les accents mélodiques sont présentés comme des *îlots de saillance* introduits dans une *forme logique* (Jackendoff, 1972). Leur interprétation ne fait pas référence à des actes de parole, comme chez Bolinger, mais aux savoirs et aux croyances que sont supposés partager (*shared, knowledge*, *beliefs*) les acteurs du discours. Dans cette optique cognitive, il est nécessaire de faire appel à des formulations théoriques explicites de cette croyance partagée, afin de pouvoir interpréter de façon cohérente les significations exprimées par l'intonation (*intonation meaning).* Pierrehumbert & Hirschberg (1990) se réfèrent ainsi à Clark & Marshall (1982) et à Joshi (1982) pour préciser l'interprétation qu'elles donnent des concepts de : *croyance partagée* (« *mutual belief* ») et *d'intention*. Elles utilisent notamment la notion de *croyance mutuelle mono-faciale* (« *one-sided mutual belief* ») pour désigner l'état mental des convictions de *A* sur les croyances qu'il suppose partager avec *B*. Selon cette conception, il apparaît que l'objectif majeur (ou l'intention majeure) d'un locuteur *L* est de modifier ce qu'il croit que l'interlocuteur *I* pense être mutuellement partagé par *L* et *I*. Il s'ensuit que les auteurs décrivent les usages qui sont faits des contours (*tunes*) par *L* en termes *d'intentions* à ajouter à la certitude que *L* a de ce que *I* pense être mutuellement partagé. Cependant, les auteurs n'analysent pas, dans leur étude, la façon dont les intentions sont reliées à la *structure intentionnelle* du discours, comme le font par exemple Grosz & Sidner (1986). Elles admettent, néanmoins, que les aspects relatifs à la *structure intentionnelle* du discours (*cf.* Cohen *et al*, 1990, pour des précisions sur les notions *d'intention* et de *structure intentionnelle*) sont véhiculés, à

l'instar de ceux qui se rapportent à sa *structure attentionnelle* (*attentional structure* : Grosz & Sidner, 1986), par le jeu des contours. Nous rappellerons, à ce propos, que dans l'approche compositionnelle de l'école américaine (Pierrehumbert, 1980 ; Pierrehumbert & Hirschberg, 1990), le contour (*tune*) de l'Unité Intonative ou du Syntagme Intonatif (*Intonational phrase*) peut être subdivisé en unités prosodiques plus petites appelées syntagmes intermédiaires (*Intermediate phrases*). Bien que les tons de frontière de ces deux unités contribuent à la signification générale du contour, ils véhiculent également des informations interprétables par rapport à leurs domaines phonologiques respectifs au sein de la hiérarchie prosodique. C'est ainsi que les tons de frontière des unités intonatives véhiculent des informations sur l'interprétation à donner au *syntagme en cours*, notamment sur sa valeur pro-active (*forward-looking*) dans le déroulement du discours, tandis que les tons de frontière des syntagmes intermédiaires informent davantage sur les liens qui existent entre ces syntagmes de rang inférieur. Ils indiquent notamment si le contenu propositionnel d'un syntagme particulier participe, en association avec celui que véhiculent d'autres syntagmes, à la construction d'une unité de sens plus large.

Dans une *approche compositionnelle hiérarchique* du type de celle que propose l'Ecole britannique, l'unité prosodique est formée d'une avant-tête (*prehead*), d'une tête (*head),* d'un noyau (*nucleus*) et d'une queue (*tail).* La configuration variable de ces éléments *participe à des degrés divers à la signification globale* de l'unité. L'élément minimal de cette dernière est cependant *le noyau* (*nucleus*) ou le *ton nucléaire* (*nuclear tone*) Il est ainsi permis d'établir le champ sémantique des significations de l'intonation (grammaticales et attitudinelles) à partir d'une typologie de ces tons nucléaires (Crystal, 1969). Dans une approche compositionnelle similaire exposée par Di Cristo (1976a, 1976b, 1978), l'Unité Intonative est fractionnée en éléments dénommés : *attaque*, *prétonique* et *tonique* (l'ensemble prétonique-tonique constituant une *cadence*). Cette division de l'Unité Intonative se fonde sur l'étude de Autesserre & Di Cristo (1972) dans laquelle il est montré au moyen de tests de perception que tous ces éléments participent à la signification portée par l'Unité Intonative que l'on pourrait alors qualifier de « *morphème discontinu* » (Rossi *et al.*, 1980).

À l'instar de l'approche compositionnelle adoptée par Pierrehumbert & Hirschberg, la conception holistique de l'interprétation du sens exprimé par les contours intonatifs connaît également des applications aux études du dialogue qui se réclament de cadres théoriques intégrant les concepts de savoirs et de certitudes - ou de croyances - partagés ou non-partagés. Les travaux récents de Beyssade *et al.* (2004), Marandin *et al.* (2004) participent de cette orientation. Ils se réfèrent en particulier au cadre interprétatif du dialogue développé par Ginzburg (à paraître, cité par les auteurs) qui promeut l'idée que les acteurs d'un échange conversationnel ne partagent pas toujours le même « contexte » (le terme étant pris ici dans sa signification la plus large). À la faveur de cette

hypothèse, ils soutiennent une approche *épistémique* de l'interprétation du sens véhiculé par les deux contours de base du français qu'identifient les modèles holistiques : le *contour montant* et le *contour descendant*. Selon cette interprétation, ces contours ne seraient corrélés ni à la *force illocutoire* (identification de la question, de l'assertion, etc.), ni à *l'engagement du locuteur*, mais à ce que le locuteur-ci « pense de ses croyances et de celles de son interlocuteur ». C'est ainsi que le contour descendant indiquerait que le locuteur pense que ses croyances et celles de son interlocuteur sont compatibles, tandis que le contour montant ou montant-descendant signalerait que le locuteur pense que sa vision du monde et celle de son interlocuteur peuvent être conflictuelles. Par ailleurs, le contour montant signifierait que le locuteur adopte son propre point de vue et le contour montant-descendant, celui de son interlocuteur (*cf.* Portes, 2004, pour une discussion de cette interprétation). Les perspectives dialogique et conversationnelle se réfèrent fréquemment au concept de « coénonciation » (Jeanneret, 1999), qui concerne notamment la façon dont le locuteur tient compte de la réception de son message par l'allocutaire. Cet aspect interactionnel constitue le thème central de l'approche développée par Morel & Danon-Boileau (1998) et par Morel (2000). Cette approche, qui est ciblée sur l'interprétation des aspects iconiques conventionnels de la prosodie (remarquablement mis en valeur par les travaux antérieurs de Bolinger, 1983 et de Fónagy, 1986, 1989), cherche notamment à établir des liens directs entre les plages hautes ou basses de F0 utilisées par les locuteurs et le degré de coénonciation exprimé par cette stratégie. C'est ainsi que l'emploi d'une plage basse signifierait un repli sur soi et une rupture de coénonciation conséquente, tandis que le recours à une plage haute suggèrerait au contraire un « maintien sur les rails » de la coénonciation impliquant une anticipation de la pensée de l'interlocuteur. Cette approche analyse de la même manière les paramètres prosodiques de durée, d'intensité et de pause pour montrer comment leur conjugaison est interprétable dans le cadre d'une conception intégrative de l'iconocité de la prosodie.

Le vaste champ de la « rhétorique du discours », qui s'avère incontournable pour le traitement des relations complexes entre les référents discursifs, notamment pour celui des relations de cohérence et de coréférence (Hobs, 1979), n'a pas encore été investi par l'analyse prosodique, laquelle est pourtant susceptible d'apporter une contribution décisive à sa compréhension. Cette absence de collusion tient probablement au fait que les modèles et les théories de la rhétorique du discours ont été initialement conçus pour interpréter des textes écrits. C'est le cas, par exemple, de l'influente théorie de la « *structure rhétorique* » : RST (« *Rethorical Structure Theory* » de Mann & Thompson (1988) qui a suscité, cependant, une tentative intéressante pour rendre compte du « texte parlé » (Van Donzel & Koopmans-Beinum, 1995). Plus récemment, le développement de la Théorie de la Représentation Discursive Segmentée, SDRT : « *Segmented Discourse Representation*

Theory » (Asher & Lascarides, 2003) paraît offrir un cadre formel propice à l'interprétation fonctionnelle de la prosodie du discours. En effet , à la différence de la RST, qui est une théorie de l'explication de la structure cohérente du texte (et donc, d'un produit fini), la SRDT (qui est un développement de la DRT : « *Discourse Representation Theory*», élaborée par Kamp dès 1981 ; *cf.* également Kamp & Reyle, 1993), s'inscrit dans une perspective plus dynamique, en ce sens qu'elle est davantage attachée à l'interprétation des processus mis en œuvre dans la construction du discours. En bref, la SDRT présente l'avantage majeur d'intégrer les concepts fondamentaux du modèle de la « *sémantique dynamique* » conçu par Kamp à une théorie de la structure du discours où chaque « *clause-proposition* » a la faculté de remplir une ou plusieurs fonctions rhétoriques. Ces vues sont, à l'évidence, compatibles avec l'idée de la *plurifonctionnalité* de la prosodie soulignée à maintes occasions dans cet article, ainsi qu'avec notre conception de la prosodie comme un *dispositif écologique* (c'est-à-dire fondamentalement adaptatif) implémenté dans une grammaire fonctionnelle de la contextualisation plurielle (Di Cristo, 2000a).

Avant de clore cette section, il convient de mentionner d'autres modèles du discours particulièrement influents qui présentent des cadres théoriques tout à fait propices à l'interprétation de la fonctionnalité prosodique et à l'élaboration d'une *sémiotique prosodique du langage situé*. À ce propos, on peut évoquer le modèle déjà cité de Grosz & Sinder (1986) qui accorde une place prépondérante à l'interprétation des *intentions des locuteurs* et aux « *attracteurs d'attention* » comme marqueurs de *focalisation*. Ce modèle, qui est particulièrement apprécié par les linguistes américains (notamment dans le domaine du traitement automatique des langues), a servi de cadre interprétatif à l'importante étude de Pierrehumbert & Hirschberg (1990) sur l'interprétation compositionnelle de la signification des contours intonatifs de l'anglo-américain dans le discours (*cf.* supra). Pour ce qui concerne le français, la référence la plus intéressante, de notre point de vue, est le modèle genevois (Roulet *et al.*, 2001), car il est conçu pour l'analyse de données attestées et à partir de ces dernières. De surcroît, ce modèle intègre de *multiples dimensions*, telles que : la structure du texte, l'organisation informationnelle, des aspects dialogiques et interactionnels, les activités discursives, la polyphonie, les stratégies communicatives, etc. Cet éclectisme de bon aloi se prête également bien à l'interprétation de l'hétérogénéité fonctionnelle de la prosodie, comme l'attestent des travaux récents qui ont tenté de montrer comment l'apport de la prosodie permet d'enrichir (voire de modifier) la structure originelle du modèle (Simon, 2003 ; Portes, 2004).

En résumé, il est établi que les contrastes prosodiques exprimés par le jeu des contours, de la proéminence et du phrasage expriment une large gamme de fonctions (et donc, de significations) dont certaines n'ont pas encore fait l'objet d'une analyse approfondie, en raison notamment de l'absence de référence à un cadre théorique d'interprétation unifié adéquat. À notre avis, le modèle

de *l'Organisation Communicative*, élaboré par Mel'cuk (2001) paraît s'approcher d'un tel cadre, car il s'efforce d'intégrer *plusieurs niveaux* d'interprétation de la signification. Par ailleurs, il est probable que les chercheurs en prosodie auront beaucoup à gagner en s'informant des grandes approches théoriques du sens telles que : la Théorie de la Représentation du Discours (DRT) dont il a déjà été question, les théories Gricéennes et Néo-Gricéennes (Levinson, 1997), la Théorie de la Pertinence (Sperber & Wilson, 1986), les théories sémantiques développées dans la mouvance de la « Linguistique Cognitive » (Jackendoff, 1983 ; Lakoff, 1987) et la théorie de la Sémantique de situation (Barwise & Perry, 1983). Cette approche se présente essentiellement comme une *théorie relationnelle de la signification*. À titre d'exemple, Bellengier (2004) se réfère à cette théorie pour modéliser la signification des énoncés comme un ensemble de contraintes qui déterminent les relations entre le contexte d'apparition d'un énoncé particulier et son interprétation. L'outil formel utilisé ici est celui des Grammaires de Propriétés (Blache, 2001). Il semble que la sémantique des situations et le formalisme de la Grammaire de Propriétés soient en mesure de fournir des cadres d'analyse propices à l'interprétation des objets linguistiques fonctionnellement et formellement complexes comme le sont par excellence les objets prosodiques.

L'évolution des recherches prosodiques dénote une tendance à s'écarter de la conception traditionnelle des *relations bijectives* entre formes et sens (qui a cherché, sans grand succès, à définir un patron prosodique spécifique pour chaque type de signification : assertion, surprise, joie, question indirecte, etc.), au profit d'une approche plus complexe de la signification prosodique. Cette nouvelle orientation s'efforce d'établir la fonctionnalité des contrastes prosodiques à *différents niveaux* de la signification qui concernent, à la fois, la sémantique et la pragmatique des actes illocutoires et perlocutoires (*cf.* le modèle pragmatique de Beun, 1994, dans lequel l'information prosodique est cruciale pour l'interprétation des actes de parole : « *speech acts* »), l'organisation informationnelle, le phrasage de la structure du discours en segments discursifs et les relations de cohérence induites de ce phrasage, les implicatures conversationnelles, le partage des connaissances et des croyances et les relations de portée. Il est manifeste que le vaste programme de recherche que suggère cette orientation est partiellement en phase avec la définition de l'intonation proposée par Ladd (*cf. supra*) et qu'il invite, par ailleurs, à une révision des systèmes de contrastes établis antérieurement sur la base : d'un cadre interprétatif limité à la phrase, d'une distinctivité exclusivement catégorielle et de l'étude univoque de l'interface phonologie/syntaxe.

Nous considérons que les relations de la prosodie (en l'occurrence, de l'intonation) au sens sont complexes, subtiles et mettent en jeu *plusieurs types de stratégies* qui ne sont pas forcément mutuellement exclusives et que nous proposons, à titre d'hypothèse de travail, de ramener aux quatre catégories suivantes :

- *holistique* : l'information est véhiculée par un contour global qui a valeur de lexème ou de morphème (Delattre, 1966 ; Rossi, 1999). La signification du contour globalement montant, descendant, ou montant/descendant est sensible à la notion de « *révision anticipée* », mentalement représentée chez le locuteur (Beyssade *et al.* 2004) ;
- *compositionnelle* : la signification du tout (contour) est égale à la somme de celle de ses parties (Pierrehumbert & Hirschberg, 1990) ;
- *superpositionnelle* : la signification de deux contours de base amalgamés est ambivalente (par exemple, le contour montant-descendant de la demande de confirmation est interprété à la fois comme la marque d'une assertion et d'une question (Fontaney, 1991 ; Purson & Di Cristo, 1998) ;
- *interactive* : la signification émerge d'une interaction entre les diverses composantes verbales et para-verbales (prosodie, sémantique syntaxe, mimo-gestualité). Le poids relatif de ces composantes est variable en fonction de divers facteurs (Blache & Di Cristo, 2000).

Nous pensons qu'une *approche plurilinéaire* des relations de la prosodie au discours conçue dans l'esprit de celle que nous proposons, c'est-à-dire, qui tente de mettre en correspondance plusieurs niveaux de représentation de la composante textuo-discursive avec plusieurs niveaux de représentation de la composante prosodique (Di Cristo *et al.*, 2004) se prête favorablement à l'évaluation des hypothèses posées par la précédente quadripartition.

3. Survol des approches linguistiques de la prosodie

Nous présentons dans cette dernière partie un survol des différentes approches linguistiques de la prosodie qui nous paraissent être les plus représentatives (*cf.* également Rossi, 2000 ; Botinis *et al.*, 2001, pour compléter ce panorama). Nous évoquerons en premier lieu les approches qui ont précédé l'émergence, dans les années 75-85, de la phonologie non-linéaire, ou qui se démarquent de ce courant. Nous consacrerons ensuite la majeure partie de cette section à l'examen des approches de la phonologie prosodique qui procèdent de cette émergence. Ces dernières se rapportent aux trois mouvances apparentées que représentent la phonologie métrique, la phonologie autosegmentale et la phonologie des domaines (ou de la constituance prosodique).

3.1. Les approches linéaires

Afin d'éviter toute méprise, il importe de préciser que toutes les approches que nous allons passer en revue dans la section 3 adhèrent à la conception fondamentale de la linéarité de l'objectivation des structures linguistiques. L'expression « approche linéaire » se rapporte donc uniquement ici à celles qui ne particpent pas de la mouvance des approches dites « plurilinéaires » ou « multilinéaires ».

Si nous laissons provisoirement de côté l'approche ancienne (mais néanmoins féconde) des structuralistes américains (Pike, Wells, Blomfield, Trager, etc. ; *cf.* Lieberman, 1967, pour une

présentation synthétique de cette mouvance), sur laquelle nous reviendrons plus tard, quand nous évoquerons des travaux plus récents de la linguistique américaine (Liberman, Pierrehumbert, Selkirk, etc.), il apparaît que les approches européennes ressortissent à deux tendances, que nous qualifierons respectivement de « *phonético-formelle* » et de « *fonctionnaliste* ». La première expression fait référence aux approches qui s'attachent principalement à la description des aspects phoniques concrets et abstraits de la prosodie. Par fonctionnaliste, nous entendons désigner les approches qui procèdent d'une analyse fonctionnelle des unités prosodiques ou qui amalgament fonctions et formes dans la description et la représentation de ces unités. Nous traiterons uniquement, dans cette partie de notre exposé, des modèles d'analyse de l'intonation, qui ont donné lieu à des développements beaucoup plus étendus que les analyses de l'accentuation dont il sera question dans la deuxième partie de ce survol.

Tous les chercheurs s'accordent sur le fait que les *courbes de F0* contiennent des indices qui reflètent l'influence de nombreux facteurs relatifs à la modalité de l'énoncé, à sa segmentation en groupes, à l'accentuation, aux caractéristiques des unités segmentales (effets microprosodiques intrinsèques et co-intrinsèques), etc. Cependant, certains chercheurs décrivent ces indices en termes de *séquences d'événements* et d'autres les interprètent comme des phénomènes *superpositionnels* (Ladd, 1995).

3.1.1. Les approches non-superpositionnelles

Parmi les approches qui sont extérieures à la *mouvance autosegmentale* et qui relèvent du premier type (non-superpositionnel) de description, il convient de citer en premier lieu le modèle de l'IPO (*Instituut for Perceptie Onderzoek*), qui a été développé à partir des années 60 au sein de l'Institut des Recherches sur la Perception d'Eindhoven ('t Hart *et al.,* 1990). Ce modèle, qui a d'abord été élaboré pour décrire l'intonation du néerlandais, mais qui a fait ensuite l'objet d'applications à d'autres langues, propose de décrire les contours ou patrons intonatifs en termes de séquences de *mouvements mélodiques*. Ces derniers, qui sont modélisés sur la base d'une distinction fondamentale entre deux *niveaux relatifs*, haut et bas, sont représentés par un contour montant-descendant qui est aligné soit par rapport à la syllabe accentuée du mot, soit par rapport à un domaine plus large constitué de plusieurs syllabes. Mis à part ces mouvements, les autres parties du contour sont assimilées à des segments de droite matérialisant une ligne supérieure et une ligne inférieure qui délimitent ainsi la dynamique de la variation mélodique du patron. Parmi les contours construits par les mouvements mélodiques figurent deux *patrons prototypiques* appelés respectivement : « chapeau pointu » (*pointed hat*) et « chapeau plat » (*flat hat*) (figure 6a.1). Les descriptions présentées dans le cadre du modèle IPO sont phonologiques, dans le sens où elles représentent des séquences d'entités *catégoriellement distinctes* attachées à des points d'ancrage spécifiques de la chaîne sonore. Elles satisfont également à la « *condition de phonologicité* » telle qu'elle est formulée par

Ladd (« *...une description phonologique complète ne doit pas se résoudre à des formules abstraites, mais elle doit aussi spécifier comment ces formules abstraites sont réalisées, c'est-à-dire qu'elle doit décrire comment les entités phonologiques catégorielles se projettent sur les paramètres acoustiques continus* » (Ladd, 1996 : 16). En effet, la description abstraite des patrons décrits par le modèle de l'IPO est mise en relation avec des données phonétiques précises se rapportant essentiellement à la perception et à l'acoustique et l'ensemble du modèle peut être mis en œuvre dans une application à la *synthèse par règles* de la prosodie.

La philosophie du modèle de l'IPO a été appliquée avec succès à la description de l'intonation de la phrase française lue, par Vaissière (1975). Selon cette analyse, la construction des contours intonatifs des phrases lues met en jeu l'emploi de quatre patrons élémentaires : P1, P2, P3 et P4 configurés au moyen de trois types de mouvements mélodiques qualifiés de chute, de montée et de plateau. Ces patrons sont dérivés d'une stylisation visuelle des courbes de F0 qui consiste à abstraire de ces variables continues des *entités catégorielles* en termes de patrons et de mouvements. Il est légitime de qualifier cette description de *pré-phonologique*, dans la mesure où elle opère sur des entités relativement abstraites, mais où il n'est pas établi, toutefois, que les quatre patrons entrent dans un paradigme de contrastes distinctifs.

Plus récemment, Taylor (1994) a proposé un modèle de *stylisation phonétique* de l'intonation qui présente des parentés avec les précédentes approches. Ce modèle, connu sous la référence de « *Tilt Intonation Model* » analyse l'intonation comme une séquence « d'événements » tels que les accents mélodiques et les tons de frontière. Cependant, ces événements ne sont pas définis par référence à des *entités catégorielles*, mais à des *propriétés phonétiques continues*. C'est ainsi que chaque événement est constituée d'une composante montante et descendante de dimension variable. Ce modèle, qui ne prétend pas être phonologique, a été conçu en réalité pour servir de cadre à *l'analyse automatique* et à la *synthèse vocale* des configurations de F0.

Le modèle de l'école britannique, dont l'origine remonte aux années 20 (Klinghart & Klem, 1920 ; Palmer, 1922 ; *cf.* Hirst & Di Cristo, 1998 et les références afférentes), est demeuré le cadre dominant de la description de l'intonation anglaise pendant près d'un siècle (figure 6a.2). Selon ce modèle, le contour intonatif abstrait (*tonal group*, *tone group*, *tone unit*, *intonation group*, selon les auteurs) est divisé en plusieurs parties (ou constituants) dénommées *prehead*, *head*, *nucleus* et *tail*. La tête (*head)* correspond à la première syllabe accentuée du contour et le noyau (*nucleus*), à la syllabe la plus proéminente, qui se trouve être généralement la dernière syllabe accentuée du contour. La queue (*tail)* désigne la (ou les) syllabe(s) inaccentuée(s) qui succèdent au noyau et dont la configuration est fortement influencée par celle de ce dernier. De tous les éléments constitutifs du *tone unit*, seul le *noyau est obligatoire* et constitue ainsi l'unique entité prosodique d'un énoncé monosyllabique. Les autres constituants sont optionnels et participent de diverses combinaisons

possibles avec le noyau. Cette variabilité combinatoire est enrichie par la diversité des formes que peuvent prendre les constituants du tone unit. L'avant-tête peut être haute ou basse ; la tête : basse, rehaussée (*stepping head*) ou glissante (*sliding head*). Pour sa part, le noyau est représenté par un ensemble de tons statiques ou cinétiques distinctifs , réalisables phonétiquement sous l'aspect de différents *allotones* (l'inventaire des tons cinétiques distinctifs comprend notamment : la chute basse (low fall), la chute haute (high fall, le ton montant-descendant (rise-fall), etc. (pour une description précise, *cf.* O'Conor & Arnold, 1961 ; Crystal, 1969 ; Cruttenden, 1997).

Halliday (1967) a exposé un autre type d'approche structurale de la prosodie de l'anglais, fondée sur la hiérarchie suivante : phonème < syllabe < Pied < tone group. Il s'agit d'une hiérarchie stricte (chaque niveau de constituance ne peut être subdivisé qu'en constituants de rang immédiatement inférieur) qui préfigure celle de la phonologie des domaines (Nespor & Vogel, 1986) dont il sera question plus loin.

Dans le même esprit que celui de l'école britannique, une approche *structurale linéaire* de la constituance prosodique a été présentée par Di Cristo (1976a) dans le cadre d'une application à l'étude du français. Cette approche, qui s'inspire des travaux de Pike (1945) sur l'anglo-américain consiste à fractionner un groupe prosodique majeur (ou une Unité Intonative) en différents éléments abstraits pouvant faire l'objet d'observations indépendantes. C'est ainsi que l'Unité Intonative peut être subdivisée en *précontour* et *contour* (figure 6a.3). Le *contour* est représenté par la configuration mélodique signifiante qui est associée à la dernière syllabe de l'Unité Intonative (la syllabe *tonique* en français), tandis que le précontour fait référence à la configuration comprise entre *l'attaque* et la *prétonique* du groupe prosodique. Afin de rendre compte de l'articulation du précontour et du contour et des effets de sens que cette relation engendre (Autesserre & Di Cristo, 1972), il est proposé d'appeler *cadence* l'ensemble prétonique/contour. Des études empiriques impliquant des expériences de perception montrent que les variations de la configuration du contour sont interprétables en termes de contrastes discrets (Di Cristo, 1971). D'autre part, chaque type de contour peut être décrit sur la base d'un ensemble de traits distinctifs fondés sur des oppositions privatives du type [± Haut, ± Mélodique, ± Retardé, ± Long, ± Intense, ± Atténué] (Di Cristo, 1978). Enfin, l'application de cette approche à la description de l'intonation du français, notamment à l'étude du vocatif (Di Cristo, 1976b), nous a conduit à formuler l'hypothèse que l'Unité Intonative est *encodée* en termes de *relais* ou de *points-clés syntagmatiques*, tels que l'attaque, la prétonique et la tonique, sur lesquels se fonde également l'auditeur pour procéder à l'identification de cette unité.

Le français a également servi de point de départ à la conception du modèle de *représentation linéaire hiérarchique* de l'intonation édifié par Martin (1977a, b). Le modèle phonologique de Martin est constitué, dans son essence, par une représentation séquentielle de contours qui procèdent d'une

hiérarchie prosodique (figure 6b.7.) Ces contours, dont la taille équivaut à l'unité prosodique minimale que constitue, selon Martin, le groupe accentuel (c'est-à-dire, une séquence ne comportant qu'un seul accent), sont définis au moyen d'un ensemble de *traits binaires* qui se distinguent sur la base de contrastes mélodiques de direction [+ Montant/- Montant], de durée [+ Long/- Long] et d'amplitude [+ Ample/-Ample]. Le modèle élaboré par Martin se présente comme une approche doublement intégrative de la prosodie, dans la mesure ou, d'une part, le contour et le groupe accentuel sont des *unités isomorphes* et où, d'autre part, les contours contractent des relations de présupposition réciproque et unilatérale motivées par la *sémantique* (aspect *phonosémantique*) et par la *syntaxe* (aspect *phonosyntaxique*).

Si nous poursuivons la revue des approches linéaires de la prosodie, nous devons faire une place à part à celles qui analysent l'intonation comme des séquences de *morphèmes*, qualifiés parfois *d'intonèmes* (Delattre, 1966) et définis par des termes qui évoquent généralement les fonctions que ces unités sont susceptibles d'assumer (continuatif, conclusif, interrogatif, etc.). La typologie amalgame parfois fonction et forme, dans la mesure où les termes de *parenthèse haute* et de *parenthèse basse*, par exemple, semblent faire davantage référence à une représentation de la forme qu'à un élément de contenu (6b.2). En ce qui concerne le français, cette approche a été particulièrement mise en valeur par Delattre (1966) et par Rossi (1999). *L'approche morphologique* de l'intonation se caractérise par le fait qu'elle prête une attention particulière aux relations entre l'expression et le contenu, à la différence des *approches phonologiques* (ou *phono-phonologiques*) évoquées précédemment, qui s'attachent surtout (voire exclusivement) à l'étude de la substance et de la forme de l'expression. Cependant, toutes les approches morphologiques ne sont pas à mettre sur un même plan. Pour certaines les morphèmes sont liés à la *signification pragmatique*, alors que pour d'autres, ils représentent la sortie d'une hiérarchie impliquant *plusieurs modules linguistiques*. Dans le second cas, l'approche morphologique se présente comme une théorie hiérarchique, dans le sens où la structure intonative reçoit des modules syntaxiques et sémantiques l'information nécessaire à son interprétation (Rossi, 1999). En ce qui concerne la forme de l'expression, celle-ci est représentée soit par des *traits distinctifs* (Delattre, 1966) ou par des *intonèmes* (Rossi, 1999), soit encore par des *gestalts* (Aubergé, 1992). Le lecteur pourra se reporter à Rossi (1999, 2000) pour une présentation raisonnée de ces différentes approches.

3.1.2. Les approches superpositionnelles

Les approches dites *superpositionnelles* (*overlay*) possèdent la particularité de concevoir le contour global de l'énoncé ou de la phrase comme le produit de l'interaction de deux ou plusieurs composantes superposées. Il est possible de distinguer deux types d'approches superpositionnelles,

suivant que la superposition se rapporte à l'interaction de contraintes linguistiques et non-linguistiques (dans le sens de non-intentionnelles) ou uniquement à l'interaction de contraintes linguistiquement motivées. Le premier type concerne notamment la superposition d'une *composante microprosodique* à une *composante macroprosodique*. Etant admis que les effets microprosodiques qui résultent des contraintes de production (*cf. supra*) sont relativement stables, il est possible de les dissocier de la composante macroprosodique (qui reflète les variations de F0 intentionnellement motivées) au moyen d'une méthode de calcul approprié. C'est ainsi que l'usage de l'algorithme MOMEL (Hirst *et al.*, 2000) permet d'extraire d'une courbe brute de F0 (comme nous l'avons déjà mentionné plus haut) deux profils distincts: le *profil segmental* qui ressortit à la composante microprosodique, et le *profil suprasegmental* qui ressortit à composante macroprosodique (Di Cristo & Hirst, 1986). Il est clair que cette approche ne constitue pas une approche phonologique stricte. Elle apporte cependant une contribution à cette dernière, dans la mesure où elle permet d'extraire du signal des informations que l'analyse phonologique peut interpréter directement.

Le second type d'approche superpositionnelle (fondée sur l'interaction de composantes *linguistiquement motivées*) est particulièrement bien illustré par le modèle de Fujisaki (Fujisaki & Nagashima, 1969), qui a été mis en œuvre dans de nombreux systèmes de synthèse vocale. Ce modèle ne constitue pas non plus une approche phonologique, mais un *modèle de production* de F0 dont la sortie peut également faire l'objet d'une interprétation phonologique. Il est à noter, en outre, que ce modèle se fonde sur l'hypothèse d'un contrôle des variations de F0 régulé par deux commandes discrètes impliquant l'activité du muscle cricothyroïdien. Dans le modèle de Fujisaki, la superposition met en œuvre deux composantes (figure 6a.4) : une composante *syntagmatique* (*phrase component*) et une composante *accentuelle* (*accent component*). La composante syntagmatique est modélisée sous la forme d'une réponse impulsionnelle se traduisant par une montée initiale rapide vers un pic, suivie d'une chute exponentielle (qui simule un effet de déclinaison). La composante accentuelle se présente comme une fonction d'incrémentation ou de décroissance par pas (*step function*) qui simule les montées et les chutes de F0 alignées avec les syllabes accentuées.

À côté du modèle de Fujisaki, qui bénéficie d'une formalisation mathématique aboutie (ce qui en fait un modèle bien adapté à des applications technologiques), il existe d'autres modèles superpositionnels d'inspiration plus linguistique. Parmi ces derniers, il convient de citer ici en particulier le *modèle de Garding* (Garding, 1983) qui a été appliqué notamment au suédois et au français et le modèle de Thorsen, alias Grønnum (Thorsen, 1983) qui s'applique uniquement au danois. Ces modèles ont en commun, à l'instar du modèle de Fujisaki, de concevoir l'intonation comme la superposition de *phénomènes prosodiques locaux* sur une *composante globale*. Cependant, la nature de ces composantes diffère sensiblement selon le modèle. Dans le modèle de Garding, la

composante globale est composée d'une *grille* formée par des lignes qui relient les maxima et les minima de la courbe de F0, tandis que la composante locale est formée d'événements tels que les *turning-points* et les *pivots* et qui se rapportent, respectivement, aux points de changement de la courbe de F0 et aux changements de direction de la grille (figure 6a.5). Dans le modèle de Thorsen (figure 6a.6), la composante globale (*sentence intonation component*) est modélisée par une ligne de pente variable sur laquelle se superposent des séquences de groupes accentuels (*stress group*) qui constituent la composante locale du modèle. La pente de la ligne permet de distinguer notamment l'énoncé déclaratif (pente fortement déclinante) de la question (pente nulle). Ces modèles superpositionnels se situent à mi-chemin entre l'analyse phonétique et l'analyse phonologique, dans la mesure où ils décrivent des données quantifiées et se prêtent à une représentation de la prosodie en termes d'entités catégorielles. Faute de place, nous ne pouvons procéder à une analyse détaillée des propriétés respectives de ces modèles. Pour obtenir davantage d'information, on pourra se reporter à Beckman (1995), Ladd (1995) et Grønnum (1995).

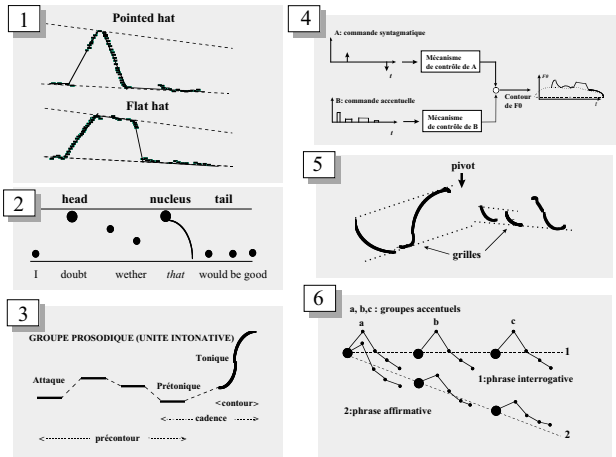

Figure 6a
Illustration de différentes approches non-superpositionnelles (colonne de gauche) et superpositionnelles (colonne de droite)

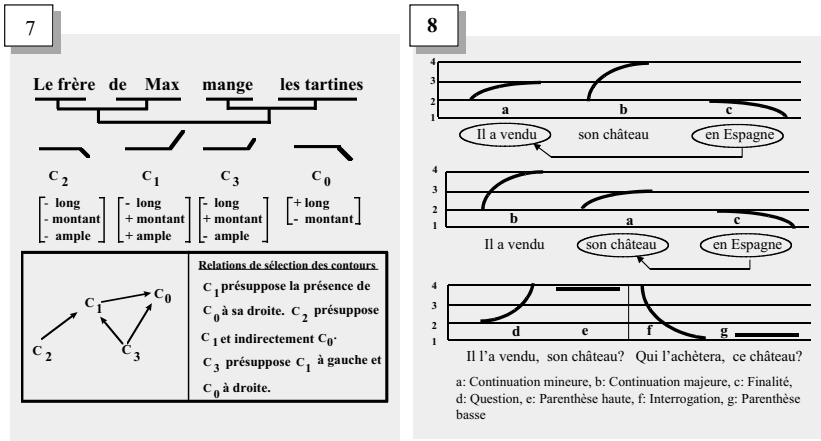

Figure 6b
Illustration des approches non-superpositionnelles (suite) : l'approche hiérarchique de Martin (7) et l'approche morphologique de Delattre (8).
Dans l'exemple 8, les flèches indiquent « l'attachement » du syntagme propositionnel « en Espagne ».

