

CILL 30 (2004) n° 1-3, 27-84

OUTILS PROSODIQUES ET ANALYSE DU DISCOURS

**Albert DI CRISTO, Cyril AURAN, Roxane BERTRAND,
Catherine CHANET, Cristel PORTES, Alain REGNIER**
Laboratoire Parole et Langage, Université de Provence

1. INTRODUCTION

Le présent article se propose de décrire les outils qui sont en cours de développement au sein du groupe de recherche PRODIGE¹ dont l'objectif est d'étudier le rôle de la prosodie dans l'élaboration et l'interprétation du discours. Nous montrerons dans ce travail que le but principal à atteindre, au cours de la première phase de ce programme, se rapporte à l'édification d'une grille d'analyse plurilinéaire permettant de confronter des représentations discursives (formelles et conceptuelles) et des représentations prosodiques (physico-formelles), sans préjuger à l'origine de la nature de leurs incidences réciproques. Il apparaîtra en effet que la conception de cette grille procède davantage d'une heuristique que de l'adhésion préalable à un modèle figé, ce qui nous semble plus adapté à la problématique que nous abordons dans le cadre de notre étude : la description d'un réseau d'interfaces dont nous ne maîtrisons pas au départ la pluralité des entités interfacées.

Il nous incombera de justifier le bien-fondé du sigle PRODIGE (PROsodie, DIscours et Grammaire Écologique) que nous avons adopté pour désigner notre groupe de travail. Compte tenu des débats que suscite le concept même de discours, nous devons aussi nous prononcer sur la façon dont nous l'appréhendons et

¹ Le groupe de recherche PRODIGE est un groupe interdisciplinaire qui s'est constitué au sein du Laboratoire Parole et Langage de l'Université de Provence et qui rassemble à ce jour des chercheurs et des doctorants de phonétique, de linguistique et de psycholinguistique.

spécifier dans quelle mesure cette conception exerce une influence sur notre démarche théorique et méthodologique. Il en ira de même de la prosodie qui demeure souvent, en dépit de l'intérêt croissant qui motive son étude, un concept flou, ou pour le moins fluctuant. Nous tenterons de montrer que les entités constitutives des systèmes prosodiques relèvent de divers niveaux interprétatifs et qu'elles doivent de ce fait être l'objet d'une représentation plurilinéaire. Il apparaîtra ainsi que la description du réseau d'interfaces formes/fonctions appliquée à l'analyse du discours prend appui sur la confrontation de deux ensembles de représentations plurilinéaires, dévolues respectivement à la prosodie et au discours, et matérialisées sous la forme d'une grille d'analyse intégrative.

L'articulation de cet article est formée de trois parties. Nous exposons, dans la première, notre approche du discours et le cadre interprétatif qui s'y rapporte. Cette première partie s'achève par une présentation de la composition de la grille d'analyse réservée à la notation des niveaux dits « textuels » et/ou « discursifs ». Dans la seconde partie, nous exposons notre conception de la prosodie ainsi que les démarches qui nous ont conduits à élaborer un modèle de représentation plurilinéaire des systèmes prosodiques. La constitution de cette représentation fait l'objet d'une description qui clôture cette seconde partie. La troisième partie est consacrée à des illustrations de la mise en pratique de la grille complète en vue de l'interprétation des relations de la prosodie au discours. Les exemples qui constituent cette partie illustrative sont extraits du corpus Fayard-Groult, qui a été mis à notre disposition par le Département de Linguistique de l'Université de Genève où s'est tenu en septembre 2002 le *Colloque Prosodie* qui a donné lieu à des échanges particulièrement fructueux.