3.2. Les approches plurilinéaires

3.2.1. Remarques générales

Les approches dont il va être question dans la dernière partie de cet article sont apparentées au paradigme de la *phonologie* dite *non-linéaire*. Il peut arriver que cette dénomination induise en erreur le non-spécialiste. C'est pourquoi il importe de préciser avant toute chose que l'approche non-linéaire ne remet pas en question le principe de la linéarité de la chaîne des événements linguistiques, mais qu'elle envisage, en revanche, la représentation de cette chaîne sous l'aspect de plusieurs lignes, rangées, ou strates (*tiers*) superposées à la manière d'une partition d'orchestre. Selon cette conception, la *synchronisation* entre les différentes *lignes-instruments* est aussi importante que la notation dévolue à chacune des *lignes* particulières de la partition. Dans ces perspectives, l'emploi des qualificatifs « *plurilinéaire* » ou « *pluridimensionnelle* » (Durand, 1993) s'avère plus judicieux pour signaler de façon non ambiguë qu'ils ne s'opposent pas à la dénomination de « *linéaire* », mais à celle de « *unilinéaire* ».

Les chercheurs qui reconnaissent leur appartenance à la phonologie plurilinéaire ou, plus généralement, à la « *phonologie néo-générative* » (qui s'est construite sur les ruines du paradigme de la phonologie générative classique de Chomsky & Halle (1968) - désormais considéré par sa filiation comme un modèle périmé) se répartissent selon deux grandes mouvances que l'on peut qualifier

de *phonologie métrique-autosegmentale* et de *phonologie des domaines* ou de la *constituance prosodique*. Ces mouvances, dont l'émergence conjuguée prend forme dans les années 70-80, partagent un ensemble de conceptions que nous proposons de résumer brièvement dans les lignes qui suivent.

La phonologie plurilinéaire est à plus d'un titre (mais pas exclusivement) une *phonologie postlexicale* ou « *phrasale* » (*phrasal phonology*) qui s'intéresse aux événements qui transgressent les limites de mots (Selkirk, 1972). Sur ce point particulier, il n'y a pas de rupture avec les options du paradigme génératif classique définies il y a une cinquantaine d'années par Chomsky, Halle & Lukoff : « *Nous limiterons notre analyse à un domaine que nous appelons la proposition phonologique* » (Chomsky *et al.*, 1956 : 70). Il est vrai, toutefois, que la perspective postlexicale a été davantage exploitée dans la phonologie des domaines et dans la phonologie intonative que dans la phonologie métrique pour laquelle l'accentuation du mot a constitué pendant longtemps un pôle d'intérêt majeur. Il n'en demeure pas moins que les propositions formelles de la phonologie plurilinéaire constituent un apport significatif par rapport à la lourdeur de l'appareil qui était mis en œuvre par Chomsky & Halle (1968) pour rendre compte notamment de l'accentuation post-lexicale de l'anglais. Sans doute convient-il de préciser ici que l'intérêt qu'a manifesté la phonologie pour l'investigation du niveau postlexical n'est pas né avec les travaux de Chomsky *et al.* Il importe en effet de rappeler le rôle joué auparavant par le Cercle de Prague, en particulier par Karcevskij (1931), à qui nous devons notamment une définition - assez « *aventureuse* » pour l'époque - de la phrase : « La phrase est une unité de communication actualisée. Elle n'a pas de structure grammaticale propre. Mais elle possède une structure phonique particulière [c'est nous qui soulignons] qui est son *intonation.* » (Karcevskij, 1934 : 190)

La phonologie plurilinéaire prône *l'autonomie des représentations prosodiques* au regard de celles des phonèmes. C'est ainsi que les entités primitives de la représentation prosodique sont envisagées comme des *autosegments*, ce qui dénote une véritable rupture avec le modèle de la phonologie générative classique (Chomsky & Halle, 1968). En effet, dans ce modèle, les traits prosodiques de configuration tonale et d'accentuation se trouvent être enfermés dans des matrices qui se révèlent être *hétérogènes*, car elles incluent également la spécification des traits distinctifs des phonèmes, comme les traits : *haut*, *nasal coronal*, etc. On pourra faire remarquer que des représentations plurilinéaires étaient déjà à l'œuvre dans les notations classiques de la prosodie tonale. Cependant cette pratique ne pouvait masquer le fait que les tons étaient alors assimilés à des *signes diacritiques* associés à des voyelles ou à des syllabes et ne constituaient donc pas des entités autonomes. Erigée en principe, l'autonomie des représentations prosodiques soulève une problématique cruciale pour l'approche plurilinéaire, sur laquelle nous reviendrons ultérieurement. Elle se rapporte à la question

de *l'association* et de *l'alignement* des structures prosodiques (métrique, intonative) avec la composante verbale.

Les approches plurilinéaires introduisent dans les représentations phonologiques des aspects concernant la *temporalité* des événements (*cf. infra* : la grille métrique) et la *valeur relative* des entités contrastives qui participent de ces représentations. Comme nous le verrons plus loin, les étiquettes *w (weak)* et *s (strong)* de la représentation métrique, et les étiquettes L *(Low)* et H *(High)* de la repr é-sentation de l'intonation, désignent en réalité des *entités catégorielles* dont la valeur est établie en fonction des *relations syntagmatiques* qu'elles contractent dans la chaîne.

Nous avons utilisé, à l'exemple de Ladd (1996), l'expression « *métrique-autosegmentale* » pour désigner une approche qui se caractérise par l'intégration de la théorie métrique et de la théorie de l'intonation dans une même composante de la Grammaire. Cet amalgame n'est pas dénué de fondement et il existe plus de similitudes que de différences entre les conceptions théoriques et méthodologiques auxquelles renvoient les deux termes de cette expression (Leben, 1982). Cependant, il est nécessaire de distinguer (McCarthy, 1982), du moins dans un premier temps, l'approche métrique, qui est regardée surtout comme une théorie de la *représentation de l'accentuation* et de *l'organisation rythmique* (mais qui peut intégrer également des aspects relatifs à l'organisation de l'intonation : *cf.* Liberman, 1975) et l'approche autosegmentale, qui prétend être une théorie générale de la *géométrie des représentations* phonétiques et phonologiques (Goldsmith, 1976a) et qui, de ce fait, s'applique à la fois à la représentation des *formes sonores segmentales* et *suprasegmentales*. En ce qui concerne les premières, la phonologie autosegmentale peut être envisagée, dans ses aspects les plus concrets, comme la théorie qui rend compte de la *coordination temporelle* des effecteurs de l'appareil articulatoire (langue, lèvres, larynx). À un niveau plus abstrait, elle propose de substituer à la représentation traditionnelle des séquences de phonèmes assortis de leurs matrices compactes de traits distinctifs, un formalisme dans lequel les différents traits sont représentés sur des lignes autonomes qualifiées de *rangées autosegmentales* (*autosegmental tiers*). En ce qui concerne *la prosodie,* la théorie autosegmentale vise à modéliser les dispositifs qui supervisent la *synchronisation* des autosegments appartenant aux différentes strates de la représentation prosodique et leur association avec le matériau verbal soit directement (par ex. attachement de tons à des syllabes), soit par l'intermédiaire d'une organisation de la structure prosodique en *domaines*. Nous proposons d'examiner à part, dans la dernière section de cette partie de l'article, la *théorie des domaines prosodiques*, qui est le plus souvent désignée (d'une manière un peu excessive) par les termes de « *phonologie prosodique* » (Nespor & Vogel, 1986) et qui se présente comme une théorie de la *constituance prosodique* de l'énoncé.

3.2.2. La théorie métrique

3.2.2.1. Présentation générale

La théorie métrique a été élaborée à l'origine par Liberman (1975) et par Liberman & Prince (1977) dans le but de décrire et de représenter formellement *l'organisation des proéminences prosodiques* de l'anglais dans un cadre structurel supérieur au mot. Au cours de son évolution, cette théorie a connu des avatars et de nombreux raffinements dont les étapes les plus significatives sont associées aux travaux de Hayes (1980), Prince (1983), Halle & Vergnaud (1987), Halle & Kenstowicz (1991) et Hayes (1991). Nous renonçons à rapporter dans le détail ces développements qui ont conduit, en définitive, à la construction d'une « machinerie » assez élaborée. Nous préférons renvoyer le lecteur aux synthèses régulièrement effectuées par plusieurs auteurs, notamment par van der Hulst & Smith (1982), Anderson (1984), Beckman (1986), Goldsmith (1990), Hesloot (1995), Kager (1995), Hayes (1995), Ladd (1993, 1996) et centrer notre exposé sur des aspects fondamentaux et assez consensuels de la théorie métrique (ci-après : TM).

L'idée originelle centrale de la TM est que l'accentuation actualisée par le jeu des proéminences n'est que le reflet d'une *organisation rythmique hiérarchique sous-jacente* ou, plus précisément, le produit dérivé d'un groupement en constituants que le locuteur/auditeur est supposé imposer à la chaîne des éléments linguistiques. La conception de l'organisation hiérarchique de la structure phonologique n'est pas nouvelle, car elle est déjà formulée explicitement par Hockett (1955) et dans l'approche firthienne de la phonologie (Firth, 1948). En ce qui concerne la distinction fondamentale entre *ordre séquentiel* et *ordre hiérarchique* dans les diverses formes de comportements, notamment dans l'usage de la parole, il convient également de mentionner l'étude de Lashley (1951), citée par Martin (1972) comme une étape déterminante. Enfin, l'interprétation linguistique de la relation de l'accentuation au rythme évoque la conception cognitive de l'organisation rythmique (Fraisse, 1974 ; Lerdahl & Jackendoff, 1983 ; Handel, 1999), suivant laquelle le traitement perceptif du rythme intègre deux processus primaires qui concernent, respectivement, la *segmentation d'un signal d'entrée en groupes* et l'extraction d'une régularité temporelle sous la forme d'une *structure de battements* (Drake, 1998).

La TM standard est une théorie du calcul phonologique de la place de l'accent et du « comptage » des syllabes et des mots (Idsardi, 1992), afin de les regrouper en constituants, ce qui justifie le terme de *métrique* (il est en effet question dans la littérature de la « *measuring mission of metrical theory* »). De plus, la TM insiste sur le fait que seuls certains éléments s'avèrent pertinents pour le calcul, cette restriction faisant ainsi de la TM une composante à part entière de la *phonologie autosegmentale*. La conception *relationnelle* de l'accentuation distingue la TM de toutes les approches

antérieures qui considèrent l'accent comme un trait *paradigmatique* qui est spécifié localement, par référence à un segment ou à une syllabe.

On notera en passant que la syllabe, qui est la « grande absente » dans le modèle de la phonologie générative exposé par Chomsky & Halle (1968), acquiert dans la TM une place de choix (« *L'accent est le produit d'un réseau de relations hiérarchiques fondées sur la syllabe* » : Liberman & Prince, 1977). En outre, parmi les éléments déterminants de l'évolution de la TM, on notera une attention particulière portée aux phénomènes qui sont sensibles à la *quantité* et au *comptage des mores*. Cette évolution a conduit à proposer une analyse métrique de la structure interne de la syllabe, qui n'implique pas nécessairement la référence à la proéminence, mais qui emprunte à la théorie standard le critère de relation de constituance pour l'appliquer à une partition de la syllabe en : [*attaque*] + [*rime*] et de la rime en : [*noyau*] + [*coda*] (figure 7a). L'un des arguments en faveur de cette analyse se rapporte au fait que les règles phonologiques peuvent faire référence à des éléments constitutifs de la syllabe, notamment à la rime *(cf.* Blevins, 1995).

Les principaux concepts qui ont fait l'objet d'un certain consensus dans le cours du développement de la TM se résument ainsi : l'accentuation requiert une *représentation hiérarchique* (*cf. infra*) capable, par exemple, de capturer les *différents niveaux* de proéminence accentuelle de l'anglais (primaire, secondaire, tertiaire et quaternaire, selon Cruttenden, 1997) ; cette représentation hiérarchique est mieux servie par une *grille* que par un *arbre* ; la grille doit s'enrichir d'une *représentation de la constituance*. Ce dernier point a été cependant débattu et a suscité des prises de position théorique divergentes. Ces dernières conduisent soit à intégrer soit à exclure les *relations de constituance* dans la construction des grilles métriques, qui ont fini par devenir les représentations formelles dominantes de la TM (Prince, 1983 ; Laks, 1993 ; Delais-Roussarie, 2000). Laks (1993) a développé une argumentation en faveur de la grille métrique simple, fondée sur l'idée qu'un système accentuel est un *système autosegmental* régi par des principes qui n'impliquent pas nécessairement que la structure accentuelle de surface soit dérivée d'une structure sous-jacente. Ces dernières remarques nous amènent naturellement à aborder la problématique des représentations formelles de la TM.

3.2.2.1.1. Les représentations formelles de la TM

3.2.2.1.2. L'arbre métrique

La TM fait appel à deux types de représentations, dénommées respectivement *arbre métrique* et *grille métrique*. Nous commençons par présenter *l'arbre métrique* sous sa forme relationnelle stricte et sous sa forme enrichie. Dans sa version première, l'arbre métrique se présente comme une structure

arborescente dans laquelle tous les nœuds (à l'exception de la racine) possèdent des *branchements binaires* étiquetés s (*strong : fort*) ou w (*weak : faible*). Cet étiquetage a une valeur relationnelle *strictement locale*, dans la mesure où un nœud est fort uniquement parce que le nœud frère est faible, et inversement. Les exemples b et c de la figure 7 illustrent des applications de ce dispositif pour la représentation des relations de proéminence syllabique (et donc, des degrés d'accentuation) dans un mot simple et un mot composé en anglais. C'est ainsi que la représentation arborescente du mot « *Alabama* » indique que l'accent principal est associé à la syllabe « *ba* » et l'accent secondaire, à la syllabe initiale « *A* ». En effet, l'élément le plus proéminent de l'arbre est celui qui n'est dominé que par des étiquettes *s* et la racine.

Étant admis que l'accentuation est regardée comme le produit d'un réseau de relations hiérarchiques fondées sur la syllabe (*cf. supra*), il est légitime que les étiquettes s et w soient attribuées prioritairement aux nœuds terminaux correspondant à ces dernières. Néanmoins, il est d'usage d'étendre l'étiquetage relationnel aux nœuds non-terminaux de l'arbre pour spécifier le degré d'accentuation d'un mot ou d'un groupe de mots, comme dans les exemples 7c et 7d.

Il ne faut pas perdre de vue que, dans sa forme originelle, l'arbre métrique exhibe un étiquetage phonologique de la structure syntaxique superficielle d'un énoncé qui exprime des relations de proéminence entre des constituants (Beckman, 1986). À cet égard, l'exemple 7d peut être considéré comme la traduction formelle de la règle qui spécifie qu'en français chaque mot possède un accent dont la force est proportionnelle à l'importance de la coupe syntaxique qui suit ce mot (Dell, 1984). Les degrés d'accentuation sont représentés ici par les nombres 1, 2, 3 (le nombre 1 indiquant le degré le plus élevé). La version initiale de l'arbre reflète donc un isomorphisme entre la structure arborescente du texte et la constituance syntaxique. Cependant, des opérations phonologiques permettent de réorganiser cette structure initiale, pour satisfaire notamment à des *contraintes rythmiques* (*cf.* Dell, 1984). Ces dernières peuvent être résumées par une « *règle rythmique* » (Vogel *et al.*, 1995) qui précise que les s et les w doivent faire l'objet d'une *alternance régulière* et qu'il convient, par conséquent, d'éviter la *collision* (*clash*) provoquée par deux syllabes adjacentes ou le *vide accentuel* (*lapse*) qui résulterait de la succession d'un trop grand nombre de syllabes inaccentuées consécutives. L'évitement de la collision consiste, par exemple, à inverser un patron accentuel (w-s), de façon à ce qu'une séquence (*w s s), mal formée métriquement au regard de la règle rythmique, soit réorganisée sous la forme (s w s). Ce dispositif de réorganisation est le plus souvent qualifié dans la littérature de règle « *d'inversion ïambique* » *(iambic reverseal)* ou de règle « *des thirteen men* ». Une application de la règle rythmique est illustrée par la figure 7e, qui montre comment une réorganisation métrique contrainte par l'alternance des temps forts et faibles entraîne une modification notable de la hiérarchie accentuelle associée au texte. Toutefois, il importe de

souligner que la règle rythmique ne peut s'appliquer en anglais que si le matériau linguistique concerné fait partie d'un *même syntagme*, ce qui dénote une relation forte entre les contraintes accentuelles et le phrasage phonologique (Hayes, 1988).

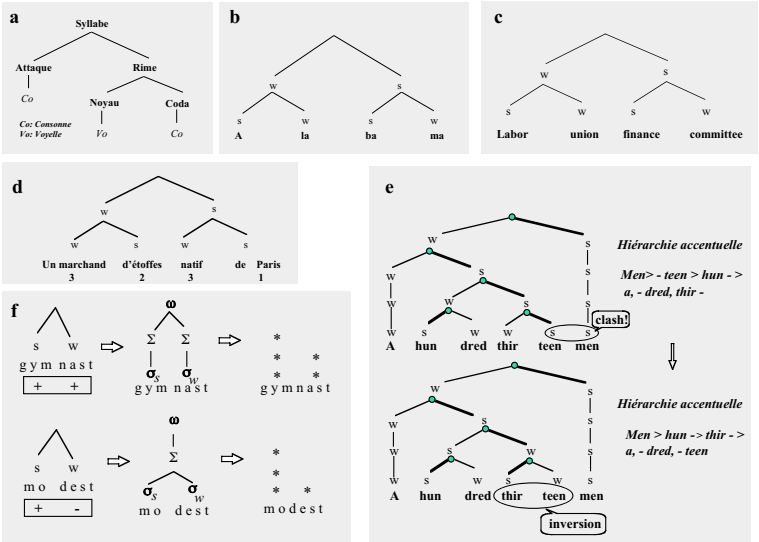

Figure 7

Représentations formelles adoptées par l'approche métrique : a) représentation métrique de la syllabe ; b) arbre métrique du mot « Alabama *» (ang.) ; c) arbre métrique du mot composé : «* Labor union finance committee *» (ang.) ; d) arbre métrique de l'énoncé : «* Un marchand d'étoffes natif de Paris *» (franç.) ; e) Illustration de l'application de la règle rythmique : «* a hundred thirteen men *» (ang.) ; f) arbres métriques enrichis de la constituance et grille métrique correspondantes (ang.)*

Il est très vite apparu que la représentation *purement relationnelle* de la version primitive de l'arbre métrique était limitée à plus d'un titre. La critique la plus sérieuse concerne une certaine inadéquation entre les objectifs de la TM et cette modalité de représentation. En effet, bien que l'arbre illustre parfaitement la proéminence relative des nœuds qui le constituent, il ne parvient pas à représenter de façon satisfaisante la structure rythmique, qui doit intégrer une *dimension temporelle* (absente de la représentation arborescente) pour rendre compte de tous les phénomènes *d'eurythmie* (Prince, 1983 ; Hayes, 1984 ; Selkirk, 1984 ; Hogg & McCully, 1987). C'est précisément ce que permet de faire la grille métrique (*cf. infra*).

Par ailleurs, le parti pris de l'isomorphisme entre la structure textuelle arborescente et la structure syntaxique, associé à la contrainte de branchement binaire, soulève un sérieux problème au regard

de l'alignement entre la structure accentuelle du texte et *l'organisation rythmique effectivement perçue*. Comment faire coïncider la richesse de la structure constituante représentée par des branchements binaires avec la granularité plus grossière du patron rythmique fondé sur l'alternance des temps forts et des temps faibles ? À la vérité, les règles de correspondance mises en œuvre dans la première version de la TM ont le désavantage de prédire plus de niveaux d'accentuation que ceux qui peuvent être réellement perçus, un défaut également attesté dans l'approche numérique des niveaux de proéminence de la phonologie générative classique (Chomsky & Halle, 1968) qui tend à prédire un nombre exagéré de niveaux d'accentuation.

Enfin, d'un point de vue plus spécifique, l'arbre purement relationnel ne parvient pas à représenter certains types de contrastes accentuels de l'anglais, comme ceux que renferment des mots comme « *contest* », « *tempest* », « *gymnast* » et « *modest* », par exemple. Il faut savoir qu'en anglais, la perception des schèmes accentuels implique une distinction entre deux catégories de voyelles, qualifiées respectivement de « *pleines* » et de « *réduites* ». C'est ainsi que les mots « *gymnast* » et « *modest* », qui possèdent une même structure métrique arborescente de type : *s w* (figure 7f), se distinguent cependant l'un de l'autre, aux niveaux acoustique et perceptif, par le fait que leur patron accentuel est construit, dans le premier cas, sur la base d'un contraste entre une voyelle accentuée et une voyelle pleine et, dans le second cas, sur celle d'un contraste entre une voyelle accentuée et une voyelle réduite. Afin de rendre compte de cette distinction, Liberman & Prince (1977) ont proposé d'utiliser le trait privatif [± Stress] comme complément au trait relationnel *s w*, ce qui donne la notation [+ +] pour le premier mot et [+ -] pour le second. Une autre solution, proposée par Selkirk (1980), consiste à introduire des nœuds *non-branchants* dans la structure arborescente, afin d'y permettre l'intégration de catégories prosodiques indépendantes hiérarchiquement supérieures à la syllabe (s), telles que le *Pied (foot)* et le *Mot Prosodique (prosodic word),* qui sont identifiées, respectivement, par les étiquettes **Σ** et w et qui ne représentent pas des unités syntaxiques, mais des *constituants phonologiques*. De ce point de vue, le Pied est considéré en anglais[1] comme l'unité phonologique formée d'une syllabe accentuée ou d'une syllabe pleine et d'un nombre optionnel de syllabes réduites, tandis que le

[1] Selon Selkirk (1977), le pied en français est généralement formé d'une *seule syllabe* à condition qu'elle ne comprenne pas un *schwa*. Si tel est le cas, la syllabe contenant le schwa s'associe à la précédente pour ne former avec elle qu'un seul pied. C'est ainsi que « partager » contient trois pieds, alors que « souvenir » n'en contient que deux, les deux premières syllabes ne formant alors qu'un seul pied. Cette interprétation s'inspire de la dichotomie traditionnelle entre langue à *isochronie syllabique (syllable-timed language)* et langue à *isochronie accentuelle (stress-timed language)* qui a provoqué et suscite encore de nombreux débats (Dauer, 1982 ; Wenk & Wiland, 1982 ; Bertinetto, 1989). De nouvelles perspectives ont été ouvertes par des recherches récentes qui semblent attester la réalité psychologique de ces « classes rythmiques » dont les études antérieures n'avaient pas réussi à rendre compte (Ramus *et al.,* 2003).

Mot Prosodique est défini par rapport à la présence de l'accent (*stress*). Il apparaît ainsi que l'organisation métrique spécifique de « *gymnast* » est illustrée par un nouvel arbre métrique dans lequel chacune des deux syllabes *s*, *w* est dominée par un Pied, tandis que celle de « *modest* » montre que les deux syllabes *s* et *w* ne sont dominées que par un seul et même Pied. Néanmoins, dans les deux cas, le Pied est dominé directement par un Mot Prosodique. L'introduction de ces unités permet de construire des *représentations enrichies* de la structure métrique arborescente qui peuvent également être traduites en grilles métriques. Cet apport, qui participe de l'idée féconde que l'arbre métrique doit représenter non pas une structure syntaxique mais une *structure phonologique*, a conduit au développement rapide de cette dernière par l'adjonction de catégories de rang supérieur au Mot Prosodique, comme le *Syntagme Phonologique* (*Phonological Phrase*) et le *Syntagme Intonatif* (*Intonational Phrase*). Bien que ces catégories participent pleinement d'une approche générale de la phonologie prosodique (*cf.* la dernière section de cette partie), elles ne concernent pas la *phonologie métrique* classique dont la version primitive ne s'intéresse pas à l'organisation des proéminences accentuelles au-delà du Pied et du Mot Prosodique. Avant d'aller plus loin, il est important de souligner qu'il existe deux interprétations de la notion de Pied : la conception « classique » (Abercrombie, 1967), d'après laquelle il équivaut au groupe accentuel (« *stress group* ») et celle de Selkirk (1980), illustrée ci-dessus, qui fait référence à une unité plus petite. Dans l'approche de Selkirk, le caractère incontournable du *principe de branchement binaire* est à l'origine des difficultés soulevées par l'inteprétation des Pieds de plus de deux syllabes. Dans ce cas, l'auteur est amenée à introduire deux niveaux de Pied dans la structure : le Pied [Σ]et le Super-pied [Σ'] (« *stress Superfoot* »). C'est ainsi que le mot « *america* », correspondant à un Mot Prosodique porteur d'un accent sur la seconde syllabe « *me* », est analysé comme étant métriquement formé, à un premier niveau, des Pieds « *a* » et « *meri* » (respectivement, faible et fort). Le Pied « *meri* » et la syllabe « *ca* » sont groupées, à un second niveau, dans un Super-pied que domine directement le Mot Prosodique. Cette analyse a été critiquée notamment par Fox (2000) qui considère qu'elle procède d'un artefact lié à la contrainte exercée par le principe de branchement binaire. On trouvera également une interprétation différente des entités primitives de l'organisation rythmique dans Jassem (1972) et une présentation des diverses approches relatives à ces entités dans Bouzon (2004).

La reconnaissance du Pied comme unité de base de l'organisation métrique (et comme domaine d'application de règles phonétiques : Selkirk, 1980) a motivé l'élaboration de *typologies paramétriques* fondées sur la comparaison des langues. Ces typologies concernent deux classes de paramètres qui se rapportent, respectivement, à la *configuration* et à la *construction* des Pieds.

Les paramètres de *configuration* concernent la *limitation* du Pied (boundedness), *la dominance* (*foot dominance*) et la *sensibilité à la quantité* (*quantity-sensitivity*). Un Pied est dit limité s'il ne contient pas plus

de deux syllabes et illimité dans le cas contraire. La dominance fait référence à la localisation de la *tête* (ou syllabe forte) dans le Pied et au sens du branchement. C'est ainsi qu'un Pied avec la *tête à droite* est branchant à gauche, alors qu'un Pied avec la *tête à gauche* est branchant à droite. S'il s'agit de Pieds limités, le premier sera qualifié de *trochaïque* et le second de *ïambique*, conformément à la terminologie métrique classique. Il a été proposé, par exemple, d'analyser le mot italien « *rappresentativamente* » comme une suite de quatre Pieds trochaïques : [rappre] [senta] [tiva] [mente] (Helsloot, 1993).

Le paramètre de *sensibilité à la quantité* se rapporte à la distribution des syllabes lourdes et légères au niveau des nœuds terminaux des Pieds.

Les paramètres de *construction* concernent la *direction* (*directionality*) suivant laquelle se développe le Pied (de gauche à droite ou dans le sens inverse) et *l'itérativité* (*iterativity*) qui précise si les Pieds sont construits de façon *itérative* ou non.

Avant d'aborder la présentation de la grille qui constitue l'objet de la prochaine section, il importe de préciser que, lors de la construction de la structure métrique, certains éléments de la chaîne des syllabes peuvent ne pas être visibles aux règles qui président à cette construction. Ce phénomène, qui est connu sous le terme *d'extramétricité* (*extrametricality*) a été décrit notamment à l'origine par Liberman & Prince (1977) et intégré ensuite à une théorie métrique plus étendue par Hayes (1980, 1982). Il a fait en outre l'objet de nombreuses discussions que nous n'avons pas la place de rapporter ici (*cf.* Hogg & McCully (1987).

3.2.2.1.3. La grille métrique

Les critiques formulées à l'encontre des représentations arborescentes classiques ont poussé les chercheurs à instituer un nouveau mode de représentation de l'organisation métrique sous la forme d'une *grille métrique simple* ou *parenthésée*. L'avantage principal de la grille métrique est qu'elle permet de représenter sur un *plan phonologique particulier* l'organisation du rythme et l'alternance des battements forts et des battements faibles en *fonction du temps*. L'une des dimensions de la grille est représentée par des colonnes de marques (* ou x) dont la hauteur est proportionnelle à la *force du battement* (et donc au degré d'accentuation). L'autre dimension, tout aussi importante, est représentée par des rangées superposées qui peuvent faire l'objet de « *tests de frappe* » individuels fondés sur l'espacement des marques à chaque étage de la grille.

La construction de la grille est asservie à un principe fondamental qui stipule que les colonnes dont elle est formée doivent être *continues*. Elle nécessite en outre la mise en œuvre d'un *mécanisme de projection* conçu comme une *interface* entre la chaîne des phonèmes et le plan phonologique autonome de la grille. Faute de place, nous renonçons à proposer des illustrations de la grille simple qui est amplement décrite dans la plupart des ouvrages d'introduction à la prosodie (*cf.* par

exemple Lacheret-Dujour & Beaugendre, 1999). Nous préférons concentrer notre attention sur les dispositifs qui président à la construction *des grilles parenthésées*.

Le recours à la grille parenthésée procède d'une conception qui considère l'accentuation comme un *principe organisateur* permettant de grouper des éléments linguistiques. Dans cette perspective, Halle & Vergnaud (1987) soutiennent l'idée que la représentation de l'accentuation implique à la fois la *proéminence* et la *constituance*. Il apparaît clairement que le mot « *autobiographic* » (*ang.*) et l'énoncé « *tous les enfants vont à l'école* » (*franç*) peuvent être segmentés en unités plus petites : [*auto*] [*bio*] [*graphic*] , [*tous les enfants*] [*vont à l'école*]. Il est également loisible de montrer que chaque groupement issu de cette segmentation recèle une syllabe plus proéminente que les autres. Cette syllabe, qui désigne la *tête métrique* du groupement, se trouve être la plus à gauche, dans le premier exemple et la plus à droite, dans le second. La représentation *conjuguée* des niveaux de proéminence et de constituance est matérialisée par une *grille métrique étiquetée* dont nous exposons ci-après les mécanismes de construction (figure 8).

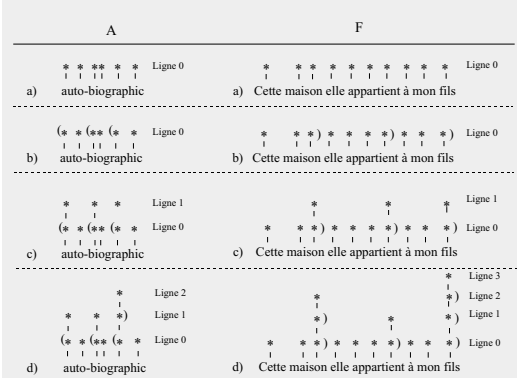

Figure 8

Grilles métriques étiquetées du mot « autobiographic *» (A : anglais) et de l'énoncé : «* Cette maison, elle appartient à mon fils *» (F : français). On notera que la seconde syllabe (prononçable) de «* cette *» n'est pas prise en compte dans la construction de la grille F. Elle pourrait y figurer, cependant, sous la notation : < * > qui signifierait que nous la traitons ici comme un élément extramétrique, c'est-à-dire comme un élément «* invisible *» aux règles d'accentuation et à la formation de la structure métrique*

Pour construire les grilles, il est proposé de représenter en *premier lieu* les unités accentuables d'une chaîne de phonèmes (telles que les syllabes) par une séquence de *marques abstraites* (*) qui constitue la *première ligne* de la grille ou ligne 0 (Aa, Fa). Afin de signaler les groupements, on utilise des parenthèses ordinaires que des règles spécifiques à la langue permettent de positionner sur la ligne 0. À la suite des propositions de Idsardi (1992) et de Halle & Idsardi (1995), il semble préférable

d'utiliser une *seule parenthèse* pour signaler la *constituance métrique*. C'est ainsi qu'une parenthèse *gauche* (exemple A) indique que le matériau placé à sa droite appartient à un même *domaine*, alors qu'une parenthèse *droite* (exemple B) a la même valeur indicatrice pour le matériau placé à sa gauche (Ad, Fb). Chaque constituant ainsi parenthésé possède une *tête* (à gauche pour A et à droite pour F) que l'on désigne comme telle en la *projetant* sur la ligne supérieure de la grille (ligne 1).

Bien que les marques de la ligne 1 (Ac, Fc) signalent un renforcement de la proéminence de certaines marques de la ligne 0, ce renforcement n'est pas uniforme, car certaines marques sont investies d'un niveau de proéminence plus élevé. Afin de capturer cette hiérarchie métrique, on construit la constituance de la ligne 1 en désignant ainsi les têtes (la syllabe la plus à droite dans A et F) qui doivent faire l'objet d'une projection sur la ligne 2 (Ad, Fd). Dans l'exemple A, le découpage de la ligne 0 correspond à une segmentation en *Pieds* dont les têtes sont projetées sur la ligne 1. La marque projetée sur la ligne 2 signale la tête du *Mot Prosodique* parenthésé au niveau de la ligne 1. Dans l'exemple B, le découpage de la ligne 0 correspond à une partition à la fois en Pieds et en Mots Prosodiques, ces deux constituants étant coextensifs dans cet exemple. Conformément à la tradition, le Pied est défini comme l'unité métrique formée d'un accent et d'un nombre optionnel de syllabes inaccentuées, tandis que le Mot Prosodique est l'unité identifiée par référence à l'accent principal. C'est ainsi qu'une séquence comme « *une des démonstrations* », dans laquelle le mot « ***dé**monstra**tions*** » comporte un accent principal final et un accent secondaire initial, peut être interprétée comme un Mot Prosodique formé de deux Pieds, ce qui peut être indiqué par le parenthésage suivant : *une des dé) monstrations*)). Toujours dans l'exemple B, le parenthésage de la ligne 1 signale un niveau de constituance supérieur et un découpage en *Unités Intonatives* dont les têtes sont à leur tour projetées sur la ligne 2. Enfin, le parenthésage de la ligne 2 délimite l'Énoncé et entraîne la projection d'une tête sur la ligne 3 qui correspond à ce que l'on appelle parfois « *accent de phrase* ».

Ces exemples montrent que la construction de la grille ne fait appel qu'à deux *mécanismes formels* qui concernent, respectivement, le placement des parenthèses et celui des marques de la grille. D'autre part, le fonctionnement de ces mécanismes est contraint par des restrictions universelles, de sorte que la marge des options possibles s'avère très limitée (Halle & Idsardi, 1995). Parmi les contraintes universelles figure en première place le principe d'eurythmie qui conditionne la bonne formation de la grille et qui fait référence à des notions comme celle de « *grille parfaite* » (Prince, 1983) ou de « *configuration idéale* » (Selkirk, 1984). Dans cet esprit, une grille est eurythmique, selon Hayes (1984), quand elle comporte une rangée sur laquelle l'espacement des marques équivaut approximativement à un intervalle de *quatre syllabes*. Il est peu concevable que dans les langues naturelles les patrons accentuels exhibent une eurythmie parfaite et en vérité peu d'évidences empiriques sont de nature à confirmer la présence de régularités quasi-métronomiques. Il en est ainsi parce que les facteurs

rythmiques ne sont pas les seuls qui motivent la place des accents et qu'il faut également tenir compte des contraintes morphologiques syntaxiques et sémantico-pragmatiques dont l'influence est loin d'être négligeable. Les structures prosodiques de surface sont l'aboutissement de compromis dont s'efforce de rendre compte une théorie comme la *Théorie de l'Optimalité* (Hammond, 1997 ; *cf.* Post, 2000 et Delais-Roussarie, 1996 pour une application au français).

3.2.3. L'approche autosegmentale de la prosodie

L'approche de la phonologie autosegmentale participe d'une théorie dont le champ d'application ne se limite pas aux fait prosodiques, dans la mesure où elle concerne également le traitement de phénomènes comme l'harmonie vocalique ou la nasalité et, plus généralement, celui de la représentation multidimensionnelle des traits phonémiques. L'approche autosegmentale s'est cependant rapidement imposée comme la clé de voûte des travaux les plus modernes sur les tons et l'intonation.

3.2.3.1. L'approche autosegmentale et la phonologie tonale

La genèse de l'approche autosegmentale de la prosodie prend appui sur une révision de l'analyse traditionnelle des systèmes tonals des langues à tons, qui a été entreprise par certains chercheurs, notamment par Woo (1969) et Leben (1973). Il découle de cette révision que certains traits phonologiques, comme les entités constitutives d'une configuration tonale donnée (montante ou descendante), ont pour domaine de spécification des empans plus petits ou plus grands que le segment phonémique traditionnel.

Cette observation a eu pour effet de motiver l'élaboration d'un nouveau cadre d'analyse dans lequel il est proposé de représenter les tons, non plus sous la forme de diacritiques associés à des segments, mais comme des unités autonomes agencées sur une ligne particulière (*tier*), parallèle à celle des unités de la ligne dite « *segmentale* ». Bien qu'il n'y ait plus de relation terme à terme entre le nombre de tons et le nombres de segments de la ligne segmentale, la contrainte de linéarisation inhérente à l'encodage de la prononciation nécessite que la ligne tonale et la ligne segmentale soient *synchronisées*. Cette opération est effectuée au moyen des *dispositifs d'association* et *d'alignement* qui spécifient comment les éléments tonals doivent être reliés aux unités de la ligne segmentale qui sont susceptibles de les recevoir (les *tone bearing units*) et qui peuvent correspondre selon le cas à des syllabes, à des rimes syllabiques ou à des mores (Goldsmith, 1976b ; Clements & Ford, 1970 ; Odden, 1995).

Dans sa version originelle (Goldsmith, 1976a, b), le dispositif d'association est assujetti à des *conditions de bonne formation* universelles qui spécifient : que les tons doivent être associés avec au moins un élément syllabique, que tous les éléments syllabiques doivent être associés avec au moins

un ton et que les lignes d'association ne doivent pas se croiser (figure 9a). Enfin, le « *principe de configuration obligatoire* » (*OCP : Obligatory Contour Principle*) précise que deux tons adjacents doivent être distincts (soit : *H H L > H L).

La figure 9a illustre ces conventions par des exemples hypothétiques. Il apparaît clairement que le nombre de tons ne correspond pas nécessairement aux nombres de syllabes (choisies ici comme unités porteuses de tons). On remarquera que dans la représentation (1), le ton H (*high*) est associé à trois syllabes consécutives, ce qui dénote un effet de *propagation tonale* (*spreading*). En revanche, on remarque que dans (2) deux tons : L (*low*) et H sont associés à une *même syllabe*. Cette dernière observation est cruciale pour l'approche autosegmentale, car elle met en lumière le fait (démontré par Woo, 1969, dans sa thèse) qu'une configuration tonale (montante ou descendante) associée à une syllabe donnée ne constitue pas une unité indivisible, mais correspond à une suite de deux tons ponctuels (*level tones*). De ce point de vue, une configuration tonale descendante ne doit pas être interprétée comme une seule unité phonologique qui serait caractérisée par le trait [+ chute], mais comme une séquence formée des deux tons [H L].

3.2.3.2. L'approche autosegmentale et la phonologie intonative

L'approche autosegmentale de la phonologie de l'intonation a tiré grand profit de l'apport du nouveau paradigme de la phonologie tonale évoqué dans la section précédente. Elle se fonde d'autre part sur la réinterprétation ou la révision de certaines approches antérieures de la prosodie considérées cependant comme des étapes marquantes. L'approche autosegmentale de l'intonation a fait l'objet de quelques variantes interprétatives qui se sont traduites notamment par le recours à l'usage de *dispositifs représentationnels* différents. Nous commencerons par présenter le modèle le plus influent, souvent qualifié de « modèle standard » dans la littérature.

3.2.3.3. Le modèle standard (par référence à l'anglais)

Bien qu'il ait motivé plusieurs tentatives d'extension à l'analyse de l'intonation de diverses langues, le modèle dit « standard » demeure une référence pour ceux qui s'intéressent principalement à l'intonation de l'anglais (et plus précisément, de l'anglais-américain). Ce modèle, qui a été exposé pour la première fois dans la thèse de Pierrehumbert (1980), a été repris et affiné par la suite dans quelques travaux (Liberman & Pierrehumbert, 1984 ; Beckman & Pierrehumbert, 1986 ; Pierrehumbert & Beckman, 1988), sur la base desquels a été édifié ultérieurement un système de notation connu sous l'acronyme de ToBI (Silverman *et al.*, 1992 ; Beckman & Ayers, 1997).

L'approche autosegmentale de l'intonation en général et celle du modèle standard que l'on vient d'évoquer en particulier, se fondent sur une conception des relations entre la substance physique

de l'intonation et l'interprétation de sa structure formelle que nous proposons d'expliciter dans les lignes suivantes. Une courbe de F0 correspondant à un tronçon de parole *dépourvu de pauses*, constitue un pseudo-continuum présentant des interruptions liées à la présences de consonnes non-voisées. Ce pseudo-continuum est cependant perçu sous la forme d'une *variation continue* de la mélodie. L'approche autosegmentale stipule que cette continuité mélodique doit être interprétée au niveau phonologique, non pas en termes de *configurations* et de *mouvements*, mais comme une *suite d'événements discrets*. Cette interprétation revient à projeter une chaîne phonologique sur une séquence de *cibles phonétiques* que la représentation du *signal physique* permet d'identifier. Cette analyse accorde donc une place prédominante à ces cibles pour le codage phonologique de l'intonation en termes d'événements discrets et un rôle secondaire aux transitions qui les relient. La conséquence de ce choix théorique est qu'il n'est pas nécessaire que chaque syllabe de la ligne segmentale soit associée à la spécification d'une unité phonologique de la chaîne tonale.

L'approche autosegmentale de l'intonation initiée par Pierrehumbert s'appuie à la fois sur les conceptions antagoniques des structuralistes américains (Pike, 1945 ; Trager & Smith, 1951) et de Bolinger (1951, 1958) dont elle parvient à neutraliser les contradictions en intégrant les apports de l'approche autosegmentale des langues à tons exposés plus haut. Les structuralistes américains proposent de décrire les contours intonatifs de l'anglo-américain au moyen de *quatre niveaux de hauteur phonémiques* (figure 9b). Les limites de cette procédure (qui ne permet pas de décrire tous les contours intonatifs de l'anglo-américain tout en rendant possible la description de contours non attestés) ont été mises en évidence par Bolinger (1951) qui suggère de décrire les patrons intonatifs en termes de *configurations* (Bolinger, 1951). Il convient dès lors de souligner que l'un des concepts fondamentaux de l'approche de Bolinger est celui *d'accent mélodique* (« *pitch accent* »). Ce dernier fait à la fois référence à l'actualisation de la proéminence et à la construction de l'intonation. Ainsi que l'a clairement formulé Bolinger, lorsqu'un élément de la chaîne est doté d'une *accentuation mélodique*, cette dernière participe de façon optimale à la *perception d'une proéminence* et s'inscrit, en même temps, par sa configuration montante ou descendante, dans la formation du *patron intonatif* global de l'énoncé. Ces changements de configurations, qui sont qualifiées par Bolinger de « *pitch obtrusions* », sont signifiantes et ont valeur de morphème. Le concept d'accent mélodique est repris à son compte par Pierrehumbert, qui lui reconnaît également la valeur de morphème, mais qui s'écarte de l'analyse formelle de Bolinger et de celle des structuralistes américains, car elle propose de réduire, en s'inspirant du modèle de la tonologie autosegmentale (*cf. supra*), l'inventaire des unités de la représentation tonale à *deux niveaux de hauteur relatifs* : les tons L et H. Aux yeux des spécialistes, cette conception « *bitonale* » de la représentation phonologique constitue une innovation fondamentale pour l'analyse phonologique de l'intonation (Ladd, 1996).

Une autre source influente de *l'approche autosegmentale de l'intonation* initiée par Pierrehumbert se rapporte à la thèse de Bruce (1977) sur la prosodie du suédois qui a attiré notamment l'attention des phonologues sur l'importance du phénomène *d'alignement* (qui se réfère à la dimension « *horizontale* » de la position des cibles tonales, la dimension « *verticale* » se rapportant pour sa part à « *l'échelle tonale* » ou au phénomène de « *scaling* »). Bruce a montré en effet que la distinctivité des accents 1 et 2 du suédois est fondée sur des différences d'alignement temporel du pic mélodique avec les limites de la syllabe accentuée, alors que, contre toute attente, le mouvement mélodique en lui-même ne joue pas un rôle décisif. La thèse de Bruce, dont la publication fait immédiatement (et harmonieusement, pourrait-on dire) suite à celle des travaux fondateurs de la mouvance de la phonologie autosegmentale et métrique (Leben, 1973 ; Goldsmith, 1976a, b ; Liberman, 1975), représente également une étape marquante de l'évolution de la recherche prosodique, pour plusieurs raisons. Elle propose d'une part un *traitement unifié* de la phonétique et de la phonologie ainsi que de l'accentuation et de l'intonation. D'autre part, elle met en lumière la notion *d'accent de syntagme* (« *phrase accent* ») et la *distinction substantielle* entre ton haut et ton bas (Pierrehumbert, 2000 ; Ladd, 2000).

Le modèle élaboré par Pierrehumbert postule que la description phonologique de l'intonation se rapporte à trois volets qui concernent, respectivement, une grammaire des *contours intonatifs syntagmatiques* (*phrasal tunes*) gérés par des séquences de tons L et H, une *représentation métrique* du texte et des *règles d'association* des contours avec le texte. Dans cette perspective, le modèle s'intègre logiquement dans le cadre théorique défini par Liberman (1975). Il existe cependant des différences substantielles entre ce cadre et celui de Pierrehumbert. Pour Liberman, les contours intonatifs (*tunes*) sont répertoriés *dans un lexique* et sont dotés d'une structure de proéminences et d'une signification propres. Les *quatre tons* qui contribuent à la construction des contours intonatifs et qui sont une réminiscence des quatre niveaux phonémiques de hauteur décrits par les structuralistes américains (Trager & Smith, 1951), participent d'une *structure arborescente* dont la fonction est supposée décrire les facteurs accentuels qui contraignent l'association entre le contour et le texte de l'énoncé (*cf.* Beckman, 1986, pour des précisions sur ce point). Or, Pierrehumbert critique le bien-fondé de la représentation arborescente pour décrire ce type de relation et elle propose de l'abandonner. De plus, elle considère que les contours ne sont pas choisis dans un lexique fermé, mais qu'ils sont en revanche engendrés par une *grammaire* opérant sur la base de *trois catégories de morphèmes* : les *accents mélodiques* (*pitch accents*), les *accents de syntagme* (*phrase accents*) et les *tons de frontière* (*boundary tones*).

Ces précisions étant apportées, il s'avère que l'appareil formel utilisé par Pierrehumbert pour décrire l'intonation de l'anglo-américain fait preuve d'une élégante simplicité. Comme nous allons

le voir, en effet, cette description ne fait appel, dans sa version première, qu'à deux symboles tonals : L et H, au signe /+/ et aux quatre diacritiques suivants : /*/,/¯/,/–/ et /%/.

On retrouve, dans l'analyse de Perrehumbert, le schéma classique de la phonologie prosodique qui distingue deux types d'événements : *internes* et *périphériques*. La notion d'événement interne fait référence à l'accentuation et celle de périphérique, aux concepts de limite (*edge*) et de frontière (*boundary*). Cette distinction est à l'œuvre aussi bien chez les structuralistes américains (où elle est actualisée par la distinction : phonèmes de hauteur/phonèmes de jointure) que dans l'œuvre de Bolinger, où elle s'exprime par la distinction entre accent et intonation. Chez Pierrehumbert, elle se traduit par le *traitement indépendant* des tons qui sont associés aux accents mélodiques, d'une part, et des tons qui sont associés à deux niveaux de frontière, d'autre part. Précisons, dès à présent, que le concept *d'association* fait référence à une opération phonologique, alors que celui *d'alignement* renvoie à la manifestation phonétique de cette opération.

Les accents mélodiques associés à une syllabe proéminente correspondant à un *pic mélodique* ou à une *vallée* sont généralement encodés par les symboles H et L assortis d'un astérisque (L* ou H*) ou d'une étoile. Cependant, les accents mélodiques peuvent être représentés par deux tons lorsque la proéminence est réalisée par un mouvement rapide correspondant à une montée ou à une chute. Si le schème associé à l'accent mélodique est *bitonal* (L H, H L), seul un des deux tons du schème est nanti de l'astérisque (ou de l'étoile). Tout ton ainsi pourvu est qualifié de *ton-étoile* (*starred-tone*). L'idée d'utiliser l'astérisque ou l'étoile pour signaler l'accent est redevable à Goldsmith (1976, a, b). Selon le formalisme proposé par ce dernier, l'accent désigné par l'astérisque est attribué à la fois à la *ligne syllabique* et à la *ligne tonale* de la représentation plurilinéaire. Dans l'exemple 9c qui illustre le patron de base /M H L/ d'un contour déclaratif de l'anglo-américain, le fait que la syllabe /pe/ et le ton H soient également pourvus d'un astérisque constitue le gage de leur *association autosegmentale*.

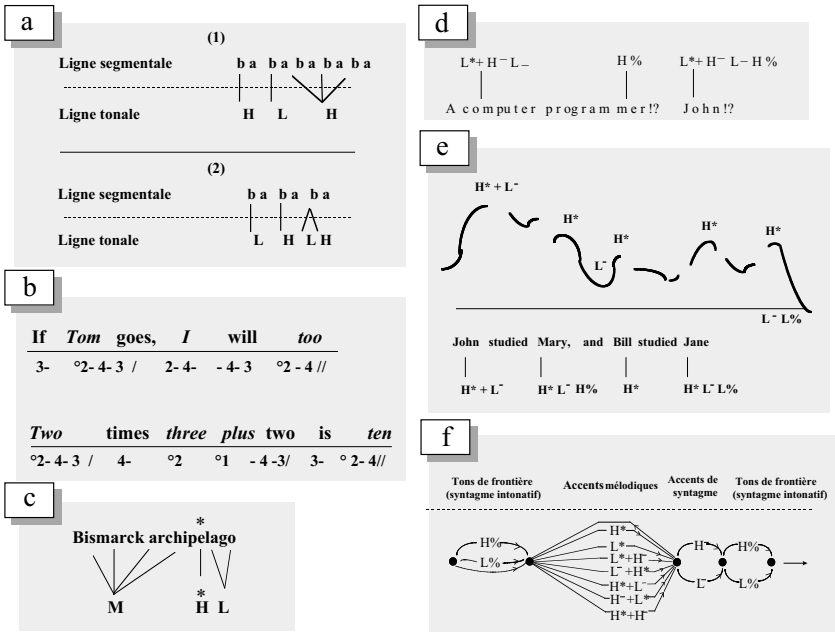

Figure 9
Illustration de différents types de représentations utilisés dans l'approche autosegmentale (a, c, d, e, f) et exemple du système de notation de l'intonation employé par les structuralistes américains (b)

Dans la première version du modèle de Pierrehumbert, le ton qui accompagne le ton-étoile est gratifié d'un tiret haut (¯). De surcroît, deux tons appartenant au même schème tonal sont reliés par le signe (+), ce qui donne lieu à une notation du type : (L*+H¯), par exemple. Par la suite ces diacritiques ont été éliminés de certains systèmes de notation, mais il ne peut en aller de même pour l'étoile. En effet, dans le modèle de Pierrehumbert, la valeur de l'étoile, telle qu'elle est illustrée par des paires comme : (L* H/L H*) ou (H* L/H L*) est *phonologique*, car elle exprime une distinction *catégorielle* entre deux façons d'associer les tons d'un même schème bitonal avec une syllabe accentuée. Cet usage de l'étoile peut être interprété phonétiquement comme le fait que le ton-étoile est aligné, du point de vue temporel, avec la voyelle accentuée, tandis que le ton non-étoile appartenant au même schème bitonal initie ou prolonge le ton-étoile dans certaines limites. Cette différence explique pourquoi le ton non-étoile est sujet à la *propagation*, ce qui est interdit pour le ton-étoile qui reste attaché à une voyelle particulière. Dans une étude récente, Grice (1995) a réexaminé la problématique

de la signification de la « *force métrique* » du ton-étoile et de la justification de son alignement avec la syllabe accentuée (*cf.* Ladd, 2000 pour une discussion sur le sujet).

L'usage de l'astérisque attribué aux tons L et H dans l'approche de Pierrehumbert permet d'établir un inventaire fermé qui comprend les six schèmes tonals distinctifs de l'accentuation mélodique de l'anglais : L*, H*, L* H, L H*, H* L, H L* (un septième schème tonal H* H, qui apparaît dans la première version du modèle (Pierrehumbert, 1980) et dans la figure 9f n'a plus été retenu par la suite).

Outre les accents mélodiques, le modèle de Pierrehumbert intègre également des entités phonologiques *périphériques* sous la forme de *l'accent de syntagme* (*phrase accent*) et du *ton de frontière* (*boundary tone*). L'accent de phrase, introduit dans la première version de la théorie, est noté par les symboles L et H, assortis d'un tiret bas (soit : L_ ou H_). Il est mentionné alors que l'accent de syntagme n'est pas associé à une syllabe particulière, mais qu'il constitue une sorte de prolongement de l'accent mélodique dans les limites d'un certain intervalle de temps, ce qui constitue une formulation assez floue. De fait le statut de l'accent de syntagme a été controversé (Ladd, 1996) et il ne figure plus dans certaines descriptions autosegmentales de l'intonation qui s'inspirent cependant fortement du modèle de Pierrehumbert (Gussenhoven, 1988). Dans une version ultérieure de ce modèle, Beckman & Pierrehumbert (1986) ont proposé de réinterpréter l'accent de syntagme comme un ton périphérique délimitatif d'un domaine appelé *syntagme intermédiaire* (« *intermediate phrase* ») (c'est-à-dire d'un constituant prosodique de rang inférieur à celui du Syntagme Intonatif (« *intonational phrase* »). Plus récemment, le concept d'accent de syntagme a été réhabilité sur la base de recherches empiriques concernant différentes langues (Grice *et al.*, 2000). Il a également fait l'objet d'une nouvelle interprétation qui consiste à ne plus l'analyser comme un ton périphérique, mais comme la marque d'une *association secondaire*. Cette conception s'efforce d'expliquer cette marque sans faire référence à la constituance prosodique, en se fondant une approche en termes d'interactions de contraintes qui s'inspire de la *Théorie de l'Optimalité* (Gussenhoven, 2000).

La troisième entité phonologique du modèle élaboré par Pierrehumbert est constituée par les *tons de frontière* (*boundary tones*). Ces derniers, qui sont associés aux limites des syntagmes intonatifs, sont également représentés par les tons L et H, assortis du diacritique (%), soit : L% et H%.

La figure 9d illustre une application du système de notation proposé par Pierrehumbert à deux énoncés consécutifs qui véhiculent la signification de la surprise (« *un programmeur en informatique* ! ? *John* ! ? ». Cet exemple utilise tous les diacritiques de la première version du modèle. Il montre notamment que : bien que l'un des deux énoncés soit plurisyllabique et l'autre, monosyllabique, ils reçoivent la même séquence de tons, car ils sont dotés du même contour (*tune*). La figure 9e illustre la notation des tons au regard de la courbe de F0 correspondant à l'énoncé : « *John observait Mary et Bill*

observait Jane ». Cet énoncé est formé de deux syntagmes intonatifs délimités respectivement par les tons H% et L%. Le premier Syntagme Intonatif comporte la séquence d'accents mélodiques : H*+ L... H*, dans laquelle le second ton H est dit *abaissé* (*downstepped*, selon l'interprétation de Pierrehumbert) par rapport au premier. Le second Syntagme Intonatif comprend une séquence de deux tons H* ayant approximativement la même hauteur. Les accents de syntagme (ou les syntagmes intermédiaires) ne sont pas spécifiés dans cet exemple. Enfin, la figure 9f représente la *grammaire intonative à états finis* que permet d'engendrer le modèle de Pierrehumbert.

L'influence grandissante exercée par ce modèle a conduit à l'élaboration d'un système de notation de l'intonation : ToBI (*Tone and Break Indices*) qui intègre à la fois la notation tonale proposée par Pierrehumbert (décrite ci-dessus) et celle des degrés de frontière, selon une échelle à quatre niveaux (Silverman *et al.*, 1992). Ce système, qui a été conçu originellement pour transcrire l'intonation de l'anglo-américain, s'est rapidement diffusé dans la communauté des intonologues, tant et si bien qu'il a donné lieu à de multiples adaptations à des langues particulières telles que : le néerlandais (ToDI), le grec (GRToBI), le coréen (K-ToBI), le japonais (J_ToBI), l'allemand (GToBI) et le russe (TORI). Par ailleurs, certains travaux sur l'intonation du français font également usage du système d'annotation initié par ToBI (Jun & Fougeron, 2000 ; Post, 2000) sans y apporter de modifications notables (on trouvera toutes les informations relatives à ToBI et à ses « dérivés », en se référant au lien *: www.ling.ohio-state.edu /~ tobi /ame_tobi/* ; *cf.* également Beckman & Ayers, 1997 et Beckman *et al.,* 2004.

Les nombreuses adaptations du système ToDI confirment le fait que le système natif n'est pas directement applicable à n'importe quelle langue car, comme l'ont souligné ses auteurs, il ne peut être regardé comme un *système de notation international.* En réalité, l'usage d'un système comme ToBI nécessité la *connaissance préalable* des patrons prosodiques distinctifs de la langue à laquelle il s'applique, ce qui constitue une contrainte particulièrement prégnante. Pour ces raisons, l'application de ToBI aux diverses variétés d'anglais fait problème (Nolan & Grabe, 1997). Le système IViE : « *Intonational Variation in English* » (Grabe *et al.*, 1998), qui se base sur le système ToBI, parvient à surmonter la difficulté en proposant d'intégrer trois niveaux de transcription prosodiques qui correspondent, respectivement, à la structure rythmique, à la structure acoustico-phonétique et à la structure phonologique. Il est ainsi possible de capturer au moyen de cette notation tripartite des aspects de la *variation régionale* que ToBI n'était pas en mesure d'identifier.

3.2.3.4. Des approches alternatives

Nous évoquerons dans cette section deux approches alternatives à celle de ToBI. La première (Mertens, 1987) s'y apparente toutefois dans la mesure où elle propose d'établir un système de

notation de l'intonation adapté à la description d'une langue particulière, en l'occurrence le français, sur la base d'un ensemble de symboles tonals. En revanche, la seconde s'en démarque à plus d'un titre, dont le principal se rapporte à la volonté d'instaurer un système de transcription de l'intonation indépendant de la langue (Hirst & Di Cristo, 1998). Aucune de ces deux approches ne peut être considérée comme une approche autosegmentale *stricte* (telle qu'elle est formulée par Goldsmith, 1976a, par exemple), bien qu'elles en conservent toutes deux l'esprit.

3.2.3.4.1. Mertens (1987)

Le système proposé par Mertens (1987) est plus complexe que celui de Pierrehumbert, car il utilise deux types de notations pour identifier les segments tonals (appelés ici : *phonèmes de hauteur*), qui se rapportent à des majuscules et à des minuscules, soit : H, h, L, l. Les majuscules servent à noter les syllabes *accentuées* et les minuscules, les syllabes *inaccentuées*. Ce système intègre en outre des distinctions de registres, grâce à l'adjonction de diacritiques (figure 10a). En définitive, les syllabes d'une chaîne peuvent se voir attribuer l'une des huit valeurs suivantes : suraigu, haut rehaussé, haut, haut abaissé, bas rehaussé, bas, bas abaissé, infra-bas. Ces valeurs établissent, selon Mertens, un *inventaire exhaustif* des phonèmes intonatifs du français. Il convient de préciser que si les valeurs H, L, H+ et L- se réfèrent à la gamme tonale du locuteur et à des intervalles mélodiques *majeurs*, en revanche, la présence d'une flèche, positive ou négative, se rapporte à des intervalles *mineurs*. La présence de deux tons sur les syllabes accentuées est destinée à indiquer la configuration mélodique (ou le contour) associée à ces dernières : HH, pour un ton de registre haut statique, H↑H, pour un ton de même registre rehaussé dans sa partie finale, etc. Dans l'approche de Mertens, il est suggéré que les relations hiérarchiques exprimées par l'organisation tonale doivent être *congruentes* avec la structure syntaxique. Le système de notation adopté par Mertens est riche et peut donc apparaître comme *sur-spécifié* au regard de *l'interprétation phonologique*, car il n'est peut-être pas nécessaire de spécifier phonologiquement la valeur tonale de *chaque syllabe*. De ce fait, le système proposé n'établit pas une distinction tranchée entre phonétique et phonologie, ce qui peut s'avérer commode dans la perspective d'une application à la synthèse de la parole (Mertens *et al.*, 2001 ; Mertens, 2002).

3.2.3.4.2. Hirst & Di Cristo (1998)

Dans l'introduction d'un ouvrage consacré à l'analyse phonologique de la prosodie du japonais, Pierrehumbert & Beckman (1988) déclarent que la subdivision entre la phonétique et la phonologie est une *question empirique* qui ne peut être résolue que par l'élaboration de *modèles complets* qui rendent *explicite* le rôle respectif de ces deux composantes dans la description des structures sonores. L'approche INTSINT (*INternational Transcription System for INTonation*), que nous

décrivons ci-après, s'efforce de fournir une réponse plausible à cette problématique en mettant en œuvre un dispositif de représentation *multidimensionnel* et *réversible* (Hirst & Di Cristo, 1998 ; Di Cristo & Hirst, 2002 ; Di Cristo *et al.,* 2002). Bien que l'approche INTSINT s'apparente à celle de ToBI en ayant également recours à des symboles tonals pour décrire l'intonation, elle s'en différencie au moins sur un point important. Alors que le système ToBI intègre la notation d'éléments *formels* (comme les tons L et H) et celle d'entité*s fonctionnelles* (comme les frontières) dans un même système de transcription, le système INTSINT ne s'attache qu'à la représentation *physico-formelle* de l'intonation. D'autre part, alors que dans ToBI la *transcription phonologique* s'effectue *directement* à partir de l'observation du signal acoustique (F0), elle se base sur un *niveau de représentation intermédiaire* dans INTSINT.

La figure 10b illustre les différents niveaux d'analyse et de représentation de l'approche INTSINT, en prenant pour exemple l'énoncé : « *Le fils de Martin a mangé les tartines* ». Le niveau de représentation le *plus concret* (ou niveau physique) correspond au tracé de la courbe brute de F0 qui se présente comme un pseudo continuum caractérisé par la présence d'interruptions correspondant à la production des consonnes non-voisés [f, t...]. D'autre part la courbe brute fait apparaître des petits sauts de F0 au début des voyelles qui succèdent à ces consonnes non voisées, ainsi que des creux sur les parties correspondant aux consonnes voisées (*cf.* la syllabe « *gé* », par exemple). Ces phénomènes constituent des microvariations (*cf. supra*) qui sont liées à des contraintes de bas niveau (de production) et qui doivent donc être écartées en vue d'une interprétation phonologique. Il importe cependant de préciser que, malgré les interruptions dues aux consonnes non-voisées, le pseudo continuum illustré par la courbe brute de F0 est *perçu* (en l'absence de pauses silencieuses) comme une *variation continue*.

Ce sont des remarques similaires à celles que nous venons de formuler qui ont conduit à proposer une *modélisation* de la courbe brute de F0 sous la forme d'une courbe *lisse* et *continue* (Hirst *et al.*, 2000). Outre les faits qu'elle représente une réalité perceptible et qu'elle permet d'éliminer des microvariations, la courbe ainsi modélisée, que nous qualifions de *représentation phonétique*, présente l'avantage de mettre en évidence des séquences de *points-cibles* (« *target points* ») dont on peut penser qu'ils constituent les éléments sur la base desquels la prononciation des patrons intonatifs est encodée. La dénomination de *représentation phonétique* se justifie par le fait que les objets de cette représentation sont des phénomènes continus. C'est notamment le cas des points-cibles qui varient à la fois sur l'échelle des fréquences et sur celle du temps. Nous proposons cependant de coder ces points-cibles à l'aide d'un *alphabet de symboles discrets* : l'alphabet INTSINT. Cet *alphabet intonétique* est constitué d'un ensemble de huit symboles (il est également possible d'utiliser une version « iconique » de cet alphabet, comme le montre la figure 10c). Trois de ces derniers servent à noter

les valeurs de hauteur des registres de *référence* qui définissent *l'espace tonal* du locuteur : M(id), T(op) et B(ottom). Cinq autres symboles sont utilisés pour noter les valeurs de hauteur *relatives* à l'intérieur de l'espace tonal : H(igher), L(ower), S(ame), U(pstepped) et D(ownstepped). Les symboles T et B constituent, respectivement, des indicateurs potentiels de la borne droite d'une Unité Intonative non-terminale (-T) ou terminale (+T). Le symbole M est l'indicateur potentiel d'une remise a niveau (*resetting*) qui caractérise généralement le début d'une Unité Intonative. L'utilisation du système INTSINT, dans l'exemple illustré par la figure 10b, conduit à la transcription de la séquence [M H D L T] [M U H D B] qui est formée de deux Unités Intonatives signalées par le parenthésage des crochets. Une séquence de ce type est qualifiée de *représentation phonologique de surface* en vertu de sa nature *catégorielle*. Précisons toutefois que cette représentation encode *formellement* la composante intonative de l'énoncé sans préjuger de la *distinctivité* des éléments qui la constituent. De ce point de vue, cette représentation ne doit pas être regardée comme une représentation *phonologique stricte*, mais plutôt comme une représentation *pré-phonologique*, qui permet, par exemple, d'informer une grammaire intonative superficielle similaire à celle de la figure 10d.

L'approche proposée par Hirst & Di Cristo (1998) intègre un dernier niveau de représentation, plus abstrait encore (partie supérieure de la figure 10b), qui est qualifié de *niveau de représentation phonologique sous-jacent* (ou profond). Cet ultime niveau est motivé par une conception théorique qui stipule que l'architecture du *système intonatif noyau* d'une langue donnée est interprétable sous la forme d'une représentation conjuguée de la *constituance prosodique* et des *gabarits tonals* qui lui sont associés (Hirst *et al.*, 2000). Cette architecture comporte deux unités : *l'Unité Intonative* et *l'Unité Tonale* auxquelles sont associés les tons L et H. L'Unité Intonative (ou Syntagme Intonatif) se définit comme une séquence caractérisée par sa cohérence mélodique (par exemple, la présence d'un schème mélodique itératif ou un phénomène d'abaissement régulier) et par la présence de marques spécifiques à ses extrémités (par exemple, l'effet de remise à niveau à sa borne (*edge*) gauche et l'association d'un ton T ou B, à sa borne droite. Deux *gabarits tonals* sous-jacents peuvent être associés à l'Unité Intonative, soit : [L L], pour l'Unité Intonative terminale et [L H], pour l'Unité Intonative non-terminale. L'Unité Tonale, qui est considérée comme l'unité *minimale de synchronisation* des segments tonals et du texte, se voit attribuer en *français* le gabarit tonal [L H] (Hirst & Di Cristo, 1984) et le gabarit tonal [H L] en anglais (Hirst, 1998).

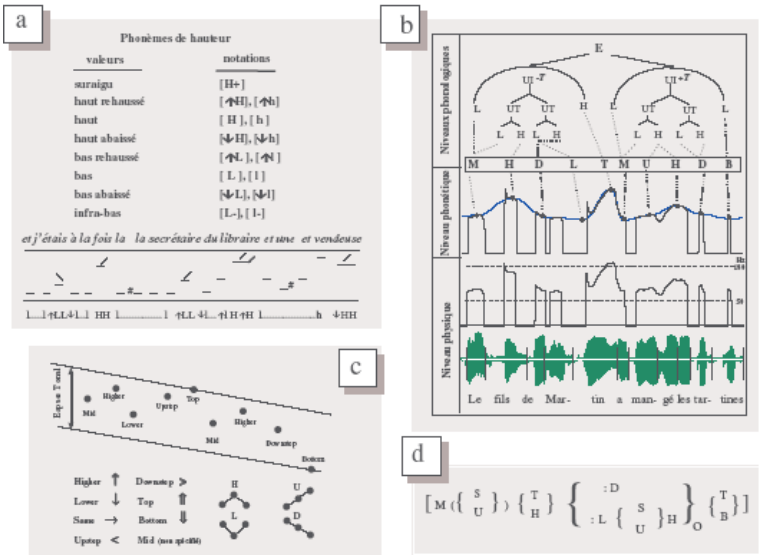

Figure 10

a) Système de notation proposé par Mertens (1987)et illustration par un exemple ; b) cadre d'analyse proposé par Hirst & Di Cristo (1998) ; c) Système de notation INTSINT ; d) grammaire stochastique de l'intonation du français fondée sur le système INTSINT

Comme le montre la figure 10b, la représentation sous-jacente de l'intonation est symbolisée par une *structure hiérarchique* qui doit faire l'objet d'une *linéarisation* pour être prononçable. Cependant, la structure linéarisée [L L H L H H] [L H L H L] correspondant à l'exemple analysé doit être réinterprétée, afin que cette représentation sous-jacente et la représentation superficielle issue du codage INTSINT soient *congruentes*. Cette réinterprétation est effectuée au moyen d'un ensemble de *règles et de principes génératifs* dont la plupart est attestée dans les travaux de *tonologie*. Dans le présent exemple, il s'agit notamment : du *principe OCP* (*Obligatory Contour Principle*) de Leben (1973) défini plus haut ; de la règle de downstep (Hyman, 1985), qui précise qu'un ton H se réécrit D, dans un contexte /H B-/ et que le ton B s'efface ensuite, dans le contexte /- D/ et de règles de conversion qui déclarent notamment que le ton L initial d'Unité Intonative se réécrit M, qu'une séquence H H se réécrit T et que le ton L final d'une Unité Intonative se réécrit B. L'application ordonnée de ces principes et règles amène à générer la séquence : [M H D L T] [M U H D B] à partir de l'entrée : [L L H L HH] [L H L H L]. Ce dispositif génératif permet ainsi de réaliser une jonction entre la démarche descendante qui procède de la théorie et la démarche ascendante qui procède de l'interprétation du signal physique.

Si nous laissons provisoirement de côté le niveau de représentation le plus élevé, il apparaît que l'approche INTSINT se présente comme une approche *physico-formelle* strictement objective, qui s'appuie uniquement sur des procédures de codages automatiques et semi-automatiques et qui n'a donc pas besoin de faire appel à des évaluations subjectives pour informer la notation. Ce qui ne signifie pas que l'approche s'interdise pour autant toute comparaison avec des analyses subjectives effectuées *séparément* (Di Cristo *et al.,* sous presse). D'autre part, à la différence de ToBI, l'approche INTSINT ne préjuge pas des *fonctions* (accentuelle, niveaux de frontière) que les indices physico-formels sont susceptibles de véhiculer. La notation INTSINT n'intègre donc pas de marques relatives à une interprétation fonctionnelle qui peut toutefois être menée de façon indépendante (*cf.* Hirst, 1977 et Di Cristo *et al.*, sous presse).