2. LE DISCOURS : CONCEPTION GENERALE, CADRE INTERPRETATIF ET CODAGE « TEXTUO-DISCURSIF »

2.1. Approches conceptuelles du discours

2.1.1. Aspects consensuels

Le concept de discours tel que l'envisagent les linguistes de l'oralité a fait l'objet de plusieurs interprétations qui ne manquent pas d'évoquer des prises de position conflictuelles. Un consensus relatif paraît cependant s'être établi, au fil des débats, sur au moins quelques-unes de ses propriétés substantielles, notamment sur le fait que le discours est le produit d'une interaction communicative (ROULET 1999) qui met en jeu des modalités d'expression diversifiées, verbales et paraverbales

(MCNEILL 2000) ; ou sur l'affirmation que le discours ne se résout pas au texte que l'enchaînement des énonciations élabore pro-activement (BERRENDONNER 1993), et encore moins à une simple juxtaposition d'énoncés (PASSONNEAU & LITMAN 1996) ; ou encore sur l'hypothèse que, à la différence de la compétence linguistique, la compétence discursive qui régule la conduite du discours ne peut être exprimée par un système de règles (AUCHLIN 1999), bien que cette affirmation ne soit pas incompatible au demeurant avec une recherche des régularités observables dans les productions discursives (ROULET *et al.* 2001). Enfin, le point de vue selon lequel, parallèlement au déploiement de son architecture structurelle, essentielle à son interprétation (ASHER 1995, 1999), le discours participe de diverses activités et exhibe des actions (LEVINSON 1979 ; BANGE 1992 ; VERNANT 1997 ; FILLIETTAZ 2002 ; EDWARDS 1997), contribue aussi à activer l'évolution d'un paradigme de l'analyse du discours – et des discours (REBOUL & MOESCHLER 1998) – qui invite à réexaminer la trame des relations que tissent ses composants.

2.1.2. Les modèles de discours et la prosodie

Dans l'optique relationnelle que l'on vient d'évoquer, et en n'écartant *a priori* aucune des propositions consensuelles précédentes, nous considérons la prosodie comme un composant de premier plan dont on commence tout juste à percevoir la pluralité des rôles qu'il est en mesure d'assumer (WICHMANN 2000). Il n'en demeure pas moins que les théories et les modèles de discours les plus influents et les plus aboutis, tels que la RST (MANN & THOMPSON 1988, 2001), la SDRT (ASHER 1995) ou encore le modèle genevois (ROULET *et al.* 2001) n'ont pas encore donné lieu à des tentatives ciblées sur son intégration systématique. Il importe cependant de mentionner les travaux de VAN DONZEL & KOOPMANS-VAN BEINUM (1995) sur la prosodie du néerlandais, qui se réfèrent à la théorie de la RST ; ceux de SIMON (2004) en ce qui concerne le français et la référence au modèle genevois, et ceux de HIRSCHBERG & PIERREHUMBERT (1986), AVESANI & VAYRA (1988), PIERREHUMBERT & HIRSCHBERG (1990), HIRSCHBERG & GROSZ (1992) qui ont en commun d'appliquer à l'analyse prosodique de l'anglais le modèle computationnel de la structure de discours conçu par GROSZ & SIDNER (1986).

2.1.3. Considérations théoriques sur le discours et fondements de l'approche du groupe PRODIGE

Les travaux du groupe PRODIGE se rattachent à un axe de recherche qui s'oriente vers la description et la modélisation de l'interface prosodie/discours. Toutefois, à la différence des études citées précédemment, notre approche ne présuppose aucunement l'adhésion à un modèle global préétabli du discours. En conséquence, son ambition se limite pour l'instant à mettre en évidence, en regard des relations de

la prosodie au discours, certaines formes de régularités qui seraient susceptibles de confirmer ou d'infirmer des hypothèses propres à des modèles de discours préexistants et/ou d'amener à en proposer de nouvelles.