Dans la mesure où les niveaux d'analyse et d'interprétation du *module objectif* ne sont pas contraints par des préalables phonologiques, la valeur heuristique de l'approche INTSINT se prête avantageusement à la recherche des patrons phonologiques d'une langue donnée, ainsi qu'à la comparaison des langues, des *variétés régionales* et des *dialectes* (Hirst & Di Cristo, 1998). De plus, l'approche INTSINT ne pose aucun problème particulier pour la notation des corpus de *parole pathologique* (Louis *et al.*, 2003). Enfin, il importe de mettre en valeur l'un des avantages majeurs de l'approche INTSINT, qui concerne sa *réversibilité*. Il est en effet possible de procéder au *codage automatique* (ou semi-automatique) d'une représentation phonologique de surface à partir de la courbe de F0, de même qu'il est permis de *dériver automatiquement* une courbe de F0 à partir d'un codage phonologique de surface. Cette seconde démarche a été implémentée avec succès dans le système de synthèse vocale à partir du texte SYNTAIX (Di Cristo & Di Cristo, 2001).

3.2.3. Questions en débat

L'intonation, que concrétisent les variations de la fréquence fondamentale (niveau physique) et de la mélodie (niveau subjectif) est un phénomène complexe dont l'interprétation paraît nécessiter la mise en œuvre d'une approche *phonético-formelle* du type : *compositionnelle interactive*. Une telle approche considère que l'implémentation phonétique de l'intonation intègre l'interaction de plusieurs facteurs partiellement indépendants. Parmi ces derniers figurent notamment le *phrasage* des domaines intonatifs (*cf. infra*), la *composition tonale* des unités (accent mélodique, marques de frontière) incluses dans ces *domaines*, les modifications locales et globales dues à la *focalisation* et à *l'emphase*, les phénomènes d'*abaissement* (downstep, downdrift, déclinaison et abaissement final : « final lowering », *cf.* Connell & Ladd, 1990 ; Laniran & Clements, 2003), ainsi que les phénomènes de *registre*. Parmi les problèmes actuellement débattus dans le champ de la phonologie intonative d'obédience

autosegmentale, celui du *statut du registre* occupe une place non négligeable (Ladd, 1990, 1994, 1996 ; Clements, 1990 ; Inkelas & Leben, 1990, Paterson, 2000). Le terme de registre (*pitch range*) fait en réalité référence à deux dimensions partiellement indépendantes : le *niveau de hauteur tonale* (*register level*) et *la gamme* couverte par les variations de la hauteur (*span*), soit deux phénomènes qui relèvent, selon Ladd (1996), de la *dimension orthogonale* de l'organisation mélodique (figure 11c). Les phénomènes de registre sont importants pour la compréhension de l'exploitation fonctionnelle de l'espace tonal, dans la mesure où ils apportent des informations à la fois sur la proéminence, le phrasage des énoncés et du discours, le style de parole et l'engagement affectif des locuteurs. De plus, il a été montré récemment que les auditeurs sont particulièrement sensibles aux changements de registre que réalisent ces locuteurs (Portes & Di Cristo, 2003 ; Carlson *et al.,* 2004).

Dans l'approche autosegmentale « standard » de la phonologie de l'intonation anglaise, initiée par Pierrehumbert (1980) et révisée par Beckman & Pierrehumbert (1986), l'interprétation linguistique du registre se limite à celle du *downstep* (ou de la *catathèse*) qui concerne l'abaissement d'un ton haut, par rapport au ton haut précédent appartenant au même syntagme. Toute autre forme de variation de registre renvoie à *l'hypothèse de la variabilité graduelle libre* (*free gradient variability hypothesis*) qui suppose que le registre varie de façon graduelle pour véhiculer ainsi différents degrés de proéminence ou d'emphase. Cette conception, qui écarte les phénomènes de registre de l'analyse phonologique (à l'exception de l'effet de *downstep* qui vient d'être évoqué), a été critiquée par plusieurs auteurs, notamment par Ladd (1996), pour qui la dimension orthogonale à la chaîne des tons L et H (en fait, « une » dimension orthogonale à cette chaîne), exprime à la fois des distinctions *graduelles* et *catégorielles*. Une façon de capturer formellement ces dernières est introduite par Ladd au moyen de la notion de *downstep emboîté* (*nested downstep*) que nous explicitons ci-après. Selon Ladd, l'effet de downstep peut s'appliquer soit entre *deux tons* H d'un même syntagme, soit entre *deux syntagmes* consécutifs. Cette double procédure participe d'un dispositif hiérarchique qu'il est possible de fixer par une *représentation arborescente à branchements binaires* (figure 11b) similaire à celle de *l'arbre métrique* (*cf. supra*). Toutefois, chaque paire de nœuds est étiquetée : h/l (*high/low*) et non plus s/w (*strong/weak*), comme dans l'arbre métrique (*cf.* supra). Cette distinction veut également signifier que l'organisation tonale et l'organisation métrique ne sont pas nécessairement isomorphes. Pour ce qui concerne le phénomène d'emphase évoqué plus haut, son intégration à la descrption phonologique *(cf.* Hirst, 1998 pour une proposition dans ce sens) ou le rejet de cette dernière, dépend fortement de la capapcité des analyses perceptives à démontrer ou pas que la marque d'emphase est un phénomène catégoriel. Sur ce point, les données disponibles ne permettent pas d'avancer une réponse définitive (Ladd & Morton, 1997 ; Remijsen & Van Heuven, 1999).

Une autre manière d'intégrer les phénomènes de registre à la description phonologique de l'intonation consiste à introduire une *nouvelle ligne* dans la représentation autosegmentale : *la ligne du registre* (*register tier*). C'est ce que proposent notamment Inkelas & Leben (1990). Nous avons adopté une approche comparable dans notre étude des relations de la prosodie au discours en français (Di Cristo *et al.*, 2004). Les grilles d'annotation que nous avons élaborées dans le cadre de cette étude comportent en effet *plusieurs lignes* dévolues d'une part à la notion des faits prosodiques qui relèvent de la composante syntagmatique (notation des tons, des allongements, des pauses, etc.) et, d'autre part, à la notation des faits prosodiques qui se rapportent à la *composante orthogonale*. Trois lignes sont réservées à cette dernière (figure 11d). La première ligne : *Dd* (*Downtrends*) enregistre les effets d'abaissements entre des domaines nommément désignés (ex. *Dd.UI* fait référence à la notation d'un abaissement de registre entre deux unités intonatives consécutives). La seconde ligne : *ReSp* (*Register Span*) sert à la notation catégorielle des variations de la gamme tonale, à l'aide des symboles : *N* (*Normal)*, *EX* (*Expanded)* et *Com* (*Compressed)*. Enfin, la troisième ligne *ReLe* (*Register Level)* permet de noter le niveau global du registre au moyen des symboles : *N* (*Normal)*, *Rai* (*Raised)*, *Low* (*Lowered)* et *Re* (*Reset)*. Ce dernier se rapporte à la notation d'un phénomène abondamment décrit dans la littérature (Wichmann, 2000), qui concerne un retour à un *niveau de référence* dont la fonction principale est de signaler *l'initialisation* d'une nouvelle unité prosodique, telle que l'Unité Intonative ou le Paraton (Yule, 1980). Il est à souligner que la méthode de codage plurilinéaire de la prosodie exposée dans Di Cristo *et al.,* 2004, intègre également une dimension orthogonale de l'organisation temporelle sous la forme d'une notation catégorielle des variations du tempo à l'aide des symboles : N (Normal), Fa (Faster) et Slo (Slower). Le choix de coder ainsi les variables orthogonales participe en réalité de l'hypothèse que ces variables n'assument pas que des fonctions paralinguisiques et qu'elles apportent également, dans des limites qui restent toutefois à préciser, des informations sur l'organisation structurelle du discours *(cf.* Batliner *et al.,* 1997).

Parallèlement à la problématique du registre, une question semble susciter aujourd'hui un intérêt accru chez les chercheurs qui s'attachent à l'investigation des liens entre la phonétique et la phonologie prosodiques. Elle concerne l'approfondissement de la connaissance relative au problème de *l'alignement*. Dans la mesure où les correspondances entre les structures prosodiques et le texte sont régies par des *principes d'association*, il devient nécessaire de décrire comment ces associations sont *implémentés phonétiquement*, en tenant compte à la fois des contraintes de production et de perception avec lesquelles cette implémentation doit nécessairement composer. Il importe également d'évaluer la *capacité distinctive* des variations de l'alignement dans les différentes langues. Il n'est pas envisageable de développer ces thèmes importants dans cet article. En conséquence, nous suggérons au lecteur de se reporter à des articles récents qui abordent

différents aspects de la question (Xu, 2002 ; d'Imperio, 2002 ; van Santen, 2002). Une dernière remarque concernera le phénomène de *déclinaison*, tel qu'il est illustré par la figure 12a. Cette figure représente l'oscillogramme et la courbe de F0 d'un *paragraphe prosodique* (qualifié parfois de *paraton*) en parole spontanée, formé d'une suite de cinq Unités Intonatives (dont quatre non-terminales et une terminale). Il apparaît clairement sur ce document que les pics de F0 qui marquent la fin des unités non-terminales s'inscrivent dans une ligne descendante que matérialisent les pointillés. On remarque également la présence d'effets d'abaissement au sein même des Unités Intonatives. Ces observations conduisent à penser qu'il existe plusieurs niveaux ou domaines de déclinaison, dont une supra-déclinaison (Wichmann, 2000) qui s'appliquerait à l'ensemble du paragraphe prosodique. Les remarques précédentes présentent un intérêt particulier dans la mesure où d'une part il a été souvent affirmé que l'effet de déclinaison était propre à la lecture (Umeda, 1982) et où, d'autre part, ce phénomène n'a pas fait l'objet d'une analyse approfondie en français, contrairement à d'autres langues comme l'anglais (Schuetze-Coburn *et al.,* 1991 ; Wichmann, 2000), le danois (Thorsen, 1985) et le suédois (Swerts *et al.,* 1996).

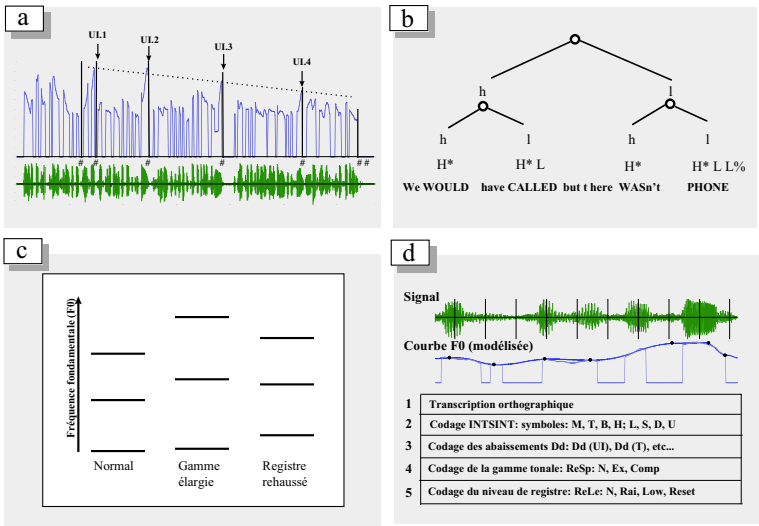

Figure 11

a) Illustration d'un effet de déclinaison dans la parole spontanée. Les pointillés relient les maxima de F0 du contour final de quatre unités intonatives consécutives (UI1-UI4)comprises dans un paragraphe prosodique ;
b) représentation arborescente du downstep emboîté (d'après la proposition de Ladd, 1990) ; c) illustration de deux effets de registre ; d) spécification du contenu d'une partie de la grille utilisée pour l'annotation de la prosodie du discours (Di Cristo et al. *2004). Les trois dernières lignes se réfèrent à des notations relatives à la composante orthogonale de l'organisation tonale.*

3.2.3.1. La phonologie des « domaines/constituants » prosodiques

3.2.3.1.1. Généralités

Dans une publication récente, Selkirk (1995a) rappelle que la question de la « *prosodie de la phrase* » (*sentence prosody*) se rapporte à trois volets distincts de la phonologie suprasegmentale qui concernent la représentation de l'intonation, celle de l'organisation métrique (ou rythmique) et celle du *phrasage* ou de la *constituance prosodique*. L'étude de la constituance prosodique, à laquelle est attachée de façon quelque peu abusive l'étiquette de *phonologie prosodique*, concerne en réalité une *sous-partie* de la phonologie prosodique, qui a été abondamment explorée notamment par Selkirk (1972, 1978, 1980, 1981, 1984) et qui a fait l'objet d'une présentation générale dans l'ouvrage de Nespor & Vogel (1986), précisément intitulé « *Prosodic Phonology* ».

Durant les dernières décennies, un consensus paraît s'être établi sur l'idée que le *phrasage prosodique* a des incidences sur l'organisation des structures tonales et accentuelles de l'énoncé ainsi que sur leurs implémentations phonétiques. D'autre part, des données empiriques ont permis de fonder l'hypothèse que la constituance prosodique est une *réalité cognitive* et qu'elle exerce une influence significative sur la planification de la parole. Dans cette perspective, on est même arrivé à supposer que la structure prosodique constituante fonctionne comme un *cadre de performance contraignant*, dans lequel doivent s'inscrire les entités de la représentation syntaxique (mots, syntagmes), afin qu'elles puissent y recevoir des marqueurs (métriques, tonales et temporelles) que gèrent la phonétique et la phonologie prosodiques (Shattuck-Hufnagel, 2000). Sur un autre versant, celui du décodage phonético-phonologique, la présence d'indices prosodiques spécifiques associée à la construction de groupements d'unités a conduit à avancer l'hypothèse que ces indices apportent des informations essentielles au *traitement perceptif* de la parole (voir ci-après pour une discussion de ces questions).

Le paradigme de la constituance prosodique est intimement mêlé à celui de la *phonologie phrasale* (*phrasal phonology*) ou *supra-lexicale*, ou encore *post-lexicale*, qui a été ouvert par Chomsky *et al.* (1956) et qui tient une place importante dans les travaux des générativistes depuis de nombreuses années (Kager & Zonneveld, 1999). La « « *phonologie phrasale* » actuelle, qui peut être regardée comme le secteur de la phonologie prosodique qui s'intéresse aux *catégories phonologiques* comprises entre le mot et l'énoncé, occupe une position privilégiée dans le modèle modulaire de la grammaire chomskyenne, en raison du fait qu'elle se confronte au problème épineux de *l'interface entre la syntaxe et la phonologie* (*cf.* sur ce sujet : Selkirk, 1984 ; Kaisse, 1985 ; Pullum & Zwicky, 1988 ; Inkelas & Zec, 1990, 1995 ; Delais-Roussarie, 2004). À cet égard, le *schéma modulaire* classique stipule que l'entrée du module phonologique reçoit l'information de sortie du module syntaxique. La prosodie, dans la mesure où elle est conçue comme un système phonologique à part entière, se trouve évidemment concernée par

cette connexion qui suscite une double problématique que l'on peut formuler en ces termes : la nature des relations entre la syntaxe et la phonologie est-elle *directe* ou *indirecte* ? Quelles sont les *informations syntaxiques* susceptibles de jouer un rôle en phonologie ? Une autre question, tout aussi cruciale, concerne plus particulièrement l'architectonie des domaines prosodiques supposés et leur raison d'être. Nous traiterons en premier lieu de cette dernière.

3.2.3.1.1.1. Le modèle de la phonologie prosodique et l'architectonie de la constituance prosodique

L'expression en usage de *phonologie prosodique* désigne une théorie explicite de la manière suivant laquelle la chaîne de la parole est organisée en un ensemble fini de *groupements* ou de *domaines phonologiques*. Elle se positionne aussi comme une approche des *relations* entre la phonologie et les autres composants de la grammaire où la syntaxe occupe une place centrale. Ce cadre de travail s'intègre logiquement dans les propositions influentes formulées par la mouvance générativiste, en ce qui concerne notamment l'organisation de la Grammaire et la place qu'y occupe la phonologie. À ce sujet, il est bon de rappeler que, conformément à la théorie de Chomsky, cette approche des relations entre les composantes adopte un modèle de grammaire dans lequel la composante syntaxique engendre des structures de phrases qui constituent l'entrée de la composante sémantique et de la composante phonologique, lesquelles sont supposées ne pas entretenir de relations directes.

Dans sa version standard *exhaustive*, l'architectonie du modèle de la phonologie prosodique comprend, de la base au sommet : la More : μ, la Syllabe : s, le Pied : S , le Mot Prosodique : w, (le Groupe Clitique : C), le Syntagme Phonologique : j, le Syntagme Intonatif : I et l'Énoncé Phonologique : U (figure 12a). Il convient de souligner que le Groupe Clitique n'est pas pris en compte par tous les auteurs, notamment par Selkirk (1978) qui considère avec d'autres que ce constituant peut ne pas être distingué du Mot Phonologique (Selkirk, 1986). C'est la raison pour laquelle nous l'avons noté entre parenthèses. L'expression « *Mot Phonologique* » est parfois utilisée à la place de celle « *Mot Prosodique* » bien que ces deux formulations ne renvoient pas à une même réalité pour de nombreux auteurs (voir par exemple l'interprétation du Mot Phonologique proposée par Hannahs, 1995, pour le français et la signification que le Mot Prosodique revêt pour la majorité des spécialistes de la prosodie du français qui l'emploient le plus souvent comme synonyme de *Groupe-*, *Syntagme-Accentuel* : « *Accentual Phrase* »).

Deux types de critères caractérisent l'organisation hiérarchique de la phonologie prosodique. D'une part les unités qui la constituent forment des *domaines* aux bornes et à l'intérieur desquels s'appliquent des règles phonologiques spécifiques. D'autre part, ces unités sont agencées selon un ordre particulier que gouverne le principe fondamental de la *hiérarchie stricte* ou de *l'étagement strict* (*Strict Layered Hypothesis* : Selkirk, 1984). Ce principe stipule en particulier que :

a) une unité non-terminale de la structure hiérarchique est obligatoirement formée d'une ou de plusieurs unités de la catégorie immédiatement inférieure ;
b) une unité d'un niveau donné de la hiérarchie est entièrement contenue dans l'unité de rang supérieur qui la domine ;
c) les structures hiérarchiques sont à branchement n-aire.

Le modèle standard de la phonologie prosodique (Nespor & Vogel, 1986) décrit ainsi une hiérarchie prosodique exhaustive (figure 12a) dont les entités constitutives se rapportent en réalité à *trois niveaux* d'interprétation différents. C'est ainsi que la syllabe et le Pied sont des entités prises en charge par la *théorie métrique*. D'autre part, le Syntagme Intonatif et l'énoncé bénéficient également d'un statut particulier, car leur interprétation est liée essentiellement à des considérations *sémantiques et pragmatiques*. Seuls les constituants de rang intermédiaire, tels que le Mot Prosodique et le Syntagme Phonologique, se trouvent réellement concernés par l'interprétation de l'interface : phonologie/syntaxe. Mais c'est le Syntagme Phonologique qui est en définitive l'objet central de cette interprétation (Truckenbrodt, 1995, 1999).

Un autre point important à souligner concerne la mise en question de certains axiomes de la théorie de la hiérarchie stricte. Il a été montré, en effet, que des constituants prosodiques d'un rang particulier peuvent dominer directement des constituants qui se situent à deux ou trois niveaux plus bas dans la hiérarchie (Shattuck-Hufnagel & Turk, 1996). De même, il est apparu qu'un Syntagme Intonatif peut dominer directement un autre Syntagme Intonatif (Ladd, 1996) et que cette forme de *récursivité* s'avère également incompatible avec la structure plate qu'impose la hiérarchie stricte dont on pourrait alors être amené à conclure qu'elle caractérise, tout au plus, une simple tendance. Ces réserves expliquent pourquoi le principe de la hiérarchie stricte a fait ultérieurement l'objet (comme nous allons l'expliquer succinctement) d'une réinterprétation (Selkirk, 1995) en termes de « *satisfaction de contraintes hiérarchisées* », conformément à l'esprit de l'influente *Théorie de l'Optimalité* (Kager, 1999).

La Théorie de l'Optimalité « *OT Theory* » de Prince & Smolensky (1993) considère qu'une grammaire s'articule selon deux composants (ou modules) distincts : un *composant génératif* qui établit, à partir d'une représentation sous-jacente, une liste de représentations de surface candidates, susceptibles d'assurer la bonne formation de la représentation de surface concernée ; et un *composant de sélection*, ou *d'évaluation*, qui choisit, parmi ces représentations superficielles, celle qui satisfait au mieux à un ensemble de *contraintes universelles* de bonne formation. Ces contraintes sont rangées selon une *hiérarchie* qui peut différer d'une langue à l'autre. Dans le cadre de cette théorie, le principe de la hiérarchie stricte se voit amendé (Selkirk, 1995b) par sa décomposition en quatre contraintes ordonnées de bonne formation prosodique, dénommées, du rang le plus haut au rang le plus bas : *Étagement*, *Dominance*, *Exhaustivité* et *Non-Récursivité*. Les contraintes de rang supérieur,

comme *Étagement* et *Dominance*, qui spécifient, respectivement, qu'un constituant de niveau inférieur ne peut dominer un constituant de rang supérieur (aucun Ci ne domine un Cj, si J > i, soit, par exemple : une syllabe ne peut dominer un Pied) et que tout constituant domine un constituant de rang inférieur (Ci doit dominer Ci-1) doivent être *respectées*. En revanche, les contraintes d'*Exhaustivité* et de *Non-Récursivité*, qui occupent un *rang inférieur* dans la hiérarchie et qui précisent, respectivement, que tous les constituants d'un niveau donné sont entièrement compris dans le constituant de rang immédiatement supérieur (aucun Ci ne domine immédiatement un Cj, J< i-1) et qu'un constituant particulier ne peut dominer un constituant de même niveau (aucun Ci ne domine un Ci), peuvent être *transgressées* (*cf.* Delais-Roussarie, 2004, pour une présentation plus détaillée des contraintes définies par Selkirk).

Enfin, il convient également d'attirer l'attention sur le fait que l'architectonie du modèle de la phonologie prosodique n'est pas figée et qu'elle est variable selon qu'elle intègre ou pas des constituants tels que : le groupe clitique, le Mot Prosodique mineur, le Mot Prosodique majeur, le Syntagme Accentuel, le Syntagme Intermédiaire, le Syntagme Phonologique mineur, le Syntagme Phonologique majeur, etc. Il peut arriver que cette profusion de termes soit interprétée comme une *absence de consensus* qui n'est pas forcément réelle sur le fond. Il n'en demeure pas moins qu'elle peut s'avérer confondante, dans la mesure où il ressort que des termes identiques peuvent être utilisés pour faire référence à des constituants différents ou que des termes différents sont employés pour désigner des constituants qui appartiennent à un même niveau de la hiérarchie (figure 12b). Nous tenterons d'apporter quelques éclaircissements sur ces divergences dans les sections suivantes.

3.2.3.1.1.2. De la raison d'être des constituants prosodiques

A priori, la raison d'être de la constituance prosodique n'est pas autre chose que l'élaboration d'un cadre de travail motivé par la force d'une hypothèse stipulant *l'universalité* de la hiérarchie prosodique. Il est donc nécessaire de vérifier si tous les niveaux postulés par la hiérarchie sont attestés dans les différentes langues et de préciser l'identité des propriétés formelles et des marques concrètes qui concourent à leur identification. Ce programme de recherche ambitieux, qui est loin d'avoir abouti, demeure l'un des thèmes privilégiés des travaux actuels sur la prosodie.

L'une des principales sources de difficulté de l'approche instaurée par la théorie de la phonologie prosodique réside dans la formulation que les règles et les principes qui fondent la constituance ne font pas uniquement référence à l'information prosodique, ce qui peut, de prime abord, paraître surprenant. En effet, certains travaux définissent les constituants prosodiques en se référant à des *critères syntaxiques*, comme : les limites des projections maximales de la syntaxe-barre (Selkirk, 1986), les relations tête/complément (Nespor & Vogel, 1986) ou la relation « *c-command* » se rapportant à la

théorie du « Liage et du Gouvernement » de Chomsky (Hayes, 1989). En général, ces études prennent également en compte des critères phonologiques et phonotactiques, tels que : le *redoublement* (*phono-*) *syntaxique* (« *Raddopiamento(fono-) sintattico* ») en italien (Nespor & Vogel, 1986 ; d'Imperio & Fivela, 2003) ; la *liaison* en français, etc. (Selkirk, 1972, 1980, 1986 ; Nespor & Vogel, 1986 ; Hayes, 1989). En revanche, d'autres études se fondent *exclusivement* sur des *critères prosodiques* relevant des organisations intonative, métrique et temporelle (Beckman & Pierrehumbert, 1986 ; Pierrehumbert & Beckman, 1988 ; Di Cristo & Hirst, 1993 ; Jun & Fougeron, 2000). Il existe donc deux approches plus ou moins concurrentes de la hiérarchie prosodique : *l'approche syntaxique* et *l'approche prosodique* (*cf.* Shattuck-Hufnagel & Turk, 1996 et Jun, 1998, pour une discussion sur le sujet).

Si nous laissons de côté la syllabe et le Pied dont il a été question plus haut, et qui font partie de la hiérarchie métrique, le premier domaine concerné par la phonologie non-segmentale est celui du mot, ou plus précisément du *Mot Phonologique* (Hayes, 1989). De nombreuses règles ont été proposées pour intégrer ce dernier à la hiérarchie, notamment la règle de dévoisement final dans diverses langues (russe, polonais, turc, catalan). Pour Chomsky & Halle (1968), le mot phonologique est une sous-partie du mot, comme la racine *sing* de *sing-ing*. D'après Hannahs (1995), le mot phonologique est formé en français d'une racine plus un suffixe. De même, dans la théorie de Selkirk (1980), le mot phonologique (ou Mot Prosodique) qu'elle qualifie de *minimal*, est formé d'une racine et des affixes étroitement liés qui affectent la place de l'accent dans cette racine. D'autres affixes (*-ing*, *-ness*), reliés de façon plus lâche à la racine n'affectent pas la place de l'accent et sont rattachés à un mot phonologique de rang plus élevé. Selkirk analyse ainsi le mot composé « *base-ball* » comme étant formé de deux mots prosodiques minimaux emboîtés dans un Mot Prosodique de rang supérieur. Toutefois, pour Hayes, la théorie de la hiérarchie prosodique doit être interprétée comme une *théorie de la constituance syntaxique* et les problèmes relatifs aux jointures internes des mots appartiennent à la *phonologie lexicale*. Par ailleurs, la définition du Mot Prosodique ne fait référence, dans certaine approches, qu'à des critères prosodiques. Dans ce cas, et en ce qui concerne le français, par exemple, le Mot Prosodique est interprété comme le *domaine d'attribution de l'accent final primaire* (Di Cristo, 1999), qui correspond à *l'Unité Rythmique* de Di Cristo & Hirst (1993), au *Syntagme Accentuel* (*Accentual phrase*) de Jun & Fougeron (2000) et à l'unité prosodique minimale du modèle de Martin (1977), c'est-à-dire, pour cet auteur, à l'unité qui ne peut recevoir qu'un seul accent.

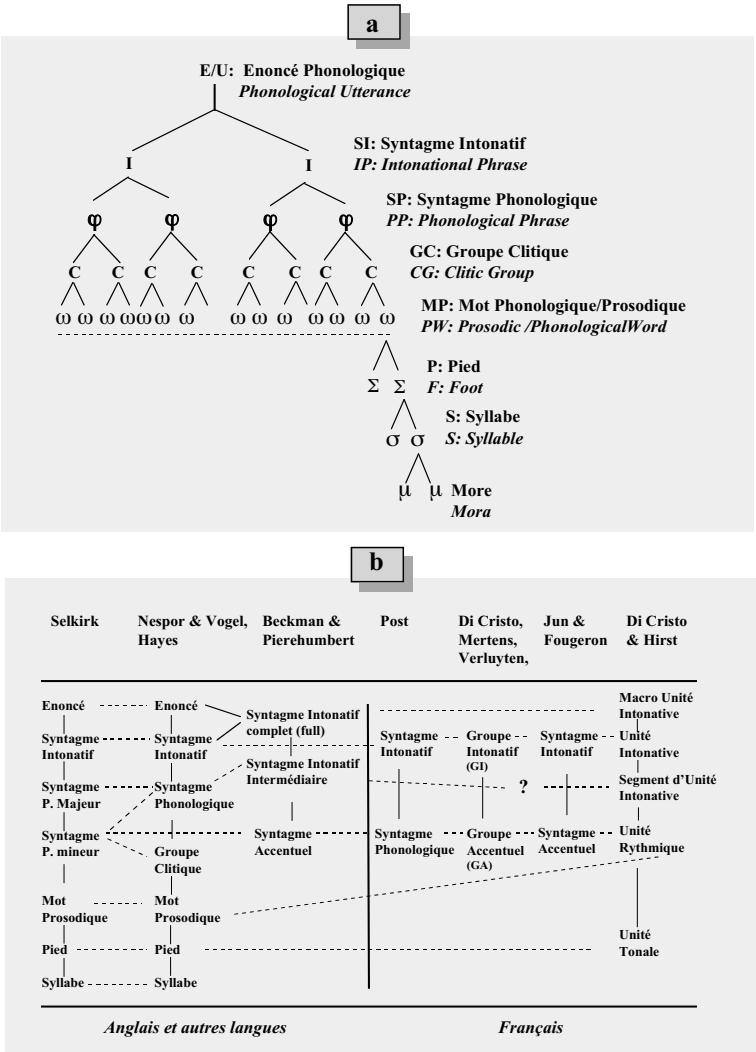

Figure 12
a) Représentation de la hiérarchie prosodique (d'après la présentation de Nespor & Vogel, 1986) ;
b) correspondances « possibles » entre les constituants/domaines prosodiques référencés dans diverses approches de la prosodie de l'anglais et du français (tableau inspiré de Shattuck- Hufnagel & Turk, 1996).

Le Syntagme - ou Groupe Clitique (*Clitic Phrase*), qui constitue la catégorie immédiatement supérieure au Mot Phonologique/Prosodique, n'est pas pris en considération dans toutes les

approches hiérarchiques. C'est ainsi qu'il ne figure pas dans la description de Selkirk (1980), alors qu'il est mis en avant dans celle de Hayes (1989). Il est défini sommairement comme le domaine formé d'un *mot lexical* (*content word*) et des *mots grammaticaux* (*functional words*) qui lui sont associés. Dans le cas du français (Garde, 1968 ; Di Cristo, 1999), la formation du groupe clitique est assujettie au principe *d'accentogénéité,* qui sélectionne les mots accentogènes, au principe de groupement qui groupe une *tête lexicale* (le mot accentogène) avec ses *dépendants* (proclitiques et enclitiques) et à la spécification du paramètre de localisation de la *tête métrique* (à droite, pour le français). Il est ainsi possible de rendre compte de la formation de constituants prosodiques dans lesquels des clitiques, par définition non-accentogènes, se trouvent être effectivement accentués : (dis-le), (vas-y), (laissez les y) (enfermés), etc. Étant donné que l'accent en question est un accent final primaire, il est difficile de distinguer sur la base de critères *strictement prosodiques* le Groupe Clitique du Mot Prosodique en français, alors que *l'approche syntaxique* opère parfois cette distinction (*cf.* par exemple, la figure 13a).

Il est déclaré dans la littérature que le *Syntagme Phonologique* (*phonological phrase*), qui occupe dans la hiérarchie le rang supérieur à celui du Groupe Clitique, peut comporter une ou plusieurs occurrences de ce dernier (Hayes, 1989). La définition du Syntagme Phonologique fait référence à plusieurs critères. Elle s'appuie notamment sur des phénomènes phonologiques/phonotactiques dont l'application n'est pas motivée par la syntaxe. Il s'agit par exemple du *redoublement syntaxique* (*raddopiamento sintaticco*) de l'italien (allongement de la consonne initiale d'un mot quand celui qui précède se termine par une voyelle accentuée), qui ne s'applique qu'à l'intérieur du Syntagme Phonologique. Il s'agit également de la règle *d'évitement de la collision accentuelle* (clash) qui n'opère que pour des éléments appartenant à un même Syntagme Phonologique. Ex. angl. : *(chi**nese dishes**) > (**chi**nese **dishes**), mais : (chi**nese**) (**dishes**) ; fr. *(sou**liers noirs**) > (**sou**liers **noirs**), mais : (des sou**liers**) (**noirs** comme...). Enfin, il s'agit de la *liaison obligatoire* en français qui, selon certains auteurs, s'applique à l'intérieur des Syntagmes Phonologiques et non à travers leurs frontières. Ces contraintes phonologiques présentées comme des *règles* doivent en réalité être validées empiriquement. De ce point de vue, si la résolution de la *collision accentuelle* paraît fournir un support empirique fiable de l'existence du Syntagme Phonologique en français, il n'en va pas de même du *phénomène de la liaison* comme l'atteste l'étude de Post (2000) dont les résultats font pencher plutôt en faveur de la *théorie de la précompilation* de Hayes (1990). Selon cette théorie, la liaison ne serait pas un processus *postlexical* à proprement parler, car elle ferait partie en réalité des formes précompilées du lexique mental. Dans ces conditions, son domaine d'application ne peut être spécifié dans le cadre interprétatif postlexical de la hiérarchie prosodique. D'autre part, la liaison est sensible à de nombreux facteurs qui ne sont pas nécessairement liés à la constituance prosodique (Post, 2000).

Il est clair que la problématique de la définition du Syntagme Phonologique (ou des Syntagmes Phonologiques, si l'on tient compte de la proposition de Selkirk (1986, 2000) de distinguer entre un Syntagme Phonologique *majeur* et un Syntagme Phonologique *mineur*) est au cœur même de celle de *l'interface* entre la prosodie et la syntaxe. On peut comprendre, dès lors, que les règles de construction de ce (ces) constituant(s) se réfèrent à la théorie syntaxique et en particulier à la syntaxe X-barre (Jackendoff, 1977). Pour Selkirk (1986), la formation du *Syntagme Phonologique majeur* regroupe une *tête lexicale* (X) avec les items lexicaux du versant *non-récursif* qui font partie de sa *projection maximale*. Conformément au « *format d'alignement généralisé* » proposé par McCarthy & Prince (1993), cette contrainte de formation est traduite par l'expression : Align (XP, R ; MaP, R), qui signifie simplement que la limite droite de toute projection maximale de la structure syntaxique (XP) doit être alignée avec la limite droite d'un Syntagme Phonologique majeur (MaP) de la structure prosodique. Une autre contrainte de formation du Syntagme Phonologique majeur, formulée par l'expression : Wrap- (XP, MaP) indique que les éléments qui font partie d'un constituant syntaxique de type XP doivent être rassemblés à l'intérieur d'un constituant prosodique de type MaP (Truckenbrodt, 1995, 1999). Cela signifie qu'un phrasage prosodique qui ne rompt pas l'unité du constituant syntaxique est préférable à tout autre phrasage, à moins qu'une rupture ne soit requise pour satisfaire à une contrainte de plus haut niveau (par exemple, par une « *contrainte de cohérence sémantique* » (Frazier *et al.*, 2004). Les contraintes de démarcation (Align) et de cohésion (Wrap) peuvent donc s'avérer conflictuelles et la dominance de l'une par rapport à l'autre est susceptible de varier en fonction de la langue. Nous ne pouvons pas nous étendre sur ce sujet, mais on trouvera un exposé circonstancié de ces contraintes dans Selkirk (2000) pour l'anglais, ainsi que dans Post (2000) et Delais-Roussarie (2000, 2004), en ce qui concerne le français.

À titre d'exemple, l'application de la contrainte d'alignement décrite par la première expression conduit à déclarer que l'exemple ci-après est formé seulement de deux syntagmes phonologiques : (*des petits enfants*) (*intelligents*), car les adjectifs pré-nominaux ne fonctionnent pas comme des têtes de Syntagme Phonologique. Il est cependant utile de préciser que le Syntagme Phonologique n'est pas défini de façon univoque par les différents auteurs. C'est ainsi que, pour Nespor & Vogel (1986), ce constituant peut inclure un complément du *versant récursif,* alors que ce n'est pas le cas pour Selkirk. Il en résulte que (*la petite fille*) constitue un seul Syntagme Phonologique pour Nespor & Vogel et pour Selkirk, tandis que « *la fille charmante* » représente un seul Syntagme Phonologique pour Nespor & Vogel : (*la fille charmante*), mais deux pour Selkirk (la fille) (charmante). Il est également admis (Verluyten, 1982) qu'un Syntagme Phonologique peut être restructuré, ce qui signifie que sa longueur est prise en compte. Dans ce cas, le premier complément d'une tête

lexicale X appartenant à son versant récursif est rattaché au syntagme qui contient X et ne forme avec lui qu'un seul Syntagme Phonologique : (des enfants) (sages) > (des enfants sages).

Le Syntagme Phonologique majeur, tel que le conçoit Selkirk, est apparemment l'équivalent du *syntagme intermédiaire* de Beckman & Pierrehumbert (1986), qui est borné par un *ton périphérique* de type L ou H. Pour Selkirk, un Syntagme Phonologique majeur (MaP) peut être constitué de un ou plusieurs *syntagme(s) phonologiques mineurs* (MiP) qui se caractérise(ent) par la présence d'une *proéminence accentuelle.* Le ton périphérique et la proéminence font ainsi référence à des marques *strictement prosodiques*. La question de savoir si la hiérarchie prosodique du français recèle une distinction formelle entre un Syntagme Phonologique mineur et un Syntagme Phonologique majeur reste ouverte. Pour l'instant, il existe un consensus sur le fait que la structure prosodique du français comporte, au-delà du Pied et en deçà de l'Énoncé, au moins deux niveaux de constituance : le Syntagme Accentuel *(Accentual Phrase)* (ou Mot Prosodique *(Prosodic Word),* selon les auteurs) et le Syntagme Intonatif *(Intonational Phrase)* ou Unité Intonative *(Intonational Unit).*

Dans le cadre de la phonologie prosodique standard, les concepts de *Syntagme Intonatif* et *d'Enoncé Phonologique* font référence à des notions larges comme la phrase et les nœuds les plus élevés de la hiérarchie syntaxique et macro-syntaxique. Au regard de la hiérarchie prosodique, l'Énoncé Phonologique doit être exhaustivement fractionné en Syntagmes Intonatifs qui doivent à leur tour être exhaustivement divisés en Syntagmes Phonologiques. En réalité, la définition de l'Énoncé se fonde sur un ensemble de *critères hétérogènes* qui se rapportent à la syntaxe (un énoncé équivaut à une phrase ou à plusieurs phrases enchâssées : Selkirk, 1981 ; Hayes, 1989), à des considérations logico-sémantiques et à des marques prosodiques, comme la présence d'une pause finale silencieuse.

Ainsi que l'admettent Hayes (1989) et Selkirk (2000), les principes qui gouvernent le phrasage en Syntagmes Intonatifs ne sont pas clairement établis. Cette opacité provient en partie du fait que le concept d'unité de sens (*sense unit*), qui est attaché à la définition de l'Unité Intonative (Selkirk, 1984), n'est pas facile à cerner. La qualification du Syntagme Intonatif sur la base de critères phonétiques et phonologiques prend en considération les propriétés suivantes : un Syntagme Intonatif est le domaine d'actualisation d'un contour mélodique perceptivement cohérent qui comprend une séquence spécifiée d'entités phonologiques telles que : l'accent mélodique, l'accent nucléaire, le ton de syntagme (*phrase tone*) et le ton de frontière (*boundary tone*). Le Syntagme Intonatif est également le domaine au terme duquel s'applique l'effet d'allongement final (*final lengthening*), ainsi que la réalisation de la pause, et à l'initiale duquel s'effectue une remise à niveau du registre (*resetting*).

Beckman & Pierrehumbert (1986), Pierrehumbert & Beckman (1988) distinguent dans leur approche deux types de Syntagmes Intonatifs *: le Syntagme Intonatif Complet* ou *majeur (Full Intonational Phrase) et le Syntagme Intonatif Intermédiaire (Intermediate Intonational Phrase*). Conformément

à la contrainte d'exhaustivité de la hiérarchie stricte, un Syntagme Intonatif complet est exhaustivement segmenté en Syntagmes Intermédiaires. Pour être reconnus comme tels, ces derniers doivent comporter un accent nucléaire et un ton de syntagme spécifique qui contrôle la valeur-cible (L ou H) du contour à la fin du syntagme. Selon les auteurs précédemment cités , le Syntagme Intonatif Intermédiaire est le domaine d'application de la *catathèse* (ou du *downstep*) en anglais et en japonais. Enfin, pour plusieurs chercheurs, la limite gauche de ce syntagme constitue un site favorable à une *accentuation précoce* dans le mot en anglo-américain (Shattuck-Hufnagel & Turk, 1996). Toutefois, il n'est pas établi, à notre connaissance, que la distinction entre Syntagme Intonatif Intermédiaire et Syntagme Intonatif Complet est applicable à d'autres langues.

Bien que la partition du texte oral en Syntagmes Intonatifs (ou Unités Intonatives) paraisse principalement motivée par des facteurs *sémantiques* et *pragmatiques* (comme ceux qui se rapportent notamment à l'organisation informationnelle des énoncés), il s'avère que cette partition peut également être sensible à des *informations syntaxiques* ainsi qu'à des *contraintes de performance*. En ce qui concerne les premières, il apparaît que la structure en constituants immédiats, les frontières de propositions, les parenthèses, les adverbes de phrases, le vocatif et les extrapositions, ont des incidences sur le phrasage intonatif des énoncés. D'autre part, des unités syntaxiques qui auraient pu former un seul Syntagme Intonatif se trouvent être scindées en plusieurs syntagmes de ce type si elles excèdent une certaine *longueur*. Ce réajustement est motivé par une contrainte de performance. La figure 13 ci-après illustre quelques types de représentation de la constituance prosodique.

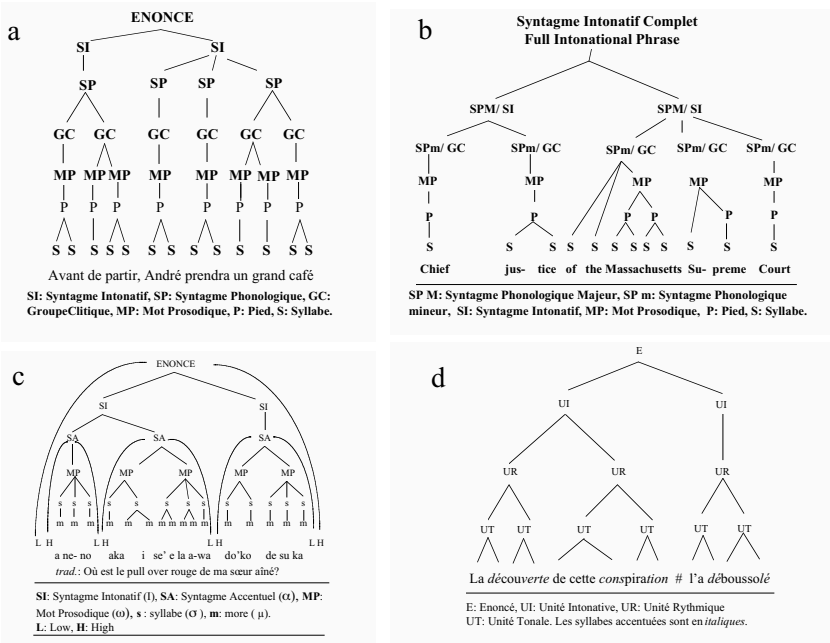

Figure 13

Illustration de quelques représentations hiérarchiques de la constituance prosodique pour diverses langues, inspirées et adaptées de la littérature : a) français : exemple de représentation exhaustive inspirée du modèle standard (on notera que chaque mot lexical est traité ici comme un Mot Prosodique) ; b) anglais : exemple adapté de Shattuck-Hufnagel. & Turk (1996) illustrant les unités décrites par Selkirk (cf. *figure 12) ; c) japonais : représentation de la constituance associant les tons : adapté de Pierrehumbert & Beckman (1988) ; d) français, d'après Di Cristo & Hirst, 1993*

3.3. Discussion et perspectives

3.3.1. Remarques générales

Dans cette dernière partie, nous proposons de fournir quelques éléments de réponse aux deux questions de portée générale qui ont été soulevées au terme de la section 3.2.3.1.1. et qui se rapportent précisément à la problématique générale des liens de la phonologie à la syntaxe. Le « ton » de cette problématique est donné par la proclamation que la structure du phrasage prosodique est indépendante mais néanmoins liée (sic) à la structure syntaxique (« *prosodic phrasing structure is independent of, but related to, the syntactic (and/or information) structure of a sentence* » : Selkirk, 2000 : 231). Cette affirmation, de prime abord paradoxale, signifie que la structure prosodique

n'est pas entièrement conditionnée par celle de la syntaxe (d'où l'absence d'isomorphisme entre les deux structures, qui est nécessaire pour justifier l'existence d'un niveau de représentation autonome de la prosodie), bien qu'elle utilise des informations transmises par cette dernière.

Pour tenter de rendre compte de la complexité de cette articulation, il est commode d'évoquer préalablement le modèle de *l'organisation modulaire* du langage qui stipule que les différentes composantes (ou les différents modules) de la grammaire sont autonomes. Au regard de cette autonomie, les échanges inter-modulaires, qui s'inscrivent naturellement dans la séquentialité des modules, activent *l'assemblage d'interfaces* où il est spécifié comment les informations transitent d'un module à l'autre. Dans cette perspective, il est postulé que le lien entre la sortie du module syntaxique et l'entrée du module phonologique subséquent n'est *pas direct.* En effet, bien que le module syntaxique « alimente » le module phonologique, une procédure de « remise en forme » ou de « restructuration » (qualifiée le plus souvent de *mapping*) est nécessaire pour que la sortie du premier puisse s'avérer compatible avec l'entrée du second. Cette opération de *reformulation* », qui est destinée à construire une structure prosodique « *acceptable* » est sensible à différents facteurs. Parmi ces derniers, il convient d'inclure notamment des phénomènes de *performance*, comme la variation du débit de parole. Bierwisch (1966) est l'un des premiers à les avoir pris en compte dans le cadre de l'approche générative des relations syntaxe/prosodie et ses propositions ont été acceptées telles quelles par Chomsky & Halle (1968).

Les contraintes associées aux effets de la performance ont d'autres incidences, qui concernent en particulier une adaptation de la structure prosodique (ou structure-P) en fonction de la longueur des groupes syntaxiques et un « aplanissement » de cette structure par rapport à la structure syntaxique qui l'informe. Ce désaccord entre la structure syntaxique et la structure-P est illustré par l'exemple bien connu de Chomsky & Halle (1968 : 372) qui montre qu'une structure syntaxique de surface-S comme :

(5) [This is [the cat that caught [the rat that stole [the cheese]]]

est restructurée sous la forme d'une structure-P tripartite :

(6) [This is the cat] [that caught the rat] [that stole the cheese]

Cette restructuration est supposée imputable à la difficulté de produire les structures branchant à droite de la construction syntaxique. Enfin, la restructuration est imposée par la *spécificité* même de la composante phonologique prosodique. En effet, cette dernière ne forme pas un système homogène, mais un ensemble de sous-systèmes interactifs *autorégulés* dont le fonctionnement participe : de la phonologie lexicale, de la phonologie métrique, de la phonologie auto-segmentale et de la phonologie des domaines. C'est à ce titre que la composante phonologique peut légitimement revendiquer une *relative* autonomie. Il n'en demeure pas moins vrai que l'information syntaxique est accessible à la

composante phonologique, alors que la position dominante dans les années 60 à 90 de l'autonomie de la syntaxe - *Phonology-free syntax* - et de l'asymétrie de l'interface phonologie/syntaxe qui en découle, prétend que la réciproque n'est pas vraie. Toutefois, cette position n'est tenable que si l'on isole la morphologie de la syntaxe, dans la mesure où il est établi que la morphologie est influencée par la phonologie, notamment par la prosodie, comme l'attestent notamment les travaux sur la morphologie prosodique de McCarthy & Prince, 1993).

3.3.2. Les trois modèles de « mapping »

L'idée que la composante phonologique a accès à l'information syntaxique étant acceptée, il se trouve que la problématique de l'interface phonologie/syntaxe s'est principalement attachée, dans les années 90, à s'interroger sur la nature de *l'information syntaxique* qui serait susceptible d'être pertinente pour influencer le phrasage prosodique. Les réponses apportées à cette question ont donné lieu à trois propositions distinctes de « *mapping* » (pour une présentation détaillée de ces propositions, *cf.* Inkelas & Zec, 1995 ; Delais-Roussarie, 2004). La première, qualifiée de « *relation-based mapping* » (Nespor & Vogel, 1986) attribue un rôle prééminent aux relations syntaxiques (tête/complément, dominance, argument/adjoint, etc.). La seconde, connue sous l'appellation de « *end-based mapping* » (Selkirk, 1986), stipule que les limites (« *edges* ») des domaines prosodiques doivent coïncider avec celles des constituants syntaxiques désignés par la syntaxe X-barre. Cette position diffère radicalement de la précédente, car elle considère que la composante phonologique-prosodique est aveugle aux informations syntaxiques relationnelles et ne prend en compte que les niveaux de la représentation X-barre et la localisation des limites des unités associées à cette représentation (*cf.* par exemple l'importance de la notion syntaxique de *projection maximale* dans la spécification du domaine prosodique qualifié de Syntagme Phonologique). Enfin, la troisième proposition, appelée « *arboreal mapping* » (Zec & Inkelas, 1990), estime que c'est l'arborescence du branchement qui influence le phrasage prosodique (alors que l'on ne prête à ce paramètre formel qu'un rôle périphérique dans le modèle proposé par Nespor & Vogel). L'approche du « *mapping arborescent* » stipule qu'au sein de l'interface l'information peut circuler dans les deux sens : de la syntaxe vers la prosodie et de la prosodie vers la syntaxe (*cf.* ci-après la discussion de ce point).

Plus récemment, à partir des années 95, on a assisté à l'émergence d'un nouveau paradigme du mapping syntaxe/phonologie. Ce dernier participe de la *Théorie de l'Optimalité* (*Optimality Theory* : OT) à laquelle se rattachent de nombreux travaux (Selkirk, 1995 ; Truckenbrodt, 1995 ; Delais-Roussarie, 1995). Le mapping syntaxe/phonologie est alors modélisé sur la base d'un ensemble de *contraintes universelles* et *compétitives* dont la hiérarchie peut différer d'une langue à l'autre. Ces

contraintes sont fonctionnellement hétérogènes et se rapportent en effet soit à des principes de bonne formation morpho-syntaxique, soit à des principes de bonne formation prosodique, soit enfin à des principes d'alignement des catégories morpho-syntaxiques et prosodiques. En ce qui concerne cette troisième option, il apparaît que *le format* des contraintes d'alignement proposé par McCarthy & Prince (1993) se présente comme un prolongement fructueux à l'approche du « *end-based mapping* » développée par Selkirk (1986).

3.3.3. Questions en débat

Avant de clore cette ultime partie de notre article, consacrée aux développements de théorie de la Phonologie Prosodique et à leurs retombées en psycholinguistique, nous souhaitons formuler quelques remarques se rapportant à des questions en débat ou qui mériteraient de l'être. Ces dernières concernent en particulier la problématique de la réalité psychologique des constituants prosodiques.

3.3.3.1. Constituant ou domaine ?

La première remarque se rapporte à des *questions de terminologie* qui dépassent en réalité de simples considérations de vocabulaire. Dans les sections précédentes, nous avons utilisé indifféremment les termes de *constituant* et de *domaine* en perpétuant de la sorte une habitude à laquelle il conviendrait peut-être de mettre fin. Les deux termes évoquent conjointement l'idée d'une unité ou d'un *groupement* nanti d'un début et d'une fin, ce qui implique la présence de marques de *bornes* et de *frontières* permettant de les délimiter. De fait, il existe une littérature abondante concernant la description des multiples indices segmentaux et prosodiques qui sont convoqués pour remplir cette fonction de « phrasage » dans les diverses langues. Le terme de *constituant* suggère, pour sa part, l'idée d'un élément faisant partie d'une *structure hiérarchiquement organisée* (*cf.* la notion de « *constituant immédiat* » en grammaire, par exemple). De plus, l'analyse syntaxique moderne a imposé une contrainte formelle (déterministe) selon laquelle un constituant contient obligatoirement un « *élément tête* » dont il est la projection (*cf.* ci-dessus, la notion de *tête métrique*). De son côté, le concept de *domaine* fait essentiellement référence à un « *empan* », défini comme une suite d'éléments où s'appliquent des *règles*, des contraintes et des *processus* particuliers (relatifs à la coarticulation, aux phénomènes de sandhi, à la liaison, aux effets prosodiques d'abaissement, etc.). Il semblerait donc que le constituant et le domaine ne partagent pas les mêmes propriétés, ce qui ne signifie pas qu'il faille pour autant les envisager comme des entités mutuellement exclusives. En effet, selon l'approche de la phonologie prosodique (Selkirk, 1978, Nespor & Vogel, 1986), les groupements formellement identifiés appartiennent à une hiérarchie stricte et actualisent la projection de têtes, tout en étant des unités au sein desquelles s'appliquent des règles phonologiques et phonétiques. Doit-on admettre que les

groupements décrits par la phonologie prosodique sont à la fois des constituants et des domaines ? Cette question a été récemment débattue par Delais-Roussarie & Fougeron (2004) qui s'appuient sur des exemples empruntés à la focalisation, à la coarticulation et à la liaison en français, pour défendre l'idée que les notions de constituant et de domaine renvoient à des *réalités différentes* et sont toutes les deux nécessaires pour rendre compte des observables de façon satisfaisante.

3.3.3.2. Primauté des constituants ?

La seconde remarque concerne le statut même des constituants prosodiques par rapport à celui qui lui est octroyé par la théorie standard de la hiérarchie prosodique. Cette dernière présente les constituants comme des entités préalables à l'analyse ou, en d'autres termes, comme des *primitives phonologiques* (dérivées des informations morphosyntaxiques). Étant donné les relations étroites qui existent entre la constituance prosodique et l'accentuation, notamment en français (Verluyten, 1982 ; Pasdeloup, 1990 ; Di Cristo & Hirst, 1993, 1997 ; Delais-Roussarie, 1995 ; Post, 2000), cette position signifie que la structure constituante est *première* et que les règles d'accentuation sont formulées à posteriori en fonction de cette structure. Cette conception est explicitée et contestée par Delais-Roussarie (2000) qui rappelle que la fragmentation en constituants prosodiques dépend à la fois de contraintes morphosyntaxiques et de contraintes rythmiques (Dell, 1984 ; Martin, 1987). Selon Delais-Roussarie, ces contraintes s'exercent *en parallèle* et la construction des patrons accentuels s'effectue ainsi de façon dynamique et progressive. Elle soutient alors que, dans ces conditions, il est difficile de maintenir que les constituants prosodiques sont des entités premières et qu'il est donc davantage plausible d'admettre que cette primauté revient aux accents. Il existe, cependant, une autre solution pour résoudre le dilemme de la « *poule et de l'oeuf* » que soulève la préséance de la constituance ou de l'accentuation. Elle consiste à proposer que constituance et accentuation participent d'une *représentation conjuguée* (Halle & Vergnaud, 1987) dont l'actualisation se traduirait par le fait que l'attribution d'un accent engendre automatiquement la formation d'un constituant et vice versa.

3.3.3.3. Réalité psychologique des constituants ?

La troisième et dernière problématique que nous souhaitons aborder est d'ordre plus général, car elle se rapporte à des *aspects cognitifs* du traitement du langage. La linguistique formelle d'obédience générativiste, qui s'est attachée notamment à l'étude de l'interface phonologie/syntaxe, prétend constituer une branche de la *psychologie cognitive* (Pollock, 1997). Cependant, comme il se doit, elle se consacre davantage à l'étude des dispositifs représentationnels de la structure interne du langage (ou à son *architecture symbolique*) qu'à celle des mécanismes qui supervisent le traitement de cette structure (et qui participent

de son *architecture fonctionnelle*), ce rôle étant traditionnellement dévolu à la *psycholinguistique*. Au regard de la *cognitique du langage*, c'est-à-dire de la branche des sciences cognitives qui s'intéresse à l'étude des relations du langage à la cognition, la communication entre les deux disciplines (voire avec cette autre discipline parente qu'est la *neurolinguistique*) est fortement souhaitable, car il est légitime de s'interroger sur la *réalité psychologique des représentations linguistiques* et de leur *interfaçage*, dans la double perspective de la *production* (Bock, 1995 ; Levelt, 1989, 1999) et de la *compréhension* (Cutler *et al.*, 1997 ; Sanderman & Collier, 1997 ; Kempsen, 1999 ; Cutler & Clifton, 1999) du langage oral.

3.3.3.3.1. Production et encodage de la parole

À notre connaissance, il existe fort peu de travaux qui s'appliquent à évaluer la validité psychologique de la théorie de la constituance prosodique dans l'optique de *l'encodage* et de la *production* de la parole (*cf.* Ferreira, 1993 ; Wheeldon & Lahiri, 1997 et Shattuck-Hufnagel, 2000, pour des synthèses récentes). En l'état des connaissances, les deux thèmes principaux qui touchent à cette problématique se rapportent d'une part à l'identité et à la taille des unités d'encodage, ainsi qu'à leur modalité d'intégration dans l'architecture fonctionnelle du système de production et, d'autre part, à l'incidence de ces questions sur l'interprétation de l'interface syntaxe/phonologie.

Le modèle de production contemporain le plus influent a été développé autour des propositions explicites de Levelt et de ses collègues (Levelt, 1989 ; Indefrey & Levelt, 2000), selon lesquelles l'élaboration de la parole procède de la mise en œuvre d'un système de traitement complexe dont la finalité consiste à établir des correspondances entre des *intentions communicatives* et des *gestes* articulatoires et phonatoires. Ces derniers engendrent à leur tour des signaux acoustiques sur lesquels se fonde en partie l'auditeur (qui utilise également les signaux para-verbaux) pour interpréter les intentions du locuteur. Le processus de production est ainsi fractionné en trois grands modules de traitement agencés séquentiellement. Au plus haut niveau de l'édifice se situe le module de *conceptualisation* qui façonne la *macroplanification* (mise en perspective de ce qui doit être dit) et la *microplanification* (traduction de la structure conceptuelle résultant de cette mise en perspective en un format propositionnel particulier) des messages. La sortie du module de conceptualisation est représentée par un *message préverbal* qui constitue l'entrée du module subséquent de *formulation*. Ce dernier prend en charge, par l'entremise de deux sous-modules successifs, *l'encodage grammatical* ou syntaxique et *l'encodage phonologique* de la forme sonore des énoncés. La sortie du module de formulation délivre une *planification phonétique* « muette » de la parole interne, laquelle forme l'entrée du *module d'articulation* qui régule la gestualité effective de la parole.

Il existe des *affinités* notables entre le modèle de production proposé par Levelt et l'approche générative classique de la phonologie prosodique, qui adhèrent de conserve à l'hypothèse fodorienne de la *modularité* (Fodor, 1983).

Nous rappelons que le processus d'encodage grammatical, tel que le conçoit Levelt, est « *guidé lexicalement* », en ce sens que les items lexicaux activés par l'information délivrée par le message préverbal, incitent l'encodeur grammatical à générer une structure syntaxique de surface appropriée. Le déploiement de cette structure de surface constitue l'entrée du module *phonologique* où se construit le *réseau hiérarchique d'unités prosodiques* qui guide, via le concours d'un hypothétique « *générateur prosodique* », la *forme phonologique* des énoncés. Le réseau fait ainsi office *d'interface* entre la syntaxe et la forme phonologique qui concerne la spécification des entités catégorielles des représentations métrique, tonale et temporelle.

Le générateur prosodique serait chargé ainsi de deux missions essentielles qui concerneraient, respectivement, la *planification métrique* et le *calcul de la configuration mélodique* de l'énoncé, ces tâches pouvant être effectuées avec un minimum d'anticipation ou de prévision (*look ahead*), selon Levelt, afin de ne pas surcharger la mémoire de travail. Un autre aspect fondamental du cadre proposé par Levelt se rapporte à la modalité de production des énoncés par *phases d'incrémentation*. Le problème est alors de savoir à quoi correspondent ces phases d'incrémentation et si les constituants prosodiques que postule la théorie linguistique sont concernés par ce processus. Dans cette perspective, Levelt soutient que le *Mot Phonologique* représente une unité d'encodage phonologique de base, le Mot Phonologique étant défini dans ce cas comme une séquence qui comprend une tête lexicale et ses dépendants, ou clitiques et qui ne possède qu'un seul accent (c'est ainsi que, d'après cette définition, la suite « *I gave it to him* » forme un seul Mot Phonologique). Il est intéressant d'observer que les *analyses empiriques* menées par Wheeldon & Lahiri (1997) confirment l'existence du Mot Phonologique comme une unité d'encodage et confortent, du même coup, l'hypothèse que l'implémentation phonétique est bien précédée par la *construction d'une structure prosodique abstraite*. La réalité psychologique du Mot Prosodique paraît donc validée sur le versant de la production (du moins, en ce qui concerne l'anglais). Mais qu'en est-il des autres constituants postulés par la théorie linguistique ? Nous n'avons pas trouvé dans la littérature d'exemples de travaux expérimentaux similaires à ceux de Wheeldon & Lahiri, mais dont l'objectif concernerait cette fois la validation psychologique du Syntagme Phonologique et du Syntagme Intonatif comme des unités sous-jacentes de l'encodage de la parole. Pourtant, selon les hypothèses formulées par Levelt, au fur et à mesure que la structure de surface se déroule de gauche à droite, le locuteur construirait par incrémentation des Mots Phonologiques et les regrouperait en Syntagmes Phonologiques et en Syntagmes Intonatifs. Nous avons vu plus haut que, pour justifier l'existence d'un constituant/domaine comme le

Syntagme Phonologique, la linguistique invoque des règles qui ne s'appliqueraient qu'à l'intérieur de cette unité et ne transgresseraient pas ses limites. Malheureusement ces règles ne sont pas toujours fiables, comme cela a été montré en particulier pour la laison en français (Post, 2000) ou le « *Raddopiamento (Fono)-Sintattico* » en italien (Agostiniani, 1992, cité par D'Imperio & Fivela, 2003). De son côté, l'analyse des paramètres phonétiques fournit une multitude d'informations relatives à des indices potentiels segmentaux et prosodiques de démarcation de domaines (Shattuck-Hufnagel, 2000). Cependant, la structure prosodique ne peut être directement déduite de ces informations de bas niveau et la problématique de *l'architecture du traitement* reste entière.

Si nous revenons à cette problématique, nous nous trouvons de nouveau confrontés à *l'interprétation de l'interface syntaxe/phonologie*. La conception modulaire la plus radicale (qui conçoit les modules comme des entités « *encapsulées* », imperméables à des informations autres que celles qui forment son entrée) considère que l'entrée du module de traitement phonologique est constitué par la sortie du module de traitement syntaxique qui correspond à une structure de surface canonique. Présenté de façon aussi simpliste, ce dispositif *séquentiel* soulève des problèmes dont nous allons débattre succinctement.

Il est désormais établi que la construction de la structure prosodique n'est pas entièrement guidée par la syntaxe et qu'elle prend en compte *d'autres types d'informations*, notamment métriques, sémantiques et pragmatiques. C'est ainsi que les informations sémantiques et pragmatiques sont cruciales pour la construction des Syntagmes Intonatifs (*cf.* ci-dessus la condition de « *sense-unit* » définie par Selkirk, 1984). Levelt suggère que l'entrée du module phonologique (et du générateur prosodique) reçoit des informations relatives à ce qu'il appelle « *intonation meaning* ». Ces dernières concernent en réalité les intentions informatives et communicatives du locuteur qui contraignent également, à l'instar de la structure syntagmatique et du déchiffrage métrique, la construction de la structure prosodique (*cf.* plus haut, la définition de l'intonation proposée par Ladd). Comment ces informations hétérogènes transitent-elles d'un module à l'autre dans le système de production ? Levelt n'est pas explicite sur ce point. Il suggère cependant que le principe général de production des énoncés par incrémentation s'applique à *différents niveaux* de traitement, suivant une *progression en parallèle*. Cette conception est apparemment difficile à concilier avec celle de la séquentialité stricte de la syntaxe vers la prosodie. En revanche, elle est davantage compatible, dans certaines limites, avec l'approche linguistique « *arboreal mapping* » de l'interface syntaxe/phonologie exposée par Zec & Inkelas (1990). Comme nous l'avons évoqué plus haut, cette approche stipule *qu'au sein de l'interface* l'information peut circuler dans *les deux sens* : de la syntaxe vers la prosodie et de la prosodie vers la syntaxe. Mais cet échange ne peut être effectif que si la construction *en parallèle* de la forme syntaxique et de la structure prosodique est également de nature *interactive*. À vrai dire, cette hypothèse concerne le problème général de la

clarification des mécanismes et de l'élucidation de l'architecture des systèmes cognitifs qui supervisent le traitement du langage (Crocker *et al.,* 2000). Il est clair que le traitement interactif enfreint le principe de la *modularité stricte* à laquelle adhèrent encore de nombreux linguistes d'obédience générativiste. À ce sujet, il convient d'accorder une attention particulière à la théorie de la *Grammaire Combinatoire* (« *Combinatory Grammar* ») élaborée par Steedman (1990, 1991). Bien que cette théorie ne réfute pas le schéma de la Grammaire Générative classique au regard de la distinction entre structure profonde et structure de surface (Chomsky, 1971 ; Jackendoff, 1972), elle propose une réinterprétation de la *structure de surface*, suivant laquelle cette dernière renferme une *structure intonative* et devient ainsi une *structure unifiée* qui est supposée servir d'entrée directe au module phonologique. L'apport de cette théorie, par référence au cadre traditionnel, est qu'elle stipule un *chemin unique* entre la *forme logique* et la *forme phonologique* des énoncés, via une *structure de surface unifiée intégrative*. La pertinence psycholinguistique de cette théorie reste toutefois à démontrer (Joshi, 1990).

Dans la perspective d'une *conception interactive* de la production de la structure syntaxique et de la structure prosodique, il nous paraît loisible de postuler (Di Cristo, 2000b) que le générateur prosodique : « *the Prosody Generator* » (dont parle abondamment Levelt, mais qui n'apparaît pas dans la représentation de l'architecture de son modèle de production, *cf.* Levelt, 1989 : 9) constitue une *unité fonctionnelle de traitement autonome* possédant la capacité d'utiliser conjointement diverses sources d'informations. Un tel dispositif devrait posséder la faculté de superviser *plusieurs niveaux* de planification prosodique. Des recherches expérimentales, qui sont conduites sur divers types de corpus, notamment sur des corpus de *parole spontanée* (Beattie, 1986) et qui se réfèrent notamment à des indices tels que la distribution et la durée des pauses dites « *d'hésitation* » (*cf.* Duez, 1997, 2001, pour la discussion de cette notion) attestent en effet de l'existence probable de phénomènes de *micro-* et de *macro-planification* (*cf.* Levelt, chap. 4 pour la discussion de ces termes), sans qu'il soit précisé pour autant quels sont *les constituants phonologiques* qui pourraient éventuellement correspondre à ces unités de planification. Plusieurs linguistes estiment que le Syntagme Intonatif (ou l'Unité Intonative) constitue une unité de planification *précoce* (Boomer, 1965 ; Bommer & Laver, 1968 ; Laver, 1970 ; Chafe, 1988 ; Blauw, 1995), qui pourrait précéder (selon notre interprétation) une planification plus tardive et à plus court terme des Mots Prosodiques et, éventuellement, d'autres unités de rang intermédiaire. Il existe de fortes présomptions en faveur de l'existence de la planification d'une *structure prosodique* susceptible de servir de *cadre préalable* à la construction de la chaîne phonologique et phonétique des énoncés (Shattuck-Hufnagel, 2000). Il reste cependant à apporter les preuves empiriques de la réalité psychologique de ce dispositif cognitif de planification et à spécifier ses modalités de traitement. Il serait également fructueux de vérifier si les *stratégies de planification prosodique* opèrent de façon similaire ou différemment dans des taches aussi diverses que la *lecture* et la *production*

de parole spontanée. Cette seconde proposition est loin d'être triviale, dans la mesure où des travaux récents postulent l'existence d'une « *prosodie implicite* » qui influencerait la résolution des ambiguïtés syntaxiques potentielles, lors de la lecture silencieuse (Fodor, 1998, 2002).

3.3.3.3.2. L'interface syntaxe/phonologie et la compréhension de la parole

3.3.3.3.2.1. Préalables

Qu'en est-il de l'interface syntaxe/phonologie sur le versant du *décodage* et de la *compréhension* de la parole ? Ce sujet alimente aujourd'hui de nombreux débats dont l'éclaircissement nécessite la définition préalable d'un cadre global de référence et la spécification de quelques repères historiques, théoriques et méthodologiques.

À l'instar de la *production* de la parole, la *compréhension* peut être envisagée comme une chaîne de traitements qui relient dans ce cas un signal acoustique d'entrée à l'interprétation du message véhiculé par ce signal. Selon la perspective psycholinguistique (Cutler & Clifton, 1999), cette chaîne comporte quatre étapes essentielles, dénommées respectivement : *décodage*, *segmentation*, *reconnaissance* et *interprétation*. Le décodage consiste à extraire le signal de parole de son environnement acoustique et à le transformer en une *représentation abstraite*. La segmentation concerne le *découpage* de l'énoncé en ses entités constitutives comme les mots et les syntagmes. L'étape de reconnaissance se rapporte à *l'identification* des mots, de la structure syntaxique et à *l'interprétation sémantique* de l'énoncé. L'ultime étape d'intégration prend en compte les *informations contextuelles* disponibles pour aboutir à la compréhension de la signification du message et des intentions de son émetteur. Sur le versant de la compréhension, la problématique de l'interface phonologie/syntaxe est donc principalement concernée par les étapes de décodage, de segmentation et de reconnaissance.

Depuis une quarantaine d'années, la question du traitement de la syntaxe occupe (influence de la théorie de la Grammaire Générative oblige) une place prépondérante dans les travaux des psycholinguistes. Les premiers travaux se sont naturellement intéressés à l'importance des *frontières* marquées dans les structures de surface. Il s'agissait d'une perspective féconde, dans la mesure où il est reconnu de nos jours que ces événements sont dotés d'une forte *prégnance perceptive*, aussi bien pour les adultes que pour les enfants (Christophe *et al.*, 2003) et qu'ils assument un rôle majeur dans l'actualisation des structures linguistiques (Hedner & Megyesi, 2003 ; Cohen *et al.*, 2001). En outre, des recherches récentes orientées vers la perception montrent que les auditeurs sont capables de prédire l'arrivée imminente d'une frontière ainsi que son degré de force, indépendamment de leur connaissance de la langue, ce qui signifie que les prévisions peuvent se fonder uniquement sur des informations prosodiques (Carlson & Swerts, 2003).