2.1.3.1. Discours et cognition située

En prenant appui sur l'idée que l'élaboration et l'interprétation du discours sont principalement concernées par la mise en correspondance de structures conceptuelles et de représentations linguistiques et paralinguistiques, nous précisons d'abord que notre approche s'intègre dans la perspective cognitive de la « cognition située » (LAVE 1988). Nous considérons en effet que le contenu conceptuel mis en jeu dans le discours comprend non seulement le « contenu idéationnel », mais également le « contenu expérientiel », qui inclut le contexte, l'affect et divers aspects de la perception. Nous adoptons ainsi les présupposés théoriques qui sous-tendent le modèle de la « sémantique cognitive » (TALMY 2000). De même, nous adhérons à la mouvance de la « pragmatisme de la sémantique » qui propose, comme l'a récemment souligné PEREGRIN (1999), de substituer à la conception triadique classique (syntaxe/sémantique/pragmatique) de la théorie du langage (telle qu'elle est formulée par MORRIS 1938 et par CARNAP 1956), celle que soutiennent des auteurs comme : WITTGENSTEIN (1953), DAVIDSON (1984) ou BRANDON (1994), et selon laquelle le langage est d'abord un outil (ou mieux : une boîte à outils) et une manifestation de l'activité humaine. Le langage instancié par le discours est une entreprise coopérative et son élaboration ne peut être construite par le travail autonome d'items indépendants. D'autre part, la signification du discours est une « construction interprétative » (ABEL 1994) qui consiste à établir comment une expression peut s'avérer utile à l'accomplissement des objectifs qui justifient l'utilisation du langage. On pourra évoquer ici la notion de « schématisation discursive » proposée par GRIZE (1996) pour exprimer l'idée que le discours se met en scène sous la forme de la construction de la représentation d'une certaine réalité. Le terme de « schématisation » désigne un processus aussi bien qu'un résultat, c'est-à-dire à la fois une construction du sens et une représentation. Cette notion de schématisation devient centrale dans une approche du « discours situé », dans la mesure où elle tente de rendre compte des modes de fonctionnement des formes observées, de leur emploi et de leur efficacité *in situ* et où elle incite à élaborer des hypothèses sur les raisons qui ont motivé leurs choix.

2.1.3.2. Hétérogénéité du discours et mouvances théoriques

À la lumière de l'ensemble des remarques précédentes, nous considérons que le discours oral spontané appréhendé dans son acception la plus large (CLARK & BLY 1995 ; BERRENDONNER 1997 ; EDWARDS 1997 ; ROULET 1999 ; BERTRAND 1999 ;

AUHLIN 1990, 1999), se montre comme un objet hétérogène dont l'interprétation relève à la fois d'une « sémiotique syncrétique » et d'une « sémiotique à géométrie variable » (DI CRISTO 2000a). Par la première, nous entendons que divers canaux (verbaux et para-verbaux) peuvent coopérer de diverses manières en vue de véhiculer un même contenu conceptuel et que plusieurs significations sont simultanément inférables d'une coopération similaire en un point donné de la chaîne discursive. Par la seconde, nous voulons souligner que des signes verbaux et para-verbaux identiques sont susceptibles de donner lieu à des interprétations différentes dans des contextes différents. Ces considérations nous ramènent en vérité au concept « d'élasticité » du discours, bien connu des sémioticiens (GREIMAS & COURTÉS 1979), qui évoque sa capacité à « mettre à plat », linéairement, des hiérarchies sémiotiques complexes et d'agencer ainsi séquentiellement, et de façon pro-active (BERRENDONNER 1993), des entités discursives (actes, segments, blocs, motifs, épisodes, etc.) qui véhiculent des informations disparates, complémentaires et amalgamées. De ce point de vue, l'élaboration du discours, contrainte par la temporalité, participe de la « cognition séquentielle » qui consiste, comme le stipulent des travaux récents, à gérer la transformation de structures abstraites isomorphes en structures sérielles variables (DOMINEY et al. 1998).