À l'origine, les travaux précurseurs ont utilisé, par exemple, le paradigme expérimental de la localisation de clicks (*cf.* Fodor & Bever, 1965) ainsi que des techniques de collage (*splicing technique*) permettant de modifier les variations de F0 (Garrett *et al.*, 1966). Parmi les conclusions tirées de ces études (Wingfield, 1975) émergent déjà les idées à résonance moderne que la parole est traitée segment par segment dans le cadre d'un *système hiérarchique* d'unités et que la segmentation perceptuelle représente un stade intermédiaire de traitement dépendant d'une *interaction* entre la *structure syntaxique* et les *indices prosodiques*, ces derniers contribuant directement à la résolution de la première. On retiendra que la formulation de ces conclusions recèle le germe des questions fondamentales actuelles relatives à l'interface phonologie/syntaxe et au rôle afférent potentiel de la prosodie. L'information prosodique est-elle transmise via des indices ou par le truchement d'une *structure* ? Y'a t-il véritablement *interaction* entre la syntaxe et la prosodie lors du traitement ? Quel rôle peut assumer la prosodie dans le « *parsing* » (ou la construction de la structure syntaxique) ?

Les termes fréquemment usités de « *parsing* » et de « *parseur* » (que nous emploierons, faute d'équivalents acceptables en français) désignent, respectivement, *l'activité cognitive* et le *système de traitement* qui participent à la segmentation et à la reconnaissance des unités grammaticales. La psycholinguistique postule plusieurs types de parseurs syntaxiques susceptibles de donner lieu à des stratégies de traitement différentes, conciliables ou antagoniques (Nicol & Pickering, 1993 ; Kempsen, 1999). Le parseur « *sériel* » ou « *stratégiquement guidé* » (*serial-*, *strategy-driven parser*) procède à partir d'informations syntaxiques (indépendamment de toute information sémantique ou pragmatique) et s'engage dans une voie (stratégie de « *garden-path* ») dès qu'un choix possible se présente, quitte à faire « machine arrière » si d'aventure une difficulté se présente. Le parseur « *en parallèle* » (*parallel-parser*) a la possibilité de calculer plusieurs configurations syntaxiques, pour chaque option qui se présente dans le cours du traitement. Le parseur en « *parallèle interactif* » (*interactive parallel-parser*) peut recourir à plusieurs types d'informations (notamment à des informations sémantiques) et les conserver en mémoire jusqu'à la sélection finale d'une interprétation. Quelle que soit la stratégie adoptée, le parseur opère par incrémentation. Cela signifie que contrairement à la conviction ancienne, le traitement n'attend pas la fin d'une séquence donnée, mais démarre avec la reconnaissance du premier mot de celle-ci.

Cette typologie élémentaire des parseurs (la littérature propose des taxinomies plus subtiles) est également à mettre en relation avec les présupposés des théories relatives à l'architecture du système cognitif, notamment en ce qui concerne la distinction entre *modularisme* et *interactionnisme*. Dans le cas d'une approche modulaire, le parseur syntaxique est *autonome* et ne se fonde que sur l'information relative à la syntaxe et à la classe des mots, que le traitement soit sériel ou en parallèle. L'apport des autres sources d'information est *anticipé* ou *différé* (*cf.* Frazier & Clifton, 1996,

pour une discussion). Toujours dans le cas d'une approche modulaire, il convient de distinguer deux propositions. D'après la première, l'analyse syntaxique s'effectue *en priorité*, sans tenir compte des autres niveaux d'analyse (y compris de la prosodie). Suivant la seconde, l'analyse syntaxique accepte *l'information prosodique* comme *entrée*. Cette seconde proposition s'avère compatible avec une architecture intégrant ce que Frazier (1990) appelle des modules « *pseudo-encapsulés* ».

À l'opposé de la conception modulaire ou autonomiste, une *approche interactive* telle que l'envisage, par exemple, Bates & MacWhinney, 1982 ; MacWhinney & Bates, 1989), considère, pour sa part, que le traitement procède d'un échange permanent des informations transmises par les différents niveaux d'analyse (syntaxe, sémantique, prosodie), à la fois de façon « *ascendante* » (*bottom-up*) et « descendante » (*top-down*). Les renseignements que nous venons de fournir sur les étapes de traitement de la parole et les modalités du parsing grammatical constituent, à notre avis, un cadre de référence utile pour interpréter les résultats des travaux dont nous allons parler ultérieurement.

Avant d'aborder la présentation de ces travaux, il est utile d'avoir un aperçu du matériau linguistique sur lequel ont porté les investigations. Dans la perspective qui nous intéresse, il apparaît que la majorité des recherches ont porté sur la résolution des ambiguïtés structurelles soit dans des *phrases* globalement ambiguës (ex : « *I read about the repayement with interest* », « *la petite brise la glace* », soit dans des *phrases* présentant une ambiguïté locale d'attachement (« *Alexander saw the girl with the binoculars* », « *J'ai appelé mes amis de Marseille* »). Ainsi que l'ont fait remarquer récemment Snedeker & Trueswell (2003), les résultats émanant de ces analyses (*cf.* ci-après) sont à prendre avec précaution pour au moins cinq raisons. La première est que les locuteurs retenus pour les expériences sont des *sujets entraînés*, *informés* des buts de l'étude et incités à *optimiser le contraste* entre les deux interprétations de la phrase structurellement ambiguë. La seconde raison est que cette situation expérimentale est très éloignée des conditions réelles de production de la parole authentique dans lesquelles les locuteurs ne fournissent pas toujours des *indices prosodiques distinctifs* susceptibles d'être exploités à des fins de désambiguïsation. Selon Snedeker & Trueswell, des expériences moins contraignantes montrent que la connaissance par le locuteur de la situation référentielle (très limitée dans le cas de la phrase isolée) affecte l'usage qu'il fait de la prosodie pour résoudre d'éventuelles ambiguïtés. Cela signifie d'abord que la relation de la prosodie à la syntaxe est *médiatisée* par le contexte de production et, ensuite, que l'information prosodique n'est peut être pas toujours présente ou nécessaire. La troisième raison est que la plupart des recherches passées ont insisté soit sur le locuteur, soit sur l'auditeur et très rarement sur leur *interaction*. La quatrième raison concerne le *caractère artificiel* des manipulations des ambiguïtés syntaxiques (*cf.* ci-dessus) et de la prosodie qui est victime d'une conception *très appauvrie* souvent limitée à la seule pause. Enfin, la cinquième raison se rapporte à la *méthodologie*. Nombre de travaux concernent le traitement de

phrases globalement ambiguës dont l'interprétation n'est résolue qu'à leur terme. Or, les recherches les plus appropriées pour évaluer le rôle éventuel de la prosodie dans le parsing sont celles qui permettent d'estimer ce rôle en « *temps réel* » ou « *à la volée* » (*on-line*), afin de montrer précisément si la prosodie peut influencer le parseur, dès que l'information qu'elle délivre s'avère disponible. Sur ce point, il semble que l'étude du traitement du langage oral soit encore en retard par rapport aux autres domaines d'investigation de la psycholinguistique. Sans doute parce que les approches sont encore influencées par des paradigmes anciens qui n'étaient pas conçus pour prendre en compte la prosodie (Cutler *et al.,* 1997). Nous ferons cependant abstraction de ces réserves dans la suite de notre article, afin d'exposer les principaux résultats acquis par les travaux des psycholinguistes sur la problématique qui nous intéresse et de faire part des réflexions que ces données suscitent.

3.3.3.3.2.2. Incidences de la prosodie sur le parsing

L'analyse des incidences de la prosodie sur le parsing se rapporte à plusieurs questions qui concernent la disponibilité efficiente de la prosodie, la nature de cette disponibilité, son « *timing* » et son impact sur le traitement de la syntaxe.

L'approche modulaire du « second type », évoquée plus haut, stipule que l'analyse syntaxique peut accepter l'information prosodique comme entrée. Encore faut-il que cette information soit disponible au moment opportun et de façon non-ambiguë. Beckman (1986) conteste à juste titre la fiabilité de cette disponibilité idéale. En effet, bien que le signal d'entrée soit normalement porteur d'une prosodie (inséparable de l'oralité), il n'est pas certain que cette prosodie soit *adaptée* à la demande du parseur, pour les simples raisons qu'elle sert à transmettre des informations sémantiques et pragmatiques et qu'elle est sujette, par ailleurs, à une large variabilité inter-individuelle et contextuelle. Contrairement à ce que l'on pourrait déduire d'une approche *algorithmique* et *déterministe* comme celle que décrivent Gee & Grosjean (1983), il n'y a pas une *structure prosodique unique* pour chaque phrase. Dès lors, il est difficile d'approuver sans réserve l'intérêt *quasi-exclusif* que les psycholinguistes ont manifesté jusqu'à présent pour l'interface syntaxe/prosodie et pour l'évaluation du rôle de la prosodie come indicateur de la seule structure syntaxique (Ferreira *et al.,* 1996).

Si nous admettons cependant qu'une information prosodique e*fficace* peut être acccessible au parseur, à un moment ou à un autre de son traitement, il est légitime de s'interroger sur la nature de l'information prise en compte. Jusqu'à une époque récente, la plupart des travaux se sont efforcés d'établir des liens entre la structure syntaxique et *les indices prosodiques* susceptibles de contribuer à l'indication de *frontières* et de *proéminences.* Parallèlement, ces indices ont fait l'objet de

nombreuses descriptions dans des études phonétiques qui ont permis notamment de mettre en évidence le rôle proéminent des facteurs temporels, tels que la pause et les allongements syllabiques pour le signalement des frontières (*cf.* par exemple les travaux précurseurs de Lehiste, 1973 ; Lehiste *et al.,* 1976). Ainsi que le soulignent Cutler *et al.* (1997), ces indices ont la capacité d'assumer une double fonction de *disjonction* et de *jonction* et d'indiquer, par là-même, soit une clôture de constituant, soit l'adjonction d'un nœud (ou d'un attachement) à un constituant donné. En ce qui concerne les proéminences, ces dernières contribuent à manifester différents *niveaux d'accentuation*, notamment l'accentuation de focalisation et il semble établi que l'identification de ces marques participe à la fois à la segmentation lexicale et à la reconnaissance de la structure sémantique des messages (Cutler, 1997 ; Cutler *et al.*, 1997 ; Cutler & Clifton, 1999).

On notera que les travaux des phonéticiens auxquels nous avons fait allusion se caractérisent par le fait qu'ils établissent des liens directs entre la forme syntaxique et des paramètres phonétiques. Or, l'analyse du versant de la production de la parole nous a conduit à présenter l'hypothèse que *l'implémentation phonétique* est guidée par une *représentation phonologique* que supervise la construction d'un *réseau hiérarchique de constituants prosodiques*. À la lumière de cette hypothèse plutôt consensuelle, il est permis de se demander si la relation entre la syntaxe et la prosodie, sur le versant de la compréhension, n'est pas également médiatisée par une *structure prosodique constituante* hiérarchique qui pourrait servir d'entrée à l'analyse syntaxique (hypothèse du *traitement modulaire*) ou faire l'objet d'un traitement conjugué avec celle-ci (hypothèse du *traitement interactif)*. La question posée ici est de savoir si l'auditeur construit à partir du signal une *représentation phonologique complète* de la prosodie (Beckman, 1996 ; Schafer, 1996) qui pourrait informer le parseur et assister du même coup l'interprétation sémantique des énoncés et du discours, car il est indéniable que la *structure prosodique* est liée à la fois à la syntaxe et à l'organisation informationnelle (Steedman, 1991 ; Croft, 1995). Il est certain qu'une réponse précise à cette question contribuerait à faire avancer nos connaissances sur la réalité psychologique des unités postulées par le modèle de la phonologie prosodique (*cf. supra)*. Toutefois, il ne semble pas que les travaux relatifs à l'exégèse du parsing se soient fondamentalement attachés à résoudre cette problématique. Probablement en raison du décalage qui existe encore entre les préoccupations de la psycholinguistique et les nouvelles théories linguistiques de la prosodie que nous avons présentées plus haut dans cet article.

Les questions relatives au « *timing* » de l'information prosodique et de son incidence sur le parsing font l'objet de ce dernier développement, dans lequel nous proposons de rendre compte de la *disparité* des résultats obtenus. Ainsi que le remarquent Cutler *et al.* (1997) dans un survol récent de la littérature, les études sur le rôle exact de la prosodie dans le parsing ne sont pas toujours explicites. Ce sont donc des tendances (souvent conflictuelles) qui ressortent principalement de ces

travaux. Les résultats qui illustrent ces tendances s'organisent selon un continuum, en fonction de l'importance attribuée au rôle de la prosodie dans le parsing.

À l'une de ces extrémités (la polarité négative, dirons-nous) se situe l'étude de Watt & Murray (1996), d'après laquelle la prosodie peut influencer différents aspects de l'interprétation, mais pas le parseur proprement dit, qui se fonde uniquement sur l'information *lexicale et syntaxique* (*cf.* Levelt, 1978). Une telle proposition est congruente avec l'idée, énoncée plus haut, que la prosodie est convoquée pour remplir d'autres fonctions et qu'elle ne peut donc pas constituer une source d'information *dédiée* au parseur. Cette interprétation est également compatible avec les résultats d'une recherche récente centrée sur le rôle de la prosodie dans le parsing intra-propositionnel des énoncés (Pynte, 1998). En effet, les données analysées dans cette étude ne confirment pas l'incidence de la prosodie sur l'analyse syntaxique. Selon l'auteur, les *ruptures prosodiques* (*breaks*) peuvent exercer une influence à un autre niveau structurel, qui relève cette fois de *l'interprétation du discours*. Il n'en demeure pas moins, comme le suggère également l'auteur, que la prosodie peut entrer en jeu au stade de *genèse de la constituance* en activant la construction « *d'unités de phrasage* » de bas niveau. Ces unités correspondent-elles à des Mots Prosodiques ? Une réponse affirmative à cette question permettrait d'effectuer un pas supplémentaire, dans la mesure ou elle contribuerait à la validation de la réalité psychologique de cette unité, dont le rôle se trouverait ainsi confirmé du double point de vue de la *production* (*cf.* supra) et de la *compréhension* du langage parlé.

Il émane des autres travaux menés par les psycholinguistes l'impression générale que la prosodie est bel et bien prise en compte dans le parsing, ce qui nous amène à aborder la problématique du « *comment* » et du « *quand* », à laquelle se greffent deux questions essentielles : celle de savoir si la prosodie *dirige* le parseur ou ne fait que *l'assister* (Stirling & Wales, 1996) et celle de la consultation à *la volée* (*on-line*) ou *différée* de la prosodie (Cutler *et al.*, 1997).

La prosodie a la faculté de diriger le parseur si elle est utilisée *précocement*. Des études récentes semblent montrer que cette influence précoce est effective (Snedeker & Trueswell, 2003). Il est particulièrement intéressant de noter que ce fait est confirmé par des travaux utilisant la méthode des « *potentiels évoqués* », qui permet précisément d'étudier l'influence de la prosodie à *la volée*, durant le défilement de l'énoncé (Jescheniak *et al.*, 1998 ; Steinhauer *et al.*, 1999). Ces travaux montrent notamment qu'une accentuation non appropriée ou qu'un groupement prosodique mal formé peut perturber le processus de parsing. Le fait que la prosodie puisse influencer l'étape initiale du parsing est compatible à la fois avec *l'approche modulaire* et *l'approche interactive*. Cependant les résultats obtenus par Snedeker & Trueswell (2003) penchent en faveur d'un *traitement interactif*, dans le sens où la prosodie peut influencer l'interprétation d'une phrase ambiguë avant même le début de la proposition dont l'attachement est à l'origine de l'ambiguïté. L'étude de Swinney & Love

(1998) présente un intérêt particulier, car elle s'intéresse au traitement des *relations syntaxiques discontinues* qui sont assez fréquentes dans la *parole spontanée* (ces relations concernent le traitement des groupes syntaxiques qui sont contigus dans la représentation syntaxique sous-jacente, mais disjoints dans les formes de surface). Selon les résultats de cette étude, il apparaît que le traitement du parseur est sensible aux *informations extra-syntaxiques,* notamment aux informations pragmatiques et prosodiques. Les auteurs en déduisent que l'information prosodique peut aider à retrouver « *on line* » l'ordre syntaxique « *canonique* » de l'énoncé.

Les travaux que nous venons de citer se positionnent à l'autre extrémité du continuum évoqué plus haut, dans la mesure où ils attribuent à la prosodie un rôle décisif dans le parsing en *contraignant* le *traitement initial* de ce dernier. En revanche, d'autres travaux, qui accordent également une certaine importance à la prosodie, semblent occuper, vis-à-vis de ce continuum, une position plutôt intermédiaire. C'est ainsi que les résultats des recherches de Beach (1991) et Pynte & Prieur (1996) plaident plus en faveur d'une relation *cumulative* de la syntaxe et de la prosodie que d'une relation *interactive*, en ce sens que la prosodie pourrait enfluencer le processus de *réanalyse* du parseur (*cf. supra*), mais pas la *décision initiale*. Enfin, l'étude souvent citée en référence de Marslen-Wilson *et al.* (1992) sur les résolutions d'ambiguïté locales permet d'avancer (Kempsen, 1999) que le parseur syntaxique utilise surtout l'information prosodique pour interpréter les *points d'attachement cruciaux*.

Comme l'avaient déjà souligné Cutler *et al.* en 1997, il est difficile, aujourd'hui encore, de tirer des conclusions définitives sur la nature exacte du traitement de l'interface prosodie/syntaxe. De l'ensemble des travaux, il se dégage cependant l'idée générale que les auditeurs possèdent la faculté d'utiliser à *bon escient* l'information prosodique pour l'analyse syntaxique, même s'ils ne le font pas systématiquement. Ces travaux montrent ainsi que ces auditeurs disposent d'une *connaissance implicite* des relations entre la syntaxe et la prosodie qui atteste pleinement de la place de celle-ci au cœur de l'interface phonologie/syntaxe de la Grammaire.

Depuis les années 70, les recherches sur le « *déchiffrage grammatical* » du système de compréhension se sont principalement centrées sur le traitement de la syntaxe, que l'on supposait constituer le *premier stade* de la compréhension de l'énoncé. Selon cette conception modulariste, que l'on pourrait qualifier d'autonomiste (ou d'obédience « *Free-syntax* » ; *cf. supra*), les informations *extra-syntaxiques* (sémantiques, pragmatiques et prosodiques) ne sont susceptibles d'exercer une influence que dans un *second stade* du traitement, ce dernier pouvant donner lieu à une réanalyse ou à une révision des choix antérieurs effectués par le parseur. Au cours des dernières années, le bien-fondé du « *modèle à deux stades* » a été remis en question. L'accumulation des données a montré en effet que le parseur est sensible à l'influence d'informations extra-syntaxiques, parmi lesquelles la prosodie occupe une place indéniable. Ces résultats plaident en faveur d'une *modalité de traitement*

interactive (Bates & MacWhinney, 1989). La question reste cependant posée de savoir si le système privilégie une *interactivité radicale* ou *relative* (Brown *et al.*, 2000). Le premier terme de l'alternative signifie qu'il n'existe pas un niveau intermédiaire de représentation syntaxique, mais des réseaux représentationnels indifférenciés. La compréhension est alors envisagée comme un *processus interactif* au cours duquel toutes les sources d'information influencent l'analyse en cours, dès qu'elles sont disponibles. Il est loisible, dans ce cas, d'envisager que le déchiffrage de la structure constituante s'effectue en une seule fois (comme cela a été suggéré récemment par Pickering *et al.*, 2002, pour la construction de cette structure sur le plan de la *production*).

Le concept *d'interactivité relative* (ou partielle) fait référence ici à un « *modèle à satisfaction de contraintes*» (*cf.* ci-dessus, les remarques au sujet de la Théorie de l'Optimalité). Ce modèle admet la nature interactive des relations entre les composants linguistiques, mais n'exclut pas l'existence de niveaux de représentation séparés. La compréhension est alors conçue comme une *compétition* fondée à la fois sur des informations *syntaxiques* et *para-syntaxiques*, parmi lesquelles il faudrait inclure les contraintes du *niveau pragmatique*, qui ne sont pas prises en compte actuellement par la Théorie de l'Optimalité. Les deux formes d'interactivité susnommées sont compatibles, selon nous, avec la proposition, que nous avons formulée plus haut, de la présence dans l'architecture cognitive d'un *compilateur prosodique*, sous la forme d'un module autonome pouvant disposer de plusieurs sources d'information simultanées. Une telle proposition n'est toutefois recevable que dans le cadre d'une théorie complète de la neurocognition du traitement du langage (*cf.* Hagoort *et al.*, 1999). Cette théorie devrait expliquer, notamment, comment le parsing s'intègre dans le processus général de compréhension de la parole et spécifier la nature et la distribution des mécanismes neuraux qui sont associés à cette intégration. Les neurosciences cognitives sont désormais en mesure d'apporter des éclaircissements sur les *niveaux de représentation* postulés par la linguistique et sur les *aspects dynamiques* des processus de compréhension qui restent encore une énigme pour la psycholinguistique (Brown *et al.*, 2000).

Épilogue

Comme il n'est pas concevable de présenter ici un résumé du contenu volontairement éclectique de cet article, nous nous bornerons à souligner l'esprit dans lequel il a été conçu. Nous nous sommes efforcé de présenter dans ce travail un éventail des problématiques que soulève l'intégration (relativement récente) de la prosodie dans le champ de la linguistique (et plus particulièrement, dans celui de la phonologie). Ces problématiques se rapportent essentiellement à des *questions d'interprétation* qui concernent : l'essence et la matérialité des formes sonores prosodiques, les fonctions que ces formes véhiculent et les relations qu'elles entretiennent avec les autres composantes de la grammaire.

Dans un souci de décloisonnement, nous avons choisi de ne pas limiter notre exposé à des considérations exclusivement « *phonologico- phonologiques* » (comme cela transparaît souvent dans des approches qui relèvent de la phonologie dite « autonomiste »). Nous avons en effet souhaité débattre d'une part des problèmes afférents aux *niveaux d'analyse* de la prosodie (ce qui soulève la question des relations entre phonétique et phonologie prosodiques) et, d'autre part, les problématiques de *l'articulation expression/contenu* et de la *fonctionnalité prosodique* (ce qui soulève la question de la gamme des significations véhiculées par la prosodie). Nous avons également accordé une place, dans la dernière partie de notre travail, à des informations émanant des recherches en psycholinguistique, car nous estimons que les entités phonologiques que décrivent les modèles linguistiques doivent être *validées psychologiquement*. L'étude récente de Vion & Colas (2004), qui propose de vérifier expérimentalement des hypothèses concernant la structuration prosodique de la question en français (Di Cristo, 1998) nous semble exemplaire de ce point de vue.

Nous ne prétendons pas avoir été exhaustif, mais seulement *informatif*. Cela signifie que nous avons cherché à attirer l'attention du lecteur sur des points qui nous paraissent névralgiques pour la compréhension du *phénomène prosodique* en général et pour celle des nombreux débats que ce phénomène suscite. Cette volonté d'information nous a poussé à accorder un place importante aux *modèles* et aux *théories* qui ont été développés récemment par la *linguistique américaine* et qui demeurent encore relativement mal connus ou mal interprétés en France. Il ne faudra donc pas s'étonner que notre texte renferme de nombreux termes anglais pour lesquels nous proposons autant que faire se peut des traductions et des explications, afin d'obvier à des difficultés de compréhension que signalent fréquemment les non-spécialistes, voire les spécialistes du domaine. Bien que nous nous soyons efforcé d'être objectif, nous n'avons pas renoncé pour autant à exposer notre point de vue en maintes occasions, aussi bien pour qui concerne les problèmes posés par l'analyse de la prosodie au regard des relations entre la phonétique et la phonologie, que pour ce qui a trait aux difficultés soulevées par l'interprétation du sens qu'exprime le jeu des contrastes prosodiques.

Nous avons renoncé à aborder dans cet article les questions qui se rapportent d'une part aux *liens entre la prosodie et les émotions* (*cf.* Arndt & Janney, 1991 ; Léon, 1993 ; Selting, 1994 ; Mozziconacci, 1998 ; Caelen-Haumont & Bel, 2001 ; Sherer & Bänziger, 2004) et, d'autre part, aux relations entre *la prosodie et la co-verbalité mimo-gestuelle (cf.* McClave, 1998 ; McNeill, 2000 ; McNeill *et al.*, 2001). Ces questions sont habituellement considérées à juste titre comme *marginales* par la phonologie. On conviendra cependant qu'elles occupent une place de plus en plus importante dans les débats actuels qui traitent du rôle de la prosodie dans la *cognitique du langage* (notamment au regard de la perspective enrichissante de la « *cognition située* »). Bien que ce rôle ne soit pas encore clairement établi, il est manifeste que les études concernées par cette nouvelle orientation motivent le foisonnement des approches

expérimentales qui cherchent à élucider la *nature des représentations mentales* de la prosodie et des liens qu'elles entretiennent avec leurs corrélats physiques. C'est à l'interprétation de ces liens que s'attache en particulier aujourd'hui la phonologie de laboratoire. Si l'investigation des corrélats physiques et leur traitement statistique a pu faire l'objet d'analyses de plus en plus abouties, les propositions se rapportant à la nature des *représentations mentales des systèmes prosodiques* demeurent hautement spéculatives et fortement dépendantes des présupposés théoriques. Nous devons convenir cependant que ces représentations ne sont que des « *représentations de représentations de représentations* » (Culioli, 1990 ; Auran, 2004), dans la mesure où elles représentent ce que les linguistes se représentent être les représentations mentales du système cognitif. Il n'est pas interdit de penser que les systèmes de représentations qui s'imposent aujourd'hui comme les plus influents et qui suscitent, de ce fait, les plus larges consensus épidémiques, se verront supplantés dans un avenir plus ou moins proche par d'autres propositions, à la faveur de l'effloraison de nouveaux paradigmes théoriques, plus largement informés par les avancées significatives des sciences de la cognition et régénérés par le développement et la mise en œuvre de nouveaux outils et de méthodologies nouvelles (jusqu'à quel point peut-on affirmer aujourd'hui avec sérénité que l'intonation est représentée mentalement par des suites de segments tonals ?). Les représentations formelles des alternances rythmiques parfaite sont-elles une réalité cognitive langagière, ou bien sont-elles simplement séduisantes pour l'esprit de celui qui les conçoit ?).

La thématique incontournable de l'articulation de la phonétique à la phonologie, soulevée au début de cet article, reste ouverte et constitue aujourd'hui encore une question centrale et non résolue de l'analyse prosodique. Il est apparu précisément que le *découplage modulaire* - classique ou révisé - des deux « disciplines » (*cf.* l'approche MESM- « *Modified Extended Standard Modularization* » - exposée par Pierrehumbert, 1994) fait problème, dans la mesure où, d'une part, il nécessite que l'on trace une ligne de démarcation franche entre les deux modules représentés par la phonétique et la phonologie et où, d'autre part, cette délimitation s'avère pour le moins épineuse. L'indétermination qu'elle suscite transparaît dans les formules métaphoriques (et poétiques !) de Pierrehumbert *et al.,* 2000, lorsque ces auteurs déclarent notamment que « *la phonétique a un parfum de phonologie et la phonologie, un parfum de phonétique* ». Afin d'apporter une modeste contribution à la résolution de ce « *dilemme olfactif* » nous avons suggéré, dans notre proposition d'interprétation des niveaux d'analyse de la prosodie, d'envisager la phonétique comme une interface entre les faits physiques et la phonologie (Hirst & Di Cristo, 1998). Il reste, bien évidemment, à définir de façon *explicite* la nature de cette interface. Parmi les propositions que nous croyons susceptibles de nous faire progresser dans cette voie, nous retiendrons en particulier celles de John Local qui invitent à accorder davantage d'importance aux « *détails phonétiques* » systématiques (laissés délibérément de

côté notamment par la phonologie prosodique classique), dans l'optique d'une modélisation des représentations cognitives *« émergentes, multi-catégorielles, multi-fonctionnelles et largement distribuées »* (Local, 2003 : 322).

Cet article reprend une partie du contenu d'un essai sur la prosodie, rédigé par l'auteur, qui paraîtra prochainement. Je remercie vivement Élisabeth Delais-Roussarie, Danielle Duez, Isabelle Marlien, Berthille Pallaud, Monique Vion, Emmanuel Bellengier, Daniel Hirst et Joël Pynte qui ont eu la patience de lire ce « texte-fleuve » et dont les remarques m'ont permis d'apporter quelques améliorations à ce travail encore perfectible.

Références bibliographiques

ABERCROMBIE, D. (1967). *Elements of General Phonetics.* Edinburgh : Edinburgh University Press.

AKMAN, V. & BAZZANELLA, C. (2003). The complexity of context. *Journal of Pragmatics,* 35 : 321-329.

ALESSANDRO (d'), C. & MERTENS, P. (1995). Automatic pitch contour stylization using a model of tonal perception. *Computer Speech and Language,* 9 (3) : 257-288.

ANDERSEN, H.L. & NØLKE, H. (2002). *Macro-syntaxe et macro-sémantique.* Berne : Peter Lang.

ANDERSON, S. (1984). A metrical interpretation of some traditional claims about quantity and stress. In Aronoff, M. & Oehrle, R.T., *Language and Sound Structure.* Cambridge (Mass.) : MIT Press : 83-106.

ANDERSON, S. (1985). *Phonology in the Twentieth Century.* Chicago : University of Chicago Press.

ARNDT, H. & JANNEY, R. (1991). Verbal, prosodic, and kinesic emotive contrasts in speech. *Journal of Pragmatics,* 15 : 521-549.

ASHER, N. (1999). Discourse structure and the logic of conversation. In Turner, K. (ed.). *The Semantics/Pragmatics Interface from Different Point of Views,* Oxford : Elsevier : 19-48.

ASHER, N. & LASCARIDES, A. (2003). *Logics of Conversation.* Cambridge University Press.

ASTESANO, C., BESSON, M. & ALTER, K. (2004). Brain potentials during semantic and prosodic processing in French. *Cognitive Brain Research,* 18 : 172-184.

AUBERGÉ, V. (1992). Developing a structured lexicon for the synthesis of prosody. In Bailly *et al.* (eds). *Talking Machines.* Amsterdam : Elsevier Science : 307-321.

AURAN, C. (2004). *Prosodie et anaphore dans le discours en anglais et en français : cohésion et attribution référentielle.* Thèse de doctorat. Université de Provence.

AURAN, C . & DI CRISTO, A. (2004). Towards an automatic annotation of temporal features in discourse : the case of syllabic duration in spontaneous French. *Proceedings of the 15th International Congress of Phonetic Sciences.* Barcelona, August 3-9.

AUTESSERRE, D. & DI CRISTO, A. (1972). Recherches psychosémantiques sur l'intonation de la phrase française. *Travaux de l'Institut de Phonétique d'Aix*, 1 : 61-98.

BARWISE, J. & PERRY, J. (1983). *Situations and Attitudes.* Cambridge (Mass.) : Cambridge University Press.

BATES, E. & MAC WHINNEY, B. (1982). Functionalist approaches to grammar, in Wanner, E. & Gleitman, L. (eds), *Language Acquisition and the State of Art,* Cambridge: Cambridge University Press : 173-217.

BATTLINER, A., KIEBLING, A., KOMPE, R., NIEMANN, H. & NÖTH, E. (1997). Tempo and its change in spontaneous speech. *Proceedings of the European Conference on Speech Communication and Technology,* (Rhodes), 2 : 763-766.

BAUM, S. & PELL, .M (1997). Production of affective and linguistic prosody by brain-damaged patients. *Aphasiology*, 11 : 177-198.

BAUM, S.R. & DWIVEDI, V. (2003). Sensitivity to prosodic structure in left-and-right-hemisphere-damaged individuals. *Brain and Language,* 87 : 278-289.

BAYBEE, J. *Phonology and Language Use.* Cambridge (UK) : Cambridge University Press.

BEACH, C.M. (1991). The interpretation of prosodic patterns at points of syntactic structure ambiguity : Evidence for cue trading relations. *Journal of Memory and Language,* 30 : 644-633.

BEATTIE, G.W. (1986). Encoding units in spontaneous speech : some implications for the dynamics of conversation, in Dechert, H.W. & Raupach, M. (eds). *Temporal Variables in Speech,* The Hague : Mouton : 131-143.

BECKMAN, M. (1986). *Stress and Non-Stress Accent.* Dordrecht : Foris.

BECKMAN, M. (1995). Local shapes and global trends. *Proceedings ICPhS* (Stockholm), 2 : 100-107.

BECKMAN, M. (1996). The parsing of prosody. *Language and Cognitive Processes,* 11 (1/2) : 17-67.

BECKMAN, M. & PIERREHUMBERT, J. (1986). Intonational structure in Japanese and English. *Phonological Yearbook,* 3 : 255-309.

BECKMAN, M. & EDWARDS, J. (1992). Intonational categories and the articulatory control of duration, in Tohkura, Y., Vatikiotis-Bateson, E. & Sagisaka, I. (eds), *Speech Perception, Production and Linguistic Structure,* Tokyo : IOS Press : 359-375.

BECKMAN, M. & AYERS, G.M. (1997). *Guidelines for ToBI labelling.* Version 3.0. Ohio State University : http ://www.ling.ohio-state.edu/phonetics/E_ToBI.

BECKMAN, M. & PIERREHUMBERT, J. (2003).Interpreting « phonetic interpretation » over the lexicon. In Local, J., Ogden, R. & Temple, R. (eds). *Phonetic Interpretation : Papers in Laboratory Phonology VI.* Cambridge University Press : 13-37.

BECKMAN, M., HIRSCHBERG, J. & SHATTUCK-HUFNAGEL, S. (2004). The original ToBI system and the evolution of the ToBI framework, in Sun-Ah Jun (ed.) Prosodic Typology- The Phonology of Intonation and Phrasing. Oxford University Press, chapter 2.

BELLENGIER, E. (2005). *Interface syntaxe-sémantique dans le cadre des Grammaires de Propriétés.* Thèse de doctorat, Université de Provence.

BERMAN, A. & SZAMOSI, M. (1972). Observations on sentential stress. *Language,* 48 : 304-332.

BERRENDONNER, A. (1991). Pour une macro-syntaxe. *Travaux de Linguistique,* 21 : 25-36.

BERRENDONNER, A. (2002). Les deux syntaxes. *Verbum,* 1-2 : 23-35.

BERTINETTO, P.M. (1989). Reflections on the dichotomy 'stress' *vs* 'syllable-timing'. *Revue de Phonétique Appliquée,* 91/93 : 99-130.

BERTRAND, R. (1998). *De l'hétérogénéité de la parole : analyse énonciative de phénomènes prosodiques et kinésiques dans l'interaction individuelle.* Thèse de doctorat : Université de Provence.

BEUN, R.J. (1994). Mental state recognition and communicative effects. *Journal of Pragmatics,* 21 : 191-214.

BEYSSADE, C., DELAIS-ROUSSARIE, E., MARANDIN, J.M., RIALLAND, A. & DE FORNEL, M. (2004). Le sens des contours intonatifs en français : croyances compatibles ou conflictuelles ? *XXVes Journées d'Etude sur la Parole* : (Fès, Maroc, avril 2004).

BIERWISCH, M. (1966). Regeln für die Intonation deutscher Sätze. *Studia Grammatica* (Berlin) 7 : 99-178.

BLACHE, P. (2001). *Les Grammaires de propriétés : Des contraintes pour le traitement automatique des langues naturelles.* Paris : Hermès Sciences.

BLACHE, P. & DI CRISTO, A. (2002). Variabilité et dépendance des composants linguistiques . *TALN 2002*, Nancy, 24-27 juin 2002.

BLAAUW, E. (1995). On the Perceptual Classification of Spontaneous and Read Speech. *OTS Publications.* University of Utrecht.

BLANCHE-BENVENISTE, C. (2002). Phrase et construction verbale. *Verbum,* 1-2 : 7-22.

BLEVINS, J. (1995). The syllable in phonological theory, in Goldsmith, J. (ed.), *The Handbook of Phonological Theory*, London, Blackwell : 206-244.

BLOOMFIELD, L. (1933). *Language.* London : Allen Unwin.

BOCK, K. (1995). Sentence production : from mind to mouth. In Miller, J.L. & Eimas, P. (eds). *Speech, Language and Communication.* New York : Academic Press : 181-216.

BOERSMA, P. (1998). *Functional Phonology : Formalizing the Interactions between articulatory and perceptual drives.* Ph.D dissertation. University of Amterdam. 504 p.

BOLINGER, D. (1951). Intonation : Levels *vs* configurations. *Word*, 7 :199-210.

BOLINGER, D. (1957). On certain functions of accents A and B. *Litera,* 4 : 80-89.

BOLINGER, D. (1958). A theory of pitch accent in English. *Word,* 14 : 109-149.

BOLINGER, D. (1961). *Generality, Gradience, and the All-or-None.* Gravenhage : Mouton.

BOLINGER, D. (1972). Accent is predictable (f you are a mind-reader). *Language,* 48 : 633-644.

BOLINGER, D. (1978). Intonation across languages. In Greenberg, J. (ed.). *Universals of Human Language,* vol. 2. Palo Alto (CAL) : Stanford University Press : 471-524.

BOLINGER, D. (1986). *Intonation and its Parts.* Palo Alto (CA) : Stanford University Press.

BOLINGER, D. (1989). *Intonation and its uses.* Palo Alto (CA) : Stanford University Press.

BOTINIS, A. (2000). *Intonation : Analysis, Modelling and Technology,* Dordrecht : Kluwer

BOTINIS, A., GRANSTRÖM, B. & MÖBIUS, B. (2001). Developments and paradigms in intonation research. *Speech Communication,* 33 (4) : 263-296.

BOOMER, D.S. (1965). Hesitation and grammatical encoding. *Language and Speech,* 8 : 148-158.

BOOMER, D.S. & LAVER, J. (1968). Slips of the tongue. *British Disorders of Communication,* 3 : 2-11.

BOUZON, C. (2004). *Rythme et structuration prosodique en anglais britannique*. Thèse de Doctorat, Université de Provence.

BRECKENRIDGE, J. (1977). The declination effect. *Journal of the Acoustical Society of America,* 60 : S90.

BRESNAN, J. (1971). Sentence stress and syntactic transformations. *Language,* 47 : 257-281.

BRESNAN, J. (1972). Stress and syntax : a reply. *Language,* 48 : 326-334.

BROWMAN, C. & GOLDSTEIN, L. (1990). Representation and reality : physical systems and phonological structure. *Journal of Phonetics,* 18 : 411-424.

BROWN, C.M., HAGOORT, P. & KUTAS, M. (2000). Postlexical integration processes in language comprehension : Evidence from brain-imaging research. In Gazzaniga, M. (ed.). *The Cognitive Neurosciences,* Cambridge (Mass.) : MIT Press.

BRUCE, G. (1977). Swedish Word Accent in Sentence Perspective. *Travaux de l'Institut de Linguistique de Lund.* Gleerup.

BYRD, D. & SALTZMAN, E. (2003). The elastic phrase : modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics,* 31 : 149-180.

CAELEN-HAUMONT, G. & BEL, B. (2001). Subjectivité et émotion dans la prosodie de la parole et du chant. *Actes du colloque international Emotions, Interaction et Développement* (Grenoble) : 141-147).

CAMPBELL, W.N. (1992). *Multi-level Timing in Speech.* PhD thesis. University of Sussex.

CAMPBELL, W.N. (1993). Durational cues to prominence and grouping, *Working Papers* (Lund University), 41 : 38-41.

CARLSON, R. & SWERTS, M. (2003). Relating perceptual judgments of upcoming prosodic breaks to F0 features. *Phonum,* 9 : 181-184.

CARLSON, R., ELENIUS, K. & SWERTS, M. (2004). Perceptual judgments of pitch range. *Speech Prosody 2004*, Nara, Japan : 689-692.

CHAFE, W. (1988). Linking intonation units. In Haiman, J. & Thomson, S. (eds). *Clause Combining in Grammar and Discourse.* Amsterdam : John Benjamins : 1-27.

CHAO, T. (2005). Fine-grained phonetic signatures of prosodic structure and their psychological reality reflected in speech comprehension. *Workshop on the Intersection between a Study of Prosody and Interaction,* Nimègue, Feb. 2005.

CHEN, M. (1990). What must phonology know about syntax ? In Inkelas, S. & Zec, D. (eds). *The Phonology-Syntax Connection.* Chicago : University of Chicago Press. :19-46.

CHOMSKY, N. (1970). Remarks on nominalization. In RA.A. Jacobs and P.S. Rosenbaum (eds). *English Transformational Grammar* : 184-221.

CHOMSKY, N. (1971). Deep structure, surface structure, and semantic interpretation. In Steinberg, D. & Jakobobovits, L. (eds). *Semantics.* Cambridge (UK) : Cambridge University Press.

CHOMSKY, N. (1995). *The Minimalist Program.* Cambridge, MA : MIT Press.

CHOMSKY, N., HALLE, MORRIS & LUKOFF, F. (1956). On accent and juncture in English. In Halle *et al.* (eds). *For Roman Jakobson*, The Hague : Mouton : 65-80.

CHOMSKY, N. & HALLE, M. (1968). *The Sound Pattern of English.* New York : Harper & Row.

CHOMSKY, N . & LASNIK, H. (1995) The theory of principle and parameters. In Chomsky, N. (ed.). *The Minimalist Program.* Cambridge, MA : MIT Press : 13-128.

CLARK, H.H. & MARSHALL, C.R. (1981). Definite reference and mutual knowledge. In Joshi, A.K., Webber, B. & Sag, I.A. (eds). *Elements of Discourse Understanding.* Cambridge : Cambridge University Press.

CLEMENTS, G.N. (1978). Tone and syntax in Ewe. In D.J. Napoli (ed.). *Elements of Tone, Stress and Intonation.* Washington DC : Georgetown University Press : 21-99.

CLEMENTS, G.N. (1990). The status of register in intonation theory. In Kingston, J. & Beckman, M.E. (eds). *Papers in Laboratory phonology*, 1 : 58-71.

CLEMENTS, G.N. & FORD, K.C. (1970). Kikuyu tone shift and its synchronic consequences. *Linguistic Inquiry,* 2 : 179-210.

COHEN, A.R., COLLIER, R. & 'T HART, J. (1982). Declination : construct or intrinsic feature of speech. *Phonetica,* 39 : 254-273.

COHEN, H., DOAIRE, J. & ELSABBAGH, M. (2001). The role of prosody in discourse processing. *Brain and Cognition,* 45 : 73-82.

COHEN, P., MORGAN, J. & POLLACK, M. (1990). *Intentions in Communication.* Cambridge (MA) : MIT Press.

CONNELL, B. (2002). Tone languages and the universality of intrinsic F0 : evidence from Africa. *Journal of Phonetics,* 30 : 101-129.

CONNELL, B. & LADD, D.R. (1990). Aspects of pitch realisation in Yoruba. *Phonology*, 7 : 1-29.

COPCEAG, D. (1966). Sur la définition de la « substance » en linguistique. *Cahiers de Linguistique Théorique et Appliquée*, 3 : 47-51.

COUPER-KUHLEN, E. (1986). *An Introduction to English prosody*. Tübingen : Niemeyer.

COUPER-KUHLEN, E. (1993). *English Speech Rhythm : Form and Function in Everyday Verbal Interaction*. Amsterdam : John Benjamins.

COUPER-KUHLEN, E. & SELTING, M. (1996). *Prosody in Conversation*. Cambridge (UK) : Cambridge University Press.

CRISTOPHE, A., GOUT, A., PEPERKAMP, S. & MORGAN, J. (2003). Discovering words in the continuous speech stream : the role of prosody. *Journal of Phonetics,* 31 : 585-598.

CROCKER, M.W., PICKERING, M. & CLIFTON, C. Jr (2000). *Architectures and Mechanisms for Language Processing*. Cambridge University Press.

CROFT, W. (1995). Intonation units and grammatical structure. *Linguistics,* 33 : 839-882.

CRUTTENDEN, A. (1997). *Intonation*. Cambridge (UK) :Cambridge University Press (second edition).

CRYSTAL, D. (1969). *Prosodic Systems and Intonation in English*. Cambridge : Cambridge University Press.

CRYSTAL, D. (1975). *The English Tone of Voice : Essays in Intonation, Prosody and Paralanguage*. London : Arnold.

CRYSTAL, D. (1991). *A Dictionary of Linguistics and Phonetics*. Oxford : Blackwell.

CULIOLI, A. (1990). La linguistique: de l'empirique au formel. In Culioli, A. *Pour une linguistique de l'énonciation : opérations et représentations,* Paris: Oprhys: 9-46.

CUTLER, A (1997). Prosody and the structure of message. In Sagisaka, Y., Campbell, N. & Higushi, N. (eds). *Computing Prosody*. New York : Springer : 63-66.

CUTLER, A. & LADD, D.R. (1983). *Prosody : Models and Measurements*. Berlin : Springer.

CUTLER, A. & CLIFTON, C. Jr (1999). Comprehending spoken language : a blueprint of the listener. In Brown, C. & Hagoort, P. (eds). *The Neurocognition of Language*. Oxford : Oxford University Press : 123-166.

CUTLER, A., DAHAN, D. & VAN DONSELAR, W. (1997). Prosody in the comprehension of spoken language : A litterature review. *Language and Speech,* 40 : 141-201.

DAINORA, A. (2001a). *An Empirically Based Probabilistic Model of Intonation in American English*. Ph D. Dissertation. University of Chicago.

DAINORA, A. (2001b). Eliminating downstep in prosodic labeling of American English. In Bacchiani, M., Hirschberg, J., Litman, D. & Ostendorf, M. (eds). *Proceedings of the Workshop on Prosody in Speech Recognition and Understanding* (oct. 22-24). Molly Pitcher Inn, Red Bank, NJ : 41-46.

DAINORA, A. (2002). Does intonational meaning come from tones or tunes ? Evidence against a compositional approach. In Bel, B. & Marlien, I. (eds). *Proceedings of Speech Prosody 2002* (Aix-en-Provence, 11-13 april 2002) : 235-238.

DAINORA, A. (2004). Modeling intonation in English : a probabilistic approach to phonological competence. In Goldstein, L.M., Whalen, D. & Best, C.T. (eds). *Papers in Laboratory Phonology* 8 : Mouton de Gruyter (forthcoming).

DAUER, R.M. (1982). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics,* 11 : 51-62.

DE BOYSSON-BARDIES, B . (1996). *Comment la parole vient aux enfants.* Paris : Odile Jacob.

DELAIS-ROUSSARIE, E. (1994). *Pour une approche parallèle de la structure prosodique.* Thèse de Doctorat : Université de Toulouse-Le-Mirail.

DELAIS-ROUSSARIE, E. (1995). Rythme et structure prosodique en français : In Lyche, C. (ed.). *French Generative Phonology : Retrospective and Perspectives.* Salford : AFLS/ ESRI : 131-150.

DELAIS-ROUSSARIE, E. (1996). Phonological phrasing and accentuation in French. In Nespor, M. & Smith, N. (eds). Dam Phonology. *HIL Phonology Papers* II, La Haye : Holland Academic Graphics : 1-38.

DELAIS-ROUSSARIE, E. (2000). Vers une nouvelle approche de la structure prosodique. *Langue Française,* 126 : 92-112.

DELAIS-ROUSSARIE, E. (2004). Interface phonologie/syntaxe : des domaines phonologiques à l'organisation de la grammaire. In Durand, J. Nguyen, N. & Rey, V. (eds). *Manuel de Phonologie et Phonétique.* Paris : Hermès (sous presse).

DELAIS-ROUSSARIE, E. & CHOI-JONIN, I. (2004). Existent-il des indices intonatifs de segmentation en unités macro-syntaxiques ? *Journées d'Etude sur la Parole* (Fès, Maroc, avril 2004).

DELAIS-ROUSSARIE, E. & FOUGERON, C. (2004). Les notions de domaine et de constituant en phonétique et en phonologie. *XXVes Journées d'Etude sur la Parole* (Fès, Maroc, avril 2004).

DELATTRE, P. (1966). Les dix intonations de base du français. *The French Review*, 40 (1) : 1-14.

DELL, F. (1984). L'accentuation dans les phrases en français. In Dell, F., Hirst, D.J. & Vergnaud, J.R. (eds). *Forme Sonore du Langage.* Paris : Hermann : 65-122.

DI CRISTO, A. (1971). *Intonation et unités discrètes.* Manuscrit, Université de Provence.

DI CRISTO, A. (1976a). Indices prosodiques et structure constituante. *Cahiers de Linguistique, d'Orientalisme et de Slavistique*, 7 : 27-40.

DI CRISTO, A. (1976b). Des indices prosodiques aux traits perceptuels : application d'un modèle d'analyse prosodique à l'étude du vocatif en français. *Travaux de l'Institut de Phonétique d'Aix*, 3 : 213-358.

DI CRISTO, A. (1978). *De la microprosodie à l'intonosyntaxe.* Thèse de doctorat d'état. Université de Provence. Publications de l'Université de Provence, 1982.

DI CRISTO, A. (1998). Intonation in French, in Hirst, D.J. & Di Cristo, A. (eds), *Intonation Systems : A Survey of Twenty Languages,* Cambridge University Press : 195-218.

DI CRISTO, A. (1999). Le cadre accentuel du français : essai de modélisation. *Langues*, vol. 2, n°3 : 184-205 et vol. 2, n°4 : 258-269.

DI CRISTO, A. (2000a). Une Grammaire Ecologique comme cadre interprétatif de la prosodie de la parole. Communication, *3rd All-Nordic and Baltic Congress on Semiotics* (June, 12-21), Imatra : Finland.

DI CRISTO, A. (2000b). La problématique de la prosodie dans l'étude de la parole dite spontanée. *Parole* (15/16) : 189-249.

DI CRISTO, A. (2000c). Interpréter la prosodie. *Actes des 23èmes Journées d'Etude sur la Parole* (Aussois, juin 2000) : 13-30.

DI CRISTO, A. (2002). Prosodie et discours : modélisation et codage de la constituance prosodique. *Verbum,* 1-2 : 37-53.

DI CRISTO, A. (2003). De la métrique et du rythme de la parole ordinaire : l'exemple du français, in Bordas, E. (ed.). *Le Rythme de la prose,* Presses Universitaires Franc-Comtoises : 25-44.

DI CRISTO, A. & CHAFCOULOFF, M. (1981). L'intonème progrédient en français : caractéristiques intrinsèques et extrinsèques. *Studia Phonetica*, 18 : 39-51.

DI CRISTO, A. & HIRST, D.J. (1986). Modelling French micromelody : analysis and synthesis. *Phonetica*, 43 : 11-30.

DI CRISTO, A. & HIRST, D. (1993). Rythme syllabique, rythme mélodique et représentation hiérarchique de la prosodie du français. *Travaux de l'Institut de Phonétique d'Aix*, vol.15 : 13-23.

DI CRISTO, A. & HIRST, D. (1997). *L'accentuation non-emphatique en français: stratégies et paramètres. Polyphonie pour Ivan Fónagy.* Paris : l'Harmattan : 71-101.

DI CRISTO, A. & HIRST, D. (2002). De l'acoustique à la phonologie, représentations et notations de l'intonation : une application au français. In Braun, A. & Masthoff, R. (eds). *Phonetics and its Applications.* Marburg : Franz Steiner Verlag : 313-322.

DI CRISTO, A., HIRST, D.J., BOUDOURESQUES, N. et LOUIS, M. (2002). Écrire l'intonation : le système INTSINT, fondements théoriques et illustrations. *Revue Parole,* 22/23/24 : 175-209.

DI CRISTO, A. & JANKOWSKI, L. (1999). Prosodic organisation and phrasing after focus in French. *Proceedings of the XIVth International Congress of Phonetic Sciences* (San Francisco): 1565-1568.

DI CRISTO, A. & DI CRISTO, P. (2001). Syntaix : une approche métrique-autosegmentale de la prosodie. *Traitement Automatique des Langues,* 42 : 69-114.

DI CRISTO, A., AURAN, C., BERTRAND, R., CHANET, C. & PORTES, C. (2004). Outils prosodiques et analyse du discours. *Cahiers de l'Institut de Linguistique de Louvain*, 30 : 27-84.

DIK, S. (1989). *Functional Grammar.* Dordrecht : Foris.

D'IMPERIO, M. (2000). Language-specific and universal constraints on tonal alignment : the nature of targets and anchors. *Speech Prosody 2002* (Aix-en-Provence) : 101-106.

D'IMPERIO, M. & GILI FIVELA, M. (2000). How many levels of phrasing ? Evidence from two varieties of italian. In Local, J., Ogden, R. & Temple, R. (eds). *Phonetic Interpretation : Papers in Laboratory Phonology VI*. Cambridge University Press : 130-143.

DOGIL, G., ACKERMANN, H., GRODD, W., HAIDER, H., MAYER, J., RIECKER, A. & WILDGRUBER, D. (2002). The speaking brain: A tutorial introduction to MRI experiments in the production of speech, prosody and syntax. *The Journal of Neurolinguistics,* 15: 59-90.

DRAKE, C. (1998). Psychological processes involved in the temporal organization of complex auditory sequences : universal and acquired processes. *Music Perception*, 16 (1) : 11-26.

DUEZ, D. (1987). *Contribution à l'étude de la structuration temporelle de la parole en français*. Thèse de docorat d'état : Université de Provence.

DUEZ, D. (1997). La signification des pauses dans la production et la perception de la parole. *Parole,* 3/4/ : 275-299.

DUEZ, D. (2001). Signification des hésitations dans la production et la perception de la parole spontanée. *Parole,* 17/18/19 : 113-137).

DURAND, J. (1990). *Generative and non-linear phonology*. London : Longmans

DURAND, J. (1993). La phonologie multidimensionnelle moderne et la description du français. *Working Papers in Language and Linguistics* (Salford University), 17 : 50p.

FANT, G., NORD, L. & KRUCKENBERG, A. (1987). Segmental and prosodic variabilities in connected speech. *Proceedings of the XIth ICPhS,* 6 : 102-105.

FERREIRA, F. (1993). Creation of prosody during sentence production. *Psychological Review,* 100 : 233-253.

FERREIRA, F., ANES, M.D. & HORINE, M.D. (1996). Exploring the use of prosody during language comprehension using the auditory moving window technique. *Journal of Psycholinguistic Research,* 25 (2) : 273-290.

FIRTH, J.R. (1948). *Sounds and Prosodies*. Transactions of the Philological Society : 127-152.

FODOR, J.A. (1983). *The Modularity of Mind*. Cambridge (Mass.) : MIT Press.

FODOR, J.A. & BEVER, T.G. (1965). The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior,* 4 : 414-420.

FODOR, J.D. (1998). Learning to parse ? *Journal of Psycholinguistic Research,* 27 :285-319.

FODOR, J.D. (2002). Psycholinguistics cannot escape prosody. *Proceedings of Speech Prosody 2002* (Aix-en-Provence, 11-13 april 2002) : 83-88.

FÓNAGY, I. (1983). *La vive voix*. Paris : Payot.

FÓNAGY, I. (1989). La Cendrillon de Dwight Bolinger. *Semiotica,* 76 (3/4) : 217-244.

FONTANEY, L. (1991). À la lumière de l'intonation. In Kerbrat Orecchioni, C. (ed.). *La question.* Lyon : PUL : 113-161.

FOUGERON, C. (1998). *Variations articulatoires en début de constituants prosodiques de différents niveaux en français.* Thèse de Doctorat. Université de Paris III.

FOUGERON, C. (1999) Prosodically conditioned articulatory variations : a review. *U.C.L.A. Working Papers in Phonetics*, 97 : 1-68.

FOX, A. (1973). Tone sequences in English. *Archivum Linguisticum,* 4 : 17-26.

FOX, A. (2000). *Prosodic Features and Prosodic Structure.* Oxford : Oxford University Press.

FRAISSE, P. (1974). *Psychologie du rythme.* Paris : Presses Universitaires de France.

FRAZIER, L. (1990). Exploring the architecture of the language system. In Altmann, G.T.M. (ed.). *Cognitive Models of Speech Processing : Psycholinguistic and Computational Perspectives.* Cambridge (Mass.) : MIT Press : 409-433.

FRAZIER, L., CLIFTON, C.Jr (1996). *Construal.* Cambridge (Mass.) : MIT Press.

FRAZIER, L., CLIFTON, C.Jr & CARLSON, K. (2004). Don't break, or do : prosodic boundary preferences. *Lingua,* 114 : 3-27.

FUCHS, C. & LE GOFFIC, P. (1992). *Les linguistiques contemporaines.* Paris : Hchette.

FUJISAKI, H. & NAGASHIMA, S. (1969). A model for the synthesis of pitch contours of connected speech. *Annual Report of the Engineering Research Institute.* Faculty of Engineering, University of Tokyo, 28 : 52-60.

GARDE, P. (1966). Forme et contexte en syntaxe. *La Linguistique*, 1, 1-15.

GARDE, P. (1968). *L'Accent.* Paris : Presses Universitaires de France.

GARDING ; E. (1983). A generative model of intonation. In Cutler, A. & Ladd, D.R. *Prosody : Models and Measurements.* Heidelberg : Springer : 11-25.

GARDING ; E. (1998). Intonation in Swedish. In Hirst, D.J. & Di Cristo, A. (1998). *Intonation Systems : A Survey of Twenty Languages.* Cambridge (UK) : Cambridge University Press : 112-130.

GARRETT, M.F., BEVER, T.G. & FODOR, J.A. (1966). The active use of grammar in speech perception. *Perception and Psychophysics,* 1 : 30-32.

GEE, J.P. & GROSJEAN, F. (1983). Performance structures : A psycholinguistic and linguistic appraisal. *Cognitive Psychology,* 15 : 411-458.

GIVÓN, T. (1995). *Functionalism and Grammar.* Amsterdam : John Benjamins.

GOBL, C. & NI CHASAIDE, A. (1992). Acoustic characteristics of voice quality. *Speech Communication,* 11 : 481-490.

GOLDSMITH, J. (1976a). *Autosemental Phonology.* Ph.D. dissertation, Cambridge, MA : MIT. Reproduced by the Indiana Linguistic Club.

GOLDSMITH, J. (1976b). An overview of autosegmental phonology. *Linguistic Analysis*, 2 : 23-68.

GOLDSMITH, J. (1990). *Autosegmental & Metrical Phonology*. Oxford : Blackwell.

GOLDSMITH, J. (1995). *The Handbook of Phonological Theory*. Oxford : Blackwell.

GOLDSMITH, J. (1995b). la phonologie actuelle. Ecole Thématique sur Les Fondements et les Perspectives en Traitement Automatique des Langues. (Luminy, 1995).

GOLDSMITH, J. (1999). Phonological Theory : The Essential Readings. Oxford : Blackwell Publishers.

GRABE, E., NOLAN, H. & FARRAR, K. (1998). IViE- a comparative transcription system for intonational variation in English. *Proceedings of the 5th international Conference on Spoken Language Processing*. Sydney, Australia).

GREIMAS, A.& COURTÈS, J. (1979). Sémiotique : *Dictionnaire raisonné de la théorie du langage*. Paris : Hachette.

GRICE, M. (1995). Leading tones and downstep in English. *Phonology,* 12 : 183-233.

GRICE, M., LADD, D.R. & ARVANITI, A. (2000). On the place of phrase accents in intonational phonology. *Phonology,* 17 : 143-185.

GRØNNUM, N. (1995). Superposition and subordination in intonation : a non-linear approach. *Proceedings ICPhS* (Stockholm), 2 : 124-131.

GRØNNUM, N. (1998). Intonation in Danish. In Hirst, D.J. & Di Cristo, A. (1998). *Intonation Systems : A Survey of Twenty Languages*. Cambridge (UK) : Cambridge University Press : 131-151.

GROSZ, B. & SINDER, C. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics,* 12 : 175-2004.

GUSSENHOVEN, C. (1984). *On the Grammar and Semantics of Sentence Accents*. Dordrecht : Foris.

GUSSENHOVEN, C. (1988). Adequacy in intonation analysis : the case of Dutch. In van der Hulst, H. & Smith, N. *Autosegmental Studies on Pitch Accent*, Dordrecht : Foris : 95-121.

GUSSENHOVEN, C. (1992). Intonational phrasing and the prosody hierarchy : *Phonologica* 1988 : 89-99.

GUSSENHOVEN, C. (1999). Discreteness and gradience in intonation contrasts. *Language and Speech,* 42 (2/3) : 283-305.

GUSSENHOVEN, C. (2000). The boundary tones are coming : on the nonperipheral realization of boundary tones. In Broe, M. & Pierrehumbert, J. (eds). *Papers in Laboratory Phonology V*. Cambridge : Cambridge University Press : 132-151.

GUSSENHOVEN, C. (2002). Intonation and interpretation : Phonetics and Phonology. *Proceedings of Speech Prosody 2002* (Aix-en-Provence) : 45-57.

GUSSENHOVEN, C. (2003). Perceiving paralinguistic meaning. *Proceedings of IP 2003* (Nantes) : 47-49.

GUSSENHOVEN, C. & RIETVELD, A.C.M. (1988). Fundemantal frequency declination in Dutch : Testing three hypotheses. *Journal of Phonetics,* 16 : 355-369.

HAGÈGE, C. (1985). *L'homme de paroles.* Paris : Fayard.

HAGOORT, P., BROWN, C.M. & OSTERHOUT, L. (1999). The neurocognition of syntactic processing. In Brown, C.M. & Hagoort, P. (1999). *The Neurocognition of Language.* Oxford : Oxford University Press : 273-316.

HALLE, M. & VERGNAUD, J.R. (1987). *An Essay on Stress.* Cambridge (MA) : MIT Press.

HALLE, M. & KENSTOWICZ, M. (1991). The free element condition and cyclic versus noncyclic stress. *Linguistic Inquiry,* 22 : 457-501.

HALLE, M. & IDSARDI, W. (1995). General properties os stress and metrical structure. In Goldsmith, J. (ed.). *The Handbook of Phonological Theory.* Oxford : Blackwell : 403-443.

HALLIDAY, M.A.K. (1967). *Intonation and Grammar in British English.* The Hague : Mouton.

HAMMOND, M. (1997). Optimality theory and prosody. In Archangeli, D. & Langendoen, D.T. (eds). *Optimality theory : an overview.* London : Blackwell : 33-58.

HANDEL, S. (1989). *Listening : An Introduction to the Perception of Auditory Events.* Cambridge (MA) : MIT Press.

'T HART, J. & COLLIER, R. (1975). Integrating different levels of intonation analysis. *Journal of Phonetics,* 3 : 235-255.

'T HART, J., COLLIER, R. & COHEN, A. (1990). *A Perceptual Study of Intonation : An Experimental-Phonetic Approach,* Cambridge (UK) : Cambridge University Press.

HANNAHS, S.J. (1995). The phonological word in French. *Linguistics*, 33 : 1125-1144.

HAYES, B. (1980) . *A Metrical Theory of Stress Rules.* PhD dissertation : MIT. Distributed by the Indiana Linguistic Club, 1981.

HAYES, B. (1982). Extrametricality and English stress. *Linguistic Inquiry,* 13 : 227-276.

HAYES, B. (1984). The phonology of rhythm in English. *Linguistic Inquiry,* 15 : 33-74.

HAYES, B. (1988). Metrics and phonological theory. In Newmeyer, F.J. (ed.). *Linguistics : The Cambridge Survey,* Cambridge : Cambridge University Press, vol. 2 : 220-249.

HAYES, B.(1989). The prosodic hierarchy in meter. In Kiparsky, P. & Youmans, G. (eds). *Phonetics and Phonology*, vol. 1. Orlando : Academic Press : 201-260.

HAYES, B. (1990). Precompiled phrasal phonology. In Inkelas, S. & Zec, D. (eds). *The Phonology-Syntax Connection.* Chicago : University of Chicago Press : 85-108.

HAYES, B. (1991). *Metrical Stress Theory.* Manuscrit : University of California at Los Angeles (UCLA).

HAYES, B. (1994). « Gesture » in Prosody. In Keating, P. (ed.). *Phonological structure and phonetic forms.* Cambridge (UK) : Cambridge University Press : 64-75.

HAYES, B. (1995). *Metrical Stress Theory : Principles and Case Studies.* The University of Chicago Press.

HAYES, B. & LAHIRI, A. (1991). Bengali Intonational Phonology. *Natural Language and Linguistic Theory*, 9 : 47-96.

HAZAËL-MASSIEUX, C. (1983). Le rôle de l'intonation dans la définition et la structuration de l'unité de discours. *Bulletin de la Société de Linguistique de Paris*, 78 : 99-160.

HELDNER, M. & MEGYESI, B. (2003) Exploring the prosody-syntax interface in conversations. *15th IsPhS.,* Barcelona.

HELSLOOT, K. (1993). The prosodic word : an endless domain ? *Proceedings of the Institute of Phonetic Sciences* (University of Amsterdam), 17 : 129-146.

HELSLOOT, K. (1995). *Metrical Prosody.* PhD dissertation. University of Amsterdam.

HERMES, D. & VAN GESTEL, J.C. (1991). The frequency scale of speech intonation. *Journal of the Acoustical Society of America*, 90 : 97-102.

HIRST, D.J. (1977). *Intonative Features.* The Hague : Mouton.

HIRST, D.J. (1987). *La représentation linguistique des systèmes prosodiques : une approche cognitive.* Thèse de Doctorat d'Etat. Université de Provence.

HIRST, D.J. (1998). Intonation in British English. In Hirst, D.J. & Di Cristo, A. (1998). *Intonation Systems : A Survey of Twenty Languages.* Cambridge (UK) : Cambridge University Press : 56-77.

HIRST, D.J. & DI CRISTO, A. (1984). French intonation : A parametric approach. *Die Neueren Sprachen*, 83 : 564-569.

HIRST, D.J. & DI CRISTO, A. (1998). *Intonation Systems : A Survey of Twenty Languages.* Cambridge (UK) : Cambridge University Press.

HIRST, D.J., DI CRISTO, A. & ESPESSER, R. (2000). Levels of representation and levels of analysis for the description of intonation systems. In Horne, M. (ed.). *Prosody : Theory and Experiment.* Dordrecht : Kluwer : 51-87.

HJELMSLEV, L. (1953).*Prolegomena to a theory of language.* Madison : University of Wisconsin Press.

HOBBS, J.R. (1979). Coherence and coreference. *Cognitive Science,* 3 : 67-90.

HOCKETT, C. (1942). A system of descriptive phonology. *Language,* 23 : 321-343.

HOCKETT, C. (1947). Componential analysis of Sierra Populuca. *International Journal of American Linguistics*, 13 : 259-267.

HOCKETT, C. (1955). *A Manual of Phonology.* Baltimore : Waverley.

HOGG, R. & MCCULLY, C.B. (1987). *Metrical Phonology : A Course Book.* Cambridge : Cambridge University Press.

HOMBERT, J.M. (1976). Phonetic explanation of the development of tones from prevocalic consonants. *Working Papers in Phonetics,* 33 : 23-39.

HONDA, K. (2004). Physiological factors causing tonal characteristics of speech : from global to local prosody. *Speech Prosody 2004* (Nara, Japan) : 739-744.

HOUSE, D. (1990). *Tonal Perception in Speech,* Lund University Press.

HYMAN, L. (1985). Word domains and downstep in Bamileke-Dschang. *Phonology Yearbook*, 2 : 45-82.

IDSARDI, J. (1992). *The Computation of Prosody.* PhD Dissertation : MIT.

INDEFREY, P. & LEVELT, W.J.M. (2000). The neural correlates of language production. In Cazzaniga, M.S. (ed.). *The New Cognitive Neurosciences.* Cambridge (Mass.) : MIT Press : 845-865.

INKELAS, S. & LEBEN, W. (1990). Where phonology and phonetics intersect : the case of Hausa intonation. In Kingston, J. & Beckman, M.E. (eds). *Papers in Laboratory Phonology*, 1 : 17-34.

INKELAS, S.& ZEC, D. (1990). *The Phonology-Syntax Connection.* Chicago : University of Chicago Press.

INKELAS, S. & ZEC, D. (1995). Syntax-phonology interface. In Goldsmith, J .A. (ed.). *The Handbook of Phonological Theory.* Oxford : Blackwell : 535-549.

JACKENDOFF, R. (1972). *Semantic Interpretation in Generative Grammar.* Cambridge, Mass : MIT Press.

JACKENDOFF, R. (1977). *X-bar Syntax.* Cambridge, Mass. : MIT Press.

JACKENDOFF, R. (1983). *Semantics and Cognition.* Cambridge (Mass.) : MIT Press.

JASSEM, W. (1952). Stress in modern English. *Bulletin de la Société Linguistique Polonaise,* 11 : 23-49.

JEANNERET, M. (1999). *La coénonciation en français.* Bern : Peter Lang.

JESCHENIAK, J.D., HAHNE, A. & FRIEDRICI, A. (1998). Brain activity patterns suggest prosodic influences on syntactic parsing in the comprehension of spoken utterances. *Music and Perception,* 16 : 55-62.

JOSHI, A.K. (1982). The role of mutual beliefs in question-answer systems. In Smith, N. (ed.). *Mutual Knowledge.* Academic Press : New York.

JOSHI, A.K. (1990). Phrase structure ans intonational phrases. In Altmann, G.M. (ed.). *Cognitive Models of Speech Processing.* Cambridge, Mass. : MIT Press : 513-531.

JUN, S.A. (1998. The accentual phrase in the Korean prosodic hierarchy. *Phonology,* 15 : 189-226.

JUN, S.A. & FOUGERON, C. (2000). A phonological model of French intonation. In Botinis, A. (ed.). *Intonation : Analysis, Modelling and Technology.* Dordrecht : Kluwer Academic Publishers : 209-242.

JUSCZYK, P. (1997). *The Discovery of Language.* Cambridge, Mass. : MIT Press.

KAGER, R. (1995). The metrical theory of word stress. In Goldsmith, J. (ed.), *The Handbook of Phonological Theory,* London : Blackwell : 206-244.

KAGER, R. (1999). *Optimality Theory.* Cambridge (U.K.) : Cambridge University Press.

KAGER, R. & ZONNEVELD, W. (1999). Phrasal phonology : an introduction. In Kager, R. & Zonneveld, W. (eds). *Phrasal Phonology*, Nijmegen University Press : 1-34.

KAISSE, E. (1985). Connected Speech : *The Interaction of Syntax and Phonology.* Orlando : Academic Press.

KAMP, H. (1981). A theory of truth and semantic representation. In Groenendjik, J.A., Janssen, T.M. & Stockhof, M.B. (eds). *Formal Methods in the Study of Natural Languages*. Amsterdam : Matematiech Centrum, Tract 133 : 277-322.

KAMP, H. & REYLE, U. (1993). *From Discourse to Logic*. Dordrecht : Kluwer.

KARCEVSKIJ, S. (1934). Sur la phonologie de la phrase. *Travaux du Cercle de Linguistique de Prague,* 4 : 188-227.

KAYE, J. (1989). *Phonology: a Cognitive View*. Hillsdale, N.J : Lawrence Erlbaum Associates Publishers.

KEATING, P. (1985). Universal phonetics and the organization of grammars. In Fromkin, V. (ed .) *Phonetic Linguistics,* New York : Academic Press : 115-132.

KEATING, P. (1994). *Phonological Structure and Phonetic Form. Papers in Laboratory Phonology III*. Cambridge (UK) : Cambridge University Press.

KEATING, P., CHO, T., FOUGERON, C. & HSU, C.S. (2003). Domain-initial articulatory strengthening in four languages. In Local, J. ; Ogden, R. & Temple, R. (eds). *Papers in Laboratory Phonology VI*. Cambridge University press : 144-162.

KEMPSEN, G. (1999). Sentence parsing. In Frederici, A. (ed.). *Language Comprehension*. New York : Springer : 211-223.