L'hétérogénéité fonctionnelle du discours provient essentiellement du fait qu'il engène des informations qui se rapportent d'une part au signalement des relations logiques que peuvent entretenir des entités discursives (telles que les topiques, les sous-topiques, des propositions, etc.) et, d'autre part, au signalement des diverses activités et actions qui se superposent à cette trame. Bien que ces informations soient intimement mêlées, les recherches qui se consacrent à leur étude participent le plus souvent de deux mouvances distinctes. La première s'attache essentiellement à la représentation de la macro-structure discursive. Il en va ainsi, par exemple, de la RST (MANN & THOMPSON 1988, 2001), qui s'intéresse plus particulièrement à la structure rhétorique du discours, de la DRT (KAMP 1981 ; KAMP & REYLE 1993), qui propose d'interpréter sémantiquement la cohérence textuelle du discours, et de la SDRT (ASHER 1995 ; LASCARIDES & ASHER 1993), qui se présente comme une théorie représentationnelle du discours prenant en compte sa segmentation et son organisation structurelle (BUSQUETS *et al.* 2001). La seconde mouvance, qui s'avère davantage éclatée, s'oriente plutôt vers l'analyse des divers processus de contextualisation discursive (GUMPERZ 1992, 1997 ; AUER & DI LUZIO 1992 ; DURANTI & GOODWIN 1992 ; GROBET 1997) que l'on peut répartir en trois classes selon qu'on les interprète comme étant plutôt : texto-, ego- ou alter-centrés (DI CRISTO 2002). C'est ainsi que les travaux qui s'inscrivent dans cette seconde mouvance portent plus précisément sur l'empaquetage de l'organisation informationnelle (VALLDUVI 1991, 1993 ; LAMBRECHT 1994 ; PROVOST 1996 ; PRINCE 1981 ; CHAFE 1976, 1987, 1993, 1994, etc.), ainsi que sur l'étude des actes énonciatifs (KERBRAT-ORECCHIONI 1980 ; SELTING 1994) et des activités

d'interlocution (KERBRAT-ORECCHIONI 1990, 1992 ; VION 1992 ; SCHEGLOFF 1995, 1996 ; FORD & THOMPSON 1996 ; OCHS *et al.* 1996 ; BERTRAND 1999 ; DOUAY 2000). À notre connaissance, seul le modèle genevois (ROULET *et al.* 2001) semble être actuellement susceptible d'ouvrir la voie à une intégration des deux mouvances dont il vient d'être question (PORTES 2001).

2.1.3.3. Les ressources verbales et para-verbales du discours et l'interprétation des indices qu'elles véhiculent

La conduite du discours, que motive l'hétérogénéité fonctionnelle évoquée précédemment, s'emploie à convoquer et à configurer, au fil de l'enchaînement des représentations conceptuelles, des ressources verbales ou textuelles (choix du lexique et des constructions syntaxiques appropriés), des ressources prosodiques pluri-systémiques (cfr *infra*) et des ressources mimo-gestuelles (MCNEILL 1992, 2000). Bien que la mise en œuvre de ces ressources hétérogènes procède des mêmes représentations conceptuelles, si l'on se rapporte à la production du discours (LEVELT 1989 ; BOCK 1995), et qu'elles participent interactivement à l'activation des inférences, si l'on se réfère à l'interprétation de ce dernier (NAKATANI 1997), il ressort que chacune d'entre elles doit composer avec des contraintes qui lui sont propres et peut bénéficier ainsi d'une relative autonomie. D'autre part, ces ressources ne sont pas obligatoirement convoquées conjointement ou de façon fusionnelle. De ce fait, elles n'entretiennent pas des relations systématiques de redondance et chacune d'entre elles a le loisir d'exprimer sa partie de la « constellation partagée ». Cette dernière remarque nous conduit à formuler l'hypothèse générale que l'interprétation du discours s'effectue sur la base d'une évaluation du poids relatif des indices linguistiques et para-linguistiques qui ont été planifiés par le locuteur, ou sélectionnés à la volée par celui-ci, et qui se trouvent être épisodiquement disponibles, soit de façon conjointe et/ou interactive, soit de manière alternative, dans le flux de parole (BLACHE & DI CRISTO 2002). Cette hypothèse et la remarque qui a conduit à la formuler sont compatibles avec l'idée que les relations qui s'instaurent entre les fonctions discursives conceptuellement motivées et les entités linguistiques et paralinguistiques qui les expriment ne sont pas bijectives, car elles sont du type : une fonction/plusieurs formes et une forme/plusieurs fonctions (DIK 1989, 1997 ; ARIEL 1998). Un exemple caractéristique de cette rupture de correspondances est celui des particules pragmatiques comme : « bon », « après tout », etc. qui peuvent donner lieu à diverses interprétations (FRASER 1999 ; ROSSARI 2001 ; SIMON & GROBET 2002). En revanche, l'hypothèse précédente est peu conciliable avec une conception modulaire stricte des relations de la prosodie au discours, ou avec une approche en termes d'interfaces binômes tels que : prosodie/syntaxe, prosodie/sémantique, etc. Elle est plus favorable à la conception d'une interface plurielle (ou d'un réseau d'interfaces) où les entités en contact