KEMPSON, R. (1988). Grammar and conversational principles. In Newmeyer, F.J. *Linguistic Theory : The Cambridge Survey,* vol . Cambridge (UK) : Cambridge University Press : 139-163.

KINGSTON, J. & BECKMAN, M. (1990). *Papers in Laboratory Phonology I*. Cambridge : Cambridge University Press.

KLINGHARDT, H. & KLEM, G. (1920). *Übungen im englischen tonfall*. Cöthen.

KOHLER, K.J. (2003). Domains of temporal control in speech and language from utterance to segment. *15th ICPhS* (Barcelona).

LACHERET-DUJOUR, A. & BEAUGENDRE, F. (1999). *La prosodie du français*. Paris : Editions du CNRS.

LACHERET-DUJOUR, A. & VICTORRI, B .(2002). La période intonative comme unité d'analyse pour l'étude du français parlé : modélisation prosodique et enjeux linguistiques. *Verbum,* 1-2 : 56-72.

LADD, D.R. (1984). Declination : a review and some hypothesis. *Phonology Yearbook,* 1 : 53-74.

LADD, D. R. (1986). *The Structure of Intonational Meaning*. Evidence from English. Bloomington : Indiana University Press.

LADD, D. R. (1986). Intonational phrasing : the case for recursive prosodic structure. *Phonology* 3 : 311-340.

LADD, D.R. (1990). Metrical representation of pitch register. In Kingston, J. & Beckman, M.E. (eds). *Papers in Laboratory Phonology*, 1 : 35-57.

LADD, D.R. (1992). An introduction to intonational phonology. In Docherty, J. & Ladd, D.R. (eds). *Papers in Laboratory Phonology* 2. Cambridge, Cambridge University Press : 321-334.

LADD, D.R. (1993). Notes on the phonology of prominence. *Working Papers* (University of Lund), 41 : 10-15.

LADD, D .R. (1994). Constraints ont he gradient variability of pitch range, or, pitch level 4 lives ! In Keating, P. (ed.). *Phonological Structure and Phonetic Forms*. Cambridge University Press : 43-63.

LADD, D.R. (1995). Linear and overlay descriptions: an autosegmental-metrical middle way*? Proc. ICPhS*. (Stockholm) : 116-123.

LADD, D.R. (1996). *Intonational Phonology*. Cambridge UK) : Cambridge University Press.

LADD, D.R. & MORTON, R. (1997). The perception of intonational emphasis : continuous or categorical. *Journal of Phonetics,* 25 : 313-342.

LADD, D.R. (2000). Bruce, Pierrehumbert, and the elements of intonational phonology. In Horne, M. (ed.). *Prosody : Theory and Experiments*. Dordrecht : Kluwer Academic Publishers : 37-50.

LAKOFF, G. (1972). The global nature of the nuclear stress rule. *Language,* 48 : 285-303.

LAMBRECHT, K. (1994). *Information Structure and Sentence Form*. Cambridge : Cambridge : Cambridge University Press.

LANIRAN, Y.O. (1992). *Intonation in Tone Languages : The Phonetic Implementation of Tones in Yoruba*. PhD. thesis : Cornell University.

LANIRAN, Y.O. & CLEMENTS, G.N. (2003). Downstep and high raising : interactinc factors in Yoruba tone production. *Journal of Phonetics,* 31 : 203-250.

LASHLEY, K.S. (1951). The problem of serial prder in behavior. In Jeffress, L.A. (ed.). *General Mechanisms in Behavior*. New York : Wiley.

LAKOFF, G. (1987). *Women, Fire, and Dangerous Things*. University of Chicago Press.

LAKS, B. (1993). La constituance revisitée. In Laks, B. & Plénat, M. (eds). *De natura sonorum* : 173-221.

LAKS, B. & PLÉNAT, M. (1993). *De natura sonorum*. Presses Universitaires de Vincennes.

LAVE, J. (1988). *Cognition in Practice*. Cambridge (UK) : Cambridge University Press.

LAVER, J. (1970). The production of speech. In Lyons, J. (ed.). *New Horizons in Linguistics*. Harmondsworth : Penguin : 53-75.

LAVER, J. (1980). *The Phonetic Description of Voice Quality*. Cambridge (UK) : Cambridge University Press.

LEBEN, W. (1973) . *Suprasegmental Phonology*. PhD. dissertation, Cambridge (MA) : MIT.

LEBEN, W. (1976). The tones in English intonation. *Linguistic Analysis*, 2 : 69-107.

LEBEN, W. (1982). Metrical or autosegmental. In van der Hulst, H. & Smith, N. (eds). *The Structure of Phonological Representations*. Dordrecht : Foris : 177-190.

LEHISTE, Ilse (1970). *Suprasegmentals.* Cambridge (MA) : MIT Press.

LEHISTE, I. (1973). Phonetic disambiguation of syntactic ambiguity. *Glossa,* 7 : 107-121.

LEHISTE, I., OLIVE, J. & STREETER, L. (1976). Role of duration in disambiguating syntactically ambiguous sentences. *Journal of the Acoustical Society of America,* 60 : 1199-1202.

LÉON, P. (1993). *Précis de phonostylistique : parole et expressivité,* Paris : Nathan Université.

LERDAHL, F. & JACKENDOFF, R. (1983). *A Generative Theory of Tonal Music.* Cambridge (MA). MIT Press.

LEVELT, W.J.M. (1978). A survey of studies in sentence perception : 1970-1976. In Levelt, W.J.M. & Flores d'Arcais, G.B. (eds). *Studies in the Perception of Language,* Chichester : John Wiley .

LEVELT, W.J.M. (1989). *Speaking : From Intention to Articulation.* Cambridge (Mass.) : MIT Press.

LEVELT, W.J.M. (1999). Producing spoken language : a blueprint of the speaker. In Brown, C. & Hagoort, P. (eds). *The Neurocognition of Language.* Oxford : Oxford University Press : 84-122.

LEVINSON, S.C. (1997). *Presumptive Meanings : The Theory of Generalized Conversational Implicature.* Nijmegen : Max Planck Institute.

LIBERMAN, A.M., COOPER, F.S., HARRIS, K.S. & MACNEILAGE, P.F. (1963). A motor theory of speech perception. *Journal of the Acoustical Society of America,* 35 : 1114.

LIBERMAN, A.M. & MATTINGLY, I.G. (1985). The motor theory of speech perception revised. *Cognition,* 21 : 1-36.

LIBERMAN, M. (1975). T*he Intonational System of English.* PhD dissertation. Massachusetts Institute of Technology.

LIBERMAN, M. & SAG I. (1974*). Prosodic form and discourse function. Proceedings of the Cicago Linguistic Society*, 10 : 416-27.

LIBERMAN, M. & PRINCE, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry* 8 : 249-336.

LIBERMAN, M. & STREETER, L.A. (1978). Use of non-sense syllable mimicry in the study of prosodid phenomena. *Journal of the Acoustical Society of America,* 63 : 231-233.

LIBERMAN, M. & PIERREHUMBERT, J. (1984). Intonational invariance under changes in pitch range and length. In Aronoff, M. & Oerhle, R. (eds). *Language Sound Structure.* Cambridge, MA : MIT Press : 157-233.

LIEBERMAN, P. (1965). On the acoustic basis of the perception of intonation by linguists. *Word*, 21 : 40-54.

LIEBERMAN, P. (1967). *Intonation, Perception and Language.* Cambridge (Mass) : MIT Press.

LILJENCRANTS, J. & LINDBLOM, B. (1972). Numerical simulation of vowel quality systems : the role of perceptual contrasts. *Language,* 48 : 839-862.

LINDBLOM, B. (1990). Explaining phonetic variation : a sketch of the H&H theory. In Hardscastle, W.J. & Marchal, A. (eds). *Speech Production and Speech Modelling.* Dordrecht : Kluwer: 403-439.

LOCAL, J. (2003). Variable domains and variable relevance : interpreting phonetic exponents. *Journal of Phonetics,* 31 : 321-339.

LOCAL, J. , OGDEN, R. & TEMPLE, R. (2003). Phonetic interpretation : *Papers in Laboratory Phonology VI*. Cambridge : Cambridge University Press.

LOUIS, M., DI CRISTO, A., HABIB, M. & HIRST, D. (2003). Transcription segmentale et suprasegmentale d'un jargon phonémique. *Revue Parole*, 23/24 : 245-266.

MAEDA, S. (1976). *A Characterization of American English Intonation.* PhD thesis, MIT.

MACWHINNEY, B. & BATES, E. (1989). *The Crosslinguisic Study of Sentence Processing,* Cambridge :Cambridge University Press.

MALMBERG , Bertil (1966). Analyse des faits prosodiques - problèmes et méthodes. *Cahiers de Linguistique Théorique et Appliquée*, 3 : 99-107.

MANN, W.C. & THOMPSON, S.A. (1988). Rhetorical Structure Theory : Toward a functional theory of text organization. *Text,* 8 (3) : 243-281.

MARANDIN, J.M., BEYSSADE, C., DELAIS-ROUSSARIE, E., RIALLAND, A. & DE FORNEL, M. (2004). *The meaning of final contours in French* (à paraître).

MARSLEN-WILSON, W.D., TYLER, L.K., WARREN, P. , GRENIER, P. & LEE, C.S. (1992). Prosodic effects in minimal attachment. *Quarterly Journal of Experimental Psychology,* 45 : 73-87.

MARTIN, J.G. (1972). Rhythmic (hierarchical) *versus* serial structure in speech and other behaviors. *Psychological Review,* 79 : 487-509.

MARTIN, P. (1977a). Syntax and intonation : an integrated theory. *Toronto Semiotic Circle Monographs* (Victoria University) n°2 : 57p.

MARTIN, P. (1977b). Résumé d'une théorie de l'intonation. *Bulletin de l'Institut de Phonétique de Grenoble*, 6 : 57-87.

MARTIN, P. (1987). Prosodic and rhythmic structures in French. *Linguistics,* 25 : 925-949.

MARTINET, A. (1949). *Phonology as a functional phonetics.* Londres : Oxford University Press.

MCCARTHY, J.J. (1982). *Nonlinear phonology : an overview.* Glow Newsletter.

MCCARTHY, J. & PRINCE, A. (1993). Generalized alignment. In Booij, G. and van Marle, J. (eds). *Yearbook of Morpholoby.* Dordrecht : Kluwer : 79-153.

MCCLAVE, E. (1998). Pitch and manual gestures. *Journal of Psycholinguistic Research,* 27 (1) : 69-89.

MCNEILL, D. (2000). *Language an Gesture.* Cambridge (UK) : Cambridge University Press.

MCNEILL, D., QUEK, F., MC CULLOUGH, K.E., DUNCAN, S., FURUYAMA, N., BRYLL, R., MA, X.F. & ANSARI, R. (2001). Catchments, prosody and discourse. *Gesture,* 1 (1) : 9-33.

MEL'CUK, I. (2001). *Communicative Organization in Natural Language.* Amsterdam : John Benjamins Publishing Company.

MERTENS, Piet (1987). *L'intonation du français : de la description linguistique à la reconnaissance automatique*. Thèse de doctorat, Université de Louvain.

MERTENS, P. (2002). Synthesizing elaborate intonation contours in text-to-speech for French. *Speech Prosody 2002* (Aix-en-Provence) : 499-502.

MERTENS, P. (2004). Le prosogramme : une transcription semi-automatique de la prosodie. *Cahiers de l'Institut de Linguistique de Louvain,* 30 (1-3) : 7-25.

MERTENS, P., GOLDMAN, J.P., WEHRLI, E. & GAUDINAT, A. (2001). La synthèse de l'intonation à partir de structures syntaxiques riches, *Traitement Automatique des Langues,* 42 (1) : 145-192.

MEYNADIER, Y. (2003). *Interaction entre prosodie et (co)articulation linguopalatale en français*. Thèse de Doctorat, Université de Provence.

MILLER, J.& WEINERT, R. (1998). *Spontaneous Spoken Language*. Oxford : Clarendon Press.

MÖBIUS, B. (1995). Components of a quantitative model of German intonation. *Proceedings of the XIIIth International Congress of Phonetic Sciences* (Stockholm), vol.2 : 108-115.

MOREL, M.A. & DANON-BOILEAU, L. (1998). *Grammaire de l'intonation ; l'exemple du français,* Paris : Ophrys.

MOREL, M.A. (2000). Intonation, coénonciation, colocution, formulation. In Berthoud, A.C. & Mondada, L. (eds), *Modèles du discours en confrontation,* Berne : Peter-Lang ; 53-78.

MORGAN, J.L. & DEMUTH, K. (1996). Signal to syntax : an overview. In Morgan, J.L. & Demuth, K. (eds). *Signal to Syntax : Bootstrapping from Speech to Grammar in Early Acquisition*. Mahwah, N.J. : Lawrence Erlbaum Associates : 1-22.

MOUNIN, G. (1974). *Dictionnaire de la Linguistique*. Paris : Presses Universitaires de France.

MOXEY, L. & SANFORD, A. (1993). *Communicating Quantities : A Psychological Perspective*. Hillsdale : Erlbaum.

MOZZICONACCI, S. (1998). *Speech Variability and Emotion : Production and Perception*. Eindhoven : The Printers of the University of Eindhoven.

NAZZI, T., BERTONCINI, J. & MEHLER, J. (1998). Language discrimination by newborns : towards an understanding of the role of rhythm. *Journal of Experimental Psychology,* 24.

NESPOR, M. & VOGEL, I. (1983). Prosodic structure above the word. In Cutler, A. & Ladd, D.R. (eds). *Prosody : Models and Measurements*. Berlin : Springer : 123-139.

NESPOR, M. & VOGEL, I. (1986). *Prosodic Phonology*. Dordrecht : Foris.

NEUHOFF, J.G. (2003). Pitch variation is unnecessary (and sometimes insufficient) for the formation of auditory objetcs. *Cognition*, 87 : 219-224.

NÍ CHASAIDE, A. & GOBL, C. (2004). Voice quality and F0 in prosody : towards a holistic account. *Speech Prosody 2004* (Nara, Japan) : 189-196.

NICOL, J.L. & PICKERING, M.J. (1993). Processing syntaxtically ambigous sentences ; evidence from semantic priming. *Journal of Psycholinguistic Research,* 22 (2) : 207-237.

NOLAN, F. & GRABE, E. (1997). Can ToBI transcribe intonational variation in British English ? *Proceedings of the ESCA Workshop* (Athens, sept. 18-21th) : 259-262.

O'CONNOR, J.D. & ARNOLD, G.F. (1961). *Intonation of Colloquial English.* London : Longmans.

ODDEN, D. (1995). Tone : African languages. In Goldsmith, L. (ed.). *The Handbook of Phonological Theory.* Oxford : Blackwell : 444-494.

ODÉ, C. (2003). Toward ToRI, a manual transcription system of russian intonation. *Proceedings 25*, Institute of Phonetics, University of Amsterdam : 1-14.

OHALA, J.J. (1978). Production of tone. In Fromkin, V. (ed.). *Tone : A Linguistic Survey.* New York : Academic : 5-39.

OHALA, J.J. (1983). Cross-language use of pitch : An ethological view. *Phonetica,* 40 :1-18.

OHALA, J.J. (1990). There is no interface between phonetics and phonology. *Journal of Phonetics*, 18 : 153-171.

OHALA, J.J. (1991). The integration of phonetics and phonology. *Proceedings of the XIIth ICPhS* :2-16.

PALMER, H.E. (1922). *English intonation with systematic exercices.* Cambridge : Heffers.

PASDELOUP, V. (1990). *Modèle de règles rythmiques du français appliqué à la synthèse de la parole.* Thèse de doctorat : Université de Provence.

PATTERSON, D. (2000). *A linguistic approach to pitch range modelling.* PhD. thesis : University of Edinburgh.

PICKERING, M.J., BRANIGAN, H.P. & MCLEAN, J.F. (2002) Constituent structure is formulated in one stage. *Journal of Memory and Language,* 46 : 586-605.

PIERREHUMBERT, J. (1979). The perception of fundamental frequency declination. *Journal of the Acoustical Society of America,* 66 (2) : 363-369.

PIERREHUMBERT, J. (1980). *The Phonology and Phonetics of English Intonation.* Ph.D. dissertation, Cambridge (MA) : MIT.

PIERREHUMBERT, J. (1990). Phonological and phonetic representation. *Journal of Phonetics,* 18 : 375-394.

PIERREHUMBERT, J. (1994). Knowledge of variation. In Beals, K. *et al.* (eds.). *Papers from the 30th Meeting of the Chicago Linguistic Society,* vol. 2. : 232-256.

PIERREHUMBERT, J. (2000). Tonal elements and their alignment. In Horne, M. (ed.). *Prosody : Theory and Experiments.* Dordrecht : Kluwer Academic Publishers : 11-36.

PIERREHUMBERT, J. (2001). Stochastic phonology. *Glot International*, 6 : 195-207.

PIERREHUMBERT, J. (2002). *Probabilistic phonology : discrimination and robustness.* www.ling.nwu.edu.

PIERREHUMBERT, J. & BECKMAN, M. (1988). *Japanese Tone Stucture.* Cambridge, MA : MIT Press.

PIERREHUMBERT, J. & HIRSCHBERG, J. (1990). The meaning of intonational contours in the interpretation of discourse. In Cohen, P.R., Morgan, J. & Pollack, M.E. (eds). *Intentions in Commmunication.* Cambridge (MA) : MIT Press.

PIERREHUMBERT, J. , BECKMAN, M. & LADD, D.R. (2000). Conceptual foundations of phonology as a laboratory science. In Burton-Roberts, N., Carr, P. & Docherty, G. (eds). *Phonological Knowledge : Its Nature and Status.* Oxford University Press.

PIKE, K. (1945). *The Intonation of American English.* Ann Arbor : University of Michigan Press.

POEPPEL, D. & HICKOK, G. (2004). Towards a new functional anatomy of language, *Cognition,* 92 : 1-12.

POLLOCK, J.Y. (1997). *Langage et cognition : introduction au programme minimaliste de la grammaire générative.* Paris : Presses Universitaires de France.

PORTES, C. (2004). *Prosodie et économie du discours : spécificité phonétique, écologie discursive et portée pragmatique de l'intonation d'implication.* Thèse de doctorat, Université de Provence.

PORTES, C. & DI CRISTO, A. (2003). Pitch range in spontaneous speech : semi-automatic approach *versus* subjective judgment. *Proceedings of the 15th International Congress of Phonetic Sciences.* Barcelona, August 3-9.

POST, B. (2000). *Tonal and Phrasal Structures in French Intonation.* Thesus, The Hague : Holland Academic Graphics.

PRIETO, P., SHIH, C. & NIBERT, H. (1996). Pitch downtrend in Spanish. *Journal of Phonetics,* 24 : 445-473.

PRINCE, A. (1983). Relating to the grid. *Linguistic Inquiry*, 14 : 19-100.

PRINCE, A. & SMOLENSKY, P. (1993). *Optimality theory : Constraint interaction in generative grammar.* Ms., Rutgers University and Johns Hopkins University.

PULLUM, G. & ZWICKY, A. (1988). The syntax-phonology interface. In F.J. Newmeyer (ed.). *Linguistics : The Cambridge Survey.* Cambridge University Press : 255-280.

PURSON , A. & DI CRISTO, A. (1998). Aspects prosodiques et pragmatiques de la demande de confirmation en français. *Travaux de l'Institut de Phonétique d'Aix,* 18 : 113-126.

PYNTE, J. (1998). The role of prosody in semantic interpretation. *Music Perception,* 16(1) : 79-97.

PYNTE, J. & PRIEUR, J. (1996). Prosodic breaks and attachment decisions in sentence parsing. *Language and Cognitive Processes,* 11 : 165-191.

RAMUS, F., NESPOR, M. & MEHLER, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition,* 73 : 265-292.

RAMUS, F., DUPOUX, E. & MEHLER, J. (2003). The psychological reality of rhythm classes : perceptual studies. *Proceedings of 15th ICPhS* (Barcelona).

REMIJSEN, B. & VAN HEUVEN, V. (1999). Gradient and categorical pitch dimensions in Dutch : diagnostic test. *Proceedings of the XIVth ICPhS* (San Francisco) : 1865-1868.

RIALLAND, A. (1997). Le parcours du « downstep », ou l'évolution d'une notion. *Bulletin de la Société de Linguistique de Paris,* XCII (1) : 207-243.

RIETVELD, A.C.M. (1980). Word boundaries in the French language. *Language and Speech,* 23 : 289-296.

ROCA, I. (1988). Constraining extrametricality. *Phonologica* : 239-248.

ROSSI, M. (1971a). Le seuil de perception des glissandos (variations mélodiques). *Phonetica,* 23 : 1-33.

ROSSI, M. (1971b). L'intensité spécifique des voyelles. *Phonetica,* 24 : 129-161.

ROSSI, M. (1977). L'intonation et la troisième articulation. *Bulletin de la Société Linguistique de Paris,* LXXII (1) : 55-68.

ROSSI, M. (1978). Interaction of intensity glides and frequency glissandos. In Honor to D.B. Fry. *Language and Speech,* 21 (4) : 384-396.

ROSSI, M. (1981). Prosodical aspects of speech production. *Proceedings FASE* (Venezia) : 125-158.

ROSSI, M. (1987). Peut-on prédire l'organisation prosodique du langage spontané ? *Etudes de Linguistique Appliquée* : 20-48.

ROSSI, M. (1999). *L'intonation : le système du français.* Paris : Ophrys.

ROSSI, M. (2000). Intonation : past, present future. In Botinis, A. (ed.). *Intonation : Analysis, Modelling and Technology*, Dordrecht : Kluwer : 13-52.

ROSSI, M. & DI CRISTO, A. (1980). Un modèle de détection automatique des frontières intonatives et syntaxiques. *Actes des XIèmes Journ ées d'Etude sur la Parole* (Strasbourg) : 213-238.

ROSSI, M., DI CRISTO, A., HIRST, D., MARTIN, P. & NISHINUMA, Y. (1981). *L'intonation : de l'acoustique à la sémantique.* Paris : Klincksieck.

ROULET, E., FILLIETTAZ, L. & GROBET, A. (2001). *Un modèle et un instrument d'analyse de l'organisation du discours.* Bern : Peter Lang.

SABIO, F. (1996). *Description prosodique et syntaxique du discours en français.* Thèse de Doctorat. Université de Provence.

SAGISAKA, Y., CAMPBELL, N. & HIGUCHI, N. (1997). *Computing Prosody.* New-York : Springer.

SANDERMAN, A. (1996). *Prosodic Phrasing. Production, Perception, Acceptability and Comprehension.* Ph D. thesis : Eindhoven University of Technology.

SANDERMAN, A. & COLLIER, R. (1997). Prosodic phrasing and comprehension. *Language and Speech,* 40 (4) : 391-408.

SCHAFER, A. (1996) : *Prosodic parsing : The role of prosody in sentence comprehension.* PhD. thesis, Amherst : University of Massachusetts.

SCHAFER, A., CARTER, J., CLIFTON, C. & FRAZIER, L. (2000). Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research,* 29 : 169-182.

SCHERER, K.R. & BÄNZIGER, T. (2004). Emotional expression in prosody. *Speech prosody 2004,* (Nara, Japan) : 359-366.

SCHIRMER, A. (2004). Timing speech : a review of lesion and neuroimaging findings. *Cognitive Brain Research,* 21 : 269-287.

SCHMERLING, S.(1976). *Aspects of English Sentence Stress.* Austin : University of Texas Press.

SCHUETZE-COBURN, S., SHAPLEY, M. & WEBER, E.G. (1991). Units of intonation in discourse : A comparison of acoustic and auditory analyses. *Language and Speech,* 34 : 207-234.

SELKIRK, E.O. (1972). *The Phrasal Phonology of English and French.* Ph.D. dissertation, MIT.

SELKIRK, E.O. (1977). The French foot : on the status of " mute " e. Paper presented at the colloquium on *Current Issues in French Phonology* (Indiana University, sept. 28-29, 1977).

SELKIRK, E.O. (1978). On prosodic structure and its relation to syntactic structure. In Fretheim, T. (ed.). *Nordic Prosody,* 2, Trondheim TAPIR : 111-140.

SELKIRK, E.O. (1980). The role of prosodic categories in English word stress. *Linguistic Inquiry*, 11 : 563-605.

SELKIRK, E.O. (1981). On the nature of phonological representation. In Myers, T., Laver, J. & Anderson, J. (eds). *The Cognitive Representation of Speech.* Amsterdam : North Holland : 379-388.

SELKIRK, E.O. (1984). *Phonology and Syntax : The Relation between Sound and Structure.* Cambridge, MA : MIT Press.

SELKIRK, E.O. (1986). On derived domains in sentence phonology. *Phonology Yearbook*, 3 : 317-405.

SELKIRK, E.O. (1995a). Sentence prosody : intonation, stress and phrasing. In Goldsmith, J. (ed.). *The Handbook of Phonological Theory.* Oxford : Blackwell : 550-569.

SELKIRK, E.O. (1995b). The prosodic structure of function words. In Beckman, J., Urbanczyk, S. & Walsh, L. (eds). *Optimality Theory Occasional Papers,* Amherst (Mas), UMOP 18 : 439-470.

SELKIRK, E.O. (2000). The interaction of constraints on prosodic phrasing. In Horne, M. (ed.). *Prosody : Theory and Experiment.* Dordrecht : Kluwer Academic Publishers : 231-261.

SELKIRK, E.O. & SHEN, T. (1990). Prosodic domains in Shanghai Chinese. In Inkelas & Zec (eds). *The Phonology-Syntax Connection.* Chicago : University of Chicago Press : 313-338.

SELTING, M. (1994). Emphatic speech style with special focus on the prosodic signalling of heightened emotive involvement in conversation. *Journal of Pragmatics,* 22 : 375-408.

SHATTUCK-HUFNAGEL, S. (2000). Phrase-level phonology in speech production planning : evidence for the role of prosodic structure. In Horne, M. (ed.). *Prosody : Theory and experiment.* Dordrecht : Kluwer : 201-229.

SHATTUCK-HUFNAGEL, S. & TURK, A. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, 25 (2) : 193-247.

SILVERMAN, K., BECKMAN, M., PITRELLI, J. , OSTENDORF, M., WIGHTMAN, C., PRICE, P., PIERREHUMBERT, J. & HIRSCHBERG, J. (1992). ToBI : a standard for labeling English prosody. *Proceedings of the 2nd International Conference on Spoken Language Processing*. Banff (Alberta), vol. 2 : 867-870.

SIMON, A.C. (2003). *Structuration prosodique du discours en français : une approche multidimensionnelle et expérientielle*. Berne : Peter Lang.

SLUIJTER, A.M.C. & VAN HEUVEN, V. (1996). Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America,* 100 (4) : 2471-2485.

SMITH, M. (2000). Conceptual structures in language production. In Wheeldon, L. (ed.). *Aspects of Language Production*. London : Psychology Press : 331-374.

SNEDEKER, J. & TRUESWELL, J. (2003). Using prosody to avoid ambiguity : effects of speaker awareness and referential context. *Journal of Memory and Language,* 48 : 103-130.

SPERBER, D. & WILSON, D. (1986). *Relevance : Communication and Cognition*. Oxford : Blackwell.

STEEDMAN, M. (1990). Syntax and intonational structure in a Combinatory Grammar. In Altmann, G.T.M. (ed.). *Cognitive Models of Speech Processing*. Cambridge (Mass.), MIT Press : 457-482.

STEEDMAN , M. (1991). Structure and Intonation. *Language*, 67 : 260-296.

STEINHAUER, K., ALTER, K. & FRIEDERICI, A. (1999). Brain responses indicate immediate use of prosody in natural speech processing. *Nature Neuroscience,* 2 : 191-196.

STEVENS, K. (1989). On the quantal nature of speech. *Journal of Phonetics*, 17 : 3-45.

STIRLING, L. & WALES, R. (1996). Does prosody support or direct sentence processing ? *Language and Cognitive Processes,* 11 : 193-212.

STOCKWELL, R.P. (1960). The place of intonation in a generative grammar of English, *Language,* 36 : 360-367.

STOCKWELL, R.P. (1972). The role of intonation : reconsiderations and other considerations. In Bolinger, D. (ed). *Intonation.* Harmondsworth (UK) : Penguin Books : 87-109.

SWERTS, M., STRANGERT, M. & HELDNER, M. (1996). F0 declination in spontaneous and read-aloud speech. *TMH Quarterly Progress Status Report,* 2 : 23-24.

SWERTS, M. & HIRSCHBERG, J. (1998). Prosody and conversation : an introduction. *Language and Speech,* 41 (3-4) : 229-233.

SWINNEY, D. & LOVE, T. (1998). The processing of discontinuous dependencies in language and music. *Music Perception,* 16 (1) : 63-78.

TAYLOR, P. (1994). The Rise/Fall/Connection model of intonation. *Speech Communication,* 15 (2) : 169-186.

THORSEN, N. (1983). Two issues in the prosody of standard danish. In Cutker, A. & Ladd, D.R. (eds). *Prosody : models and measurements.* Heidelberg : Springer : 27-38.

THORSEN, N. (1985). Intonation in text in standard Danish. *Journal of the Acoustical Society of America,* 77 : 1205-1216.

TITZE, I.R. (1989). On the relation between subglottal pressure and fundamental frequency in phonation. *Journal of the Acoustical Society of America,* 85 (2) : 901-906.

TRAGER, G.L. (1958). Paralanguage : a first approximation. *Studies in Linguistics,* 13 (1-2) : 1-12.

TRAGER, G.L. (1964). Paralanguage and other things. *Le Maître Phonétique*, 122 : 21-23.

TRAGER, G.L. & SMITH, H.L. (1951). An Outline of English Structure. *Studies in Linguisitics Occasional Papers* 3. Norman, Oklahoma : Battenburg Press.

TROUBETZKOY, N.S. (1933). La phonologie actuelle. *Journal de Psychologie,* XXX : 227-246.

TRUCKENBRODT, H. (1995). *Phonological Phrases : their Relation to Syntax, Prominence and Focus.* Doctoral Dissertation, MIT.

TRUCKENBRODT, H. (1999). On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry,* 30 : 219-255.

TROUBETZKOY, N. (1939). *Principes de phonologie.* Paris : Klincksieck.

UMEDA, N. (1982). F0 declination is situation-dependent. *Journal of Phonetics,* 10 : 279-290.

VAISSIÈRE, J. (1975). On French prosody. *MIT Quarterly Progress report*, 114 et 115 : 212-223 et 251-261.

VAISSIÈRE, J. (1983). Language independant prosodic features. In Cutler, A. & Ladd, D.R. (eds). *Prosody : models and measurements.* Heidelberg : Speinger : 53-66.

VAISSIÈRE, J. (2004). Perception of intonation. In Pisoni, D.B. & Remez, R.E. (eds). *Handbook of Speech Perception.* Oxford : Blackwell. Forthcoming.

VALLDUVÍ, E. (1990). *The Informational Component.* Ph.D. dissertation : Philadelphia : University of Pennsylvania. Distributed by Garland (New-York).

VALLDUVÍ, E. & ENGDAHL, E. (1996). The linguistic realization of information packaging. *Linguistics*, 35 : 459-519.

VAN DER HULST, H. & SMITH, N. (1982). *The Structure of Phonological Representations.* Dordrecht : Foris.

VANDERSLICE, R. & LADEFOGED, P. (1971). Binary suprasegmental features. *UCLA Working Papers in Phonetics,* 17 : 6-24.

VAN DONZEL, M.E. & KOOPMANS-VAN BEINUM, F.J. (1995). Evaluation of discourse structure on the basis of written *vs* spoken material. *Proceedings of the XIIIth International Congress of Phonetic Sciences,* 3 : 258-261.

VAN HEUVEN, V. (1994). What is the smallest prosodic domain ? In Keating, P. (ed.). Phonological Structure an Phonetic Form. *Papers in Laboratory Phonology III*. Cambridge (UK) : Cambridge University Press : 76-98.

VAN LANCKER, D. & SIDITIS, J.J. (1992). The identification of affective-prosody stimuli by left-and-right-hemisphere-damaged subjects : All errors are not created equal. *Journal of Speech and Hearing Research,* 35 : 963-970.

VAN SANTEN, J.P. (2002). Quantitative modeling of pitch alignment. *Speech Prosody 2002* (Aix-en-Provence) : 107-112.

VERLUYTEN, P. (1982). *Recherches sur la prosodie et la métrique du français*. Thèse de doctorat : Université d'Anvers.

VERSCHUEREN, J. (1994). The pragmatic perpesctive. In Verschueren, J., Östman, J.O. & Blommaert, J. (eds). *Handbook of Pragmatics,* Amsterdam: John Benjamins: 1-19.

VION, M. & COLAS, A. (2004). Pitch cues for the identification of yes-no questions in French. *Soumis pour publication*.

VOGEL, I. & KENESEI, I. (1990). Syntax and Semantics in Phonology. In Ikelas, S. & Zec, D. (eds). *The Phonology-Syntax Connection*. Chicago : Chicago University Press : 339-363.

VOGEL, I., BUNNELL, H.T. & HOSKINS, S. (1995). The phonology and phonetics of the rhythm rule. In Connell, B. & Arvanti, A. (eds). Phonology and Phonetic Evidence : *Papers in Laboratory Phonology IV,* Cambridge University Press : 111-140.

VON HEUSINGER, K. (1999). *Intonation and Information Structure*. Habilitationsschrift. Faculty of Philosophy : University of Konstanz.

WATT, S.M. & MURRAY, W. (1996). Prosodic form and parsing commitments. *Journal of Psychological Research,* 25 (2) : 291-318.

WARREN, P. (1985). *The Temporal Organization and Perception of Speech*. Doctoral dissertation, UK : University of Cambridge.

WARREN, P. (1996). Prosody and parsing : An introduction. *Language and Cognitive Processes,* 11 (1/2) : 1-16.

WATT, D.L. (1994). *Phonology and Semology of Intonation : An Instrumental and Systemic Perspective*. Bloomington : Indiana University Linguistic Club Publications.

WENK, B.J. & WIOLAND, F. (1982). Is French really syllable-timed ? *Journal of Phonetics,* 10 : 193-216.

WELLS, R. (1945). The pitch phonemes of English. *Language*, 21 : 27-40.

WHEELDON, L. & LAHIRI, A. (1997). Prosodic units in speech production. *Journal of Memory and Language,* 37 : 356-381.

WHITE, L.S. (2002). *English speech timing : A domain and locus approach*. PhD dissertation : University of Edinburgh.

WICHMANN, A. (2000). *Intonation in text and discourse*. London : Longman.

WICHMANN, A. (2002). Intonation and the inferential process, *Proceedings of Speech Prosody 2002,* Aix-en-Provence : 11-15.

WINGFIELD, A. (1975). The intonation-syntax interaction : prosodic features in perceptual processing of sentences. In Cohen, A. & Nooteboom, S.G. *Structure and Process in Speech Perception,* New York : Springer : 146-160.

WOO, N. (1969). *Prosody and Phonology.* PhD thesis. MIT. Distributed by IUCL.

XU, Y. (2002). Articulatory constraints and tonal alignement. *SpeechProsody 2002* (Aix-en-Provence) : 91-100.

XU, Y. & XUEJING, S. (2002). Maximum speed of pitch change and how it may relate to speech, *Journal of the Acoustical Society of America,* 111 (3) : 1399-1413.

XU Y. (2004). Understanding tone from the perspective of production and perception. Submitted to : *Language and Linguistics.*

YULE, G. (1980). Speaker's topics and major paratones. *Lingua*, 52: 33-47.

ZATORRE, R.J., EVANS, A.C, & MEYER, E. & GJEDDE, A.(1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science,* 256 : 846-849.