interagissent de façon complexe selon des stratégies qui favorisent tantôt la conspiration, tantôt la compétition. Nous nous référons ici au modèle du « *mapping* » des coalitions des formes de surface et des coalitions de fonctions de BATES & MAC WHINNEY (1989) qui considère que le système cognitif du traitement du langage peut placer sur un pied d'égalité des indices aussi divers que les marques lexicosémantiques, morphologiques, syntaxiques et prosodiques.

2.2. Un cadre interprétatif de référence : la « Grammaire Écologique »

2.2.1. Écologie du discours et flexibilité prosodique

Le cadre interprétatif de référence que nous adoptons pour tenter d'explicitier le rôle de la prosodie dans l'élaboration et l'interprétation du discours est celui d'une « Grammaire Écologique » (DI CRISTO 2000a, 2001), ci-après : GE. Il importe de préciser dès à présent que cette expression est empruntée à LAMBRECHT (1994) et que le terme de grammaire n'est pas à prendre ici dans son acception usuelle et restrictive de « grammaire formelle » (bien que les formes grammaticales syntaxiques et prosodiques constituent un composant central de la GE), mais dans la signification extensive de : cadre descriptif des modes d'existence et de fonctionnement d'une langue naturelle (CRYSTAL 1980). Le qualificatif d'écologique que nous associons à celui de grammaire fonde sa légitimité sur la base axiomatique que le discours est un écosystème, c'est-à-dire un mode d'expression qui s'adapte en permanence au milieu dans lequel il se déploie, en fonction de la versatilité des situations, de l'évolution de la mémoire discursive, de l'état des connaissances mutuellement partagées et non partagées, des pressions exercées par les forces interactionnelles, de la subjectivité et de l'intersubjectivité des interlocuteurs. Le terme d'écologique fait également référence à l'idée de concevoir la prosodie mise en jeu dans le discours comme un dispositif essentiellement adaptatif (GIBBON 1984) qui bénéficie par sa texture même, d'une exceptionnelle flexibilité.

2.2.2. Paradigmes de la Grammaire Écologique

Telle qu'elle est conçue, la GE s'inscrit à la fois dans le paradigme d'une théorie pragmatique large de la communication (cfr la notion de « perspective pragmatique », définie par VERSCHUEREN 1994) et dans le paradigme des sciences cognitives qui stipule que « les capacités intellectuelles et les grandes fonctions adaptatives ou perceptives sont des systèmes de traitement de l'information en provenance de l'environnement » (COSNIER 1998). La « perspective cognitive »

nous invite à envisager le discours du double point de vue de sa production – ou de son élaboration (en termes d'activités de construction, de co-construction et de domaines de planification) – et de sa compréhension/interprétation, fondée à la fois sur la mise en œuvre de processus de décodage et d'inférence. Pour résumer, la perspective globale de la GE s'inscrit dans la paradigme cognitif de l'adaptation contextuelle (ASHBY 1960 ; BATESON 1972) et de la compétition (BATES & MAC WHINNEY 1989). Ce qui revient à dire que la pragmatique n'est pas traitée dans la GE comme un module autonome (une sorte d'analogie conversationnel de la phonologie, de la morpho-syntaxe et de la sémantique), mais comme un système compétitif interactif permettant d'établir des liens entre des expressions (linguistiques et paralinguistiques) et des fonctions discursives.

Du point de vue de son architecture, la GE se présente comme l'interfaçage d'une « grammaire de la contextualisation » et d'une « grammaire des expressions ». La première appréhende le terme de contextualisation dans sa signification la plus large (AUER & DI LUZIO 1992 ; DURANTI & GOODWIN 1992 ; LAMBRECHT 1994 ; DI CRISTO 1999 ; AKMAN 2000). Elle doit s'efforcer de rendre compte (en cherchant à développer des outils de description et de représentation appropriés) à la fois d'une pragmatique de la pertinence et de la subjectivité – et de l'intersubjectivité – dans les usages du langage. Soit d'un ensemble de dispositifs et de stratégies de contextualisation qui sont révélateurs de la compétence discursive et communicative des locuteurs, et par lesquels ces derniers manifestent ostensiblement leurs intentions informatives et communicatives et reconnaissent celles de leurs interlocuteurs, élaborent et interprètent des messages communicativement bien formés (c'est-à-dire, adaptés à leurs objectifs et à leurs contextes), se mettent en scène personnellement dans les actes de discours qu'ils profèrent et les activités discursives qu'ils accomplissent, et participent interactivement à la mise en forme et à la mise en scène des échanges interpersonnels dans lesquels ils sont impliqués.

La grammaire des expressions est formée des primitives et des constructions constitutives des systèmes verbaux (lexique, syntaxe) et para-verbaux (prosodiques et mimo-gestuels) que sélectionnent et configurent les opérateurs fonctionnels de la grammaire de la contextualisation discursive. À l'instar de la grammaire de la contextualisation, la grammaire des expressions est un ensemble hétérogène. La GE est donc une approche qui propose de mettre en correspondance deux ensembles hétérogènes. L'interface de la grammaire des expressions et de la grammaire de la contextualisation se présente en réalité comme un réseau d'interfaces formes / fonctions qui, dans l'optique de l'analyse prosodique qui nous intéresse, transcende les interfaces traditionnelles : prosodie / syntaxe, prosodie / sémantique, prosodie / pragmatique, et dont l'étude s'inscrit alors dans le paradigme d'une « Théorie de l'Optimalité Généralisée ». Dans cette perspective, l'interprétation du réseau d'interfaces vise notamment à rendre compte de la façon dont les contraintes de la contextualisation et les contraintes des systèmes linguistiques (syntaxique et proso-

dique) cohabitent ou entrent en conflit dans le choix des configurations prosodiques qui participent à l'élaboration et à l'interprétation du discours. Enfin, ce réseau d'interfaces peut être envisagé comme un locus où s'articulent une linguistique des états et une linguistique des opérations et où se coordonnent deux aspects essentiels des activités cognitives mises en jeu dans le langage (le troisième étant évidemment l'aspect conceptuel) : l'aspect représentationnel et l'aspect procédural, au service de la dynamique communicative du discours (BLAKEMORE 1987 ; WILSON & SPERBER 1993 ; NICOLLE 1997).

2.2.3. L'articulation formes/fonctions et les approches concernées

Il apparaîtra sans équivoque que la GE ambitionne d'établir des passerelles entre deux mouvances souvent considérées comme antagoniques : le formalisme et le fonctionnalisme (NUYTS 1994 ; NEWMAYER 1998). En se positionnant de la sorte, la GE invite à dialoguer des théories et des approches qui émanent de champs disciplinaires divers, mais qui présentent cependant une authentique parenté conceptuelle. C'est ainsi que, outre les récentes approches de la phonologie prosodique (LADD 1996), qui s'avèrent incontournables pour édifier une partie des représentations constitutives de la *grammaire des expressions*, la GE convoque également, à titre d'exemple, pour informer la *grammaire de la contextualisation* : les travaux de GUMPERZ (1992), les théories de VALLDUVI (1991, 1993) et de LAMBRECHT (1994) sur la mise en forme informationnelle, la *théorie de la pertinence* de SPERBER & WILSON (1986), la *théorie des actes de sens* de BRUNER (1990), la *grammaire fonctionnelle* de DIK (1989), le *modèle de compétition* de BATES & MAC WHINNEY (1989), le modèle psycholinguistique de production de LEVELT (1989) et la *théorie présuppositionnelle Hyper/Hypo* de la variabilité phonétique de LINDBLOM (1990), l'approche de l'intégration temporelle de la hiérarchie des actions communicatives de MCNEILL et al. (2001) ainsi que les diverses approches cognitives et pragmatiques de l'interlocution (VION 1992 ; COSNIER 1998 ; DOUAY 2000).

2.3. Vers l'élaboration d'une grille multilinéaire de représentation des indices « textuo-discursifs »

2.3.1. Remarques préliminaires

Nous avons souligné dans les sections précédentes l'hétérogénéité fonctionnelle du discours. Nous avons également proposé un cadre interprétatif propice à la confrontation de cette hétérogénéité (dont peut rendre compte une interprétation plurielle de la contextualisation) avec la diversité des ressources expressives

(verbales et paraverbales) qui sont mises en œuvre pour la manifester. Nous avons enfin mis en avant l'idée que les relations entre les fonctions discursives et ces ressources ne sont pas bijectives et que l'interprétation de leurs correspondances implique l'examen d'un réseau d'interfaces. Concrètement, dans la mise en œuvre de notre approche, l'accès à ce réseau passe par l'élaboration d'une grille de représentations multilinéaire dont une partie est réservée au codage des faits prosodiques et l'autre à celui des faits de discours. Il importe de préciser que les propositions que nous énumérons pour édifier la grille discursive ne sont que provisoires dans la mesure où elles participent d'une heuristique et non d'un modèle de discours préétabli. D'autre part, on remarquera que les grilles que nous décrivons n'intègrent pas encore le composant mimo-gestuel qui constitue cependant un élément très important de l'interprétation du discours (on se reportera à ce propos aux actes des deux derniers colloques sur l'oralité et la gestualité : SANTI *et al.* 1998 ; CAVE *et al.* 2001). Une telle intégration fait néanmoins partie des projets que nous entendons mettre à exécution ultérieurement dans le déploiement des activités de notre groupe. Dans l'immédiat, pour se faire une idée de la façon dont il est possible d'associer des représentations prosodiques et des représentations mimo-gestuelles pour capturer des aspects dynamiques de la conduite du discours, on pourra se reporter aux travaux novateurs de MCNEILL *et al.* (2001).

2.3.2. Principes de construction de la grille « textuo-discursive »

La partie non prosodique de la grille correspondant à la notation des phénomènes « textuels » et/ou discursifs s'attache à inventorier les indices de différents niveaux qui sont présents dans le texte et qui permettent d'effectuer des inférences sur le discours même en tant qu'activité ayant donné naissance à ce texte. Les niveaux répertoriés sur le plan de l'analyse du texte considéré comme un produit discursif concernent le lexique, la syntaxe, les phénomènes énonciatifs, les phénomènes référentiels et leur combinaison. Ces différents niveaux, qui participent à la construction de l'interprétation, sont complétés par des informations que peuvent fournir différents modèles de discours, tels que la RST (MANN & THOMPSON 1988, 2001) ou le modèle genevois (ROULET *et al.* 2001).

Une (ou plusieurs) ligne(s) de cette partie « textuo-discursive » de la grille est (sont) consacrée(s) à consigner des informations issues de l'interprétation des concordances et des discordances d'indices répertoriés sur les lignes supérieures. Cette interprétation peut concerner les aspects argumentatifs du discours ou ses aspects structurels et organisationnels. Par mesure de commodité, il est loisible de présenter les différents aspects de ces niveaux « textuo-discursifs » comme relevant de deux ordres de phénomènes distincts : ceux qui interviennent dans la construction du sens d'un énoncé particulier et ceux qui tissent des relations entre les informations produites par les différents énoncés et qui participent ainsi à la fois à

l'établissement d'une cohérence discursive et à la construction des représentations cognitives qui en résulte.

Étant donné la diversité des échantillons de discours qui peuvent faire l'objet de ces analyses, il apparaît que toutes les lignes potentielles de la grille « textuo-discursive » n'ont pas besoin d'être systématiquement convoquées pour l'étude d'un exemple donné, sans que cela ne remette en cause l'architecture de la représentation.

2.3.3. Notation des phénomènes internes à l'énoncé

Une première partie de la grille « textuo-discursive » concerne les « micro-informations » que l'on peut tirer d'unités de rang inférieur à celui de l'énoncé, défini comme une construction syntaxique maximale, parfois supposée – dans certaines approches – constituer le résultat d'un seul acte d'énonciation. Ces « micro-informations » vont être d'ordre syntaxique, lexical, sémantique et énonciatif.

Les lignes réservées à la syntaxe font apparaître les dépendances et relations d'ordre syntaxique entre les mots du texte. À chaque prédicat verbal est associé un numéro qui est aussi affecté à chaque argument du même prédicat verbal. Deux types d'attributs sont marqués : des attributs d'ordre catégoriel et des attributs d'ordre relationnel. Parmi les attributs relationnels figurent les fonctions syntaxiques et à ce titre les fonctions sujet et objet sont indiquées sur la grille. Les attributs d'ordre catégoriel indiquent la catégorie syntaxique associée au mot ou au groupe de mots si celle-ci s'applique à une parenthèse. La catégorie la plus élevée est celle de la phrase notée P. Ainsi Pron Obj33 signifie Pronom qui a la fonction syntaxique d'Objet du verbe 33. En ce qui concerne la représentation sur la grille, une ligne propose le texte retranscrit parenthésé qui met en évidence les dépendances d'ordre syntaxique entre les éléments du corpus. Une autre ligne utilise le parenthésage précédent sans le texte, dans lequel on insère les catégories syntaxiques pour chaque élément lexical ou groupe d'éléments entre parenthèses.

Le contenu des lignes attribuées aux phénomènes lexicaux peut être variable d'une grille à l'autre, car tous les phénomènes lexicaux n'ont pas une égale pertinence dans la construction de l'interprétation du discours, selon les exemples retenus. Comme pour les lignes qui relèvent des autres niveaux, nous avons donc affaire ici à des observables sélectionnés dans la masse des données sur la base de critères qu'il importera de spécifier. Les lignes affectées aux phénomènes lexicaux vont pouvoir concerner : des traits sémantiques attachés aux items lexicaux (animé/inanimé, comptable/non comptable, etc.), des catégorisations sémantiques de sous-classes de verbes (verbes de parole, verbes porteurs de présupposés aspectuels : « *cesser de* », « *continuer à* » ou factifs : « *savoir que* »), ainsi qu'une interprétation de ces phénomènes : présence d'un discours rapporté pour les verbes de parole, présupposés véhiculés par les verbes présupposants.

La sémantique actantielle (**sém act**) représente une interprétation des prédicats et de leurs arguments en fonction de leur emploi en énonciation. En plus clair, un verbe (prédicat) est interprété comme un procès et ses arguments sont interprétés comme des actants jouant un certain rôle dans ce procès. Le type du procès (processus, état, etc.) et les rôles actantiels (agent, expérienceur, patient, bénéficiaire, etc.) résultent de l'interaction de la syntaxe et du lexique pour lequel entrent en ligne de compte une catégorisation du verbe et les traits sémantiques qui peuvent être attachés à ses arguments.

Les phénomènes énonciatifs (**énonciat**^o) qui semblent particulièrement intéressants à noter sont divers et peuvent occuper plusieurs lignes concernant en particulier les déictiques, par le biais des repérages personnels, spatiaux et temporels relatifs à la situation d'énonciation, laquelle est définie comme un repère constitué des personnes du discours (locuteur et allocutaire), du lieu de l'énonciation et du moment de l'énonciation **M**^o. Ces repérages permettent notamment de distinguer, en se fondant sur les temps verbaux et les caractéristiques aspectuelles des verbes, les procès accomplis au moment de l'énonciation de ceux qui sont encore en cours au même moment (cfr grille 1, Figure 3 dans la section 6 et sur le CD-ROM).

Une autre ligne peut être consacrée aux gloses et autres formules méta-énonciatives (AUTHIER-REVUZ 1995 ; GÜLICH & KOTSCHI 1995).

Enfin, le niveau énonciatif concerne également la pluralité des voix dans le discours d'un locuteur, telle qu'elle s'exprime notamment dans le discours rapporté (BERTRAND 1999). Selon DUCROT (1984), « le sens d'un énoncé correspond à la description de son énonciation [...] consistant en diverses informations de type argumentatif, illocutoire et des indications dites plus primitives présupposées par tout ce que l'on peut dire sur les aspects illocutoires, argumentatifs et expressifs du langage. Il s'agit des renseignements que l'énoncé apporte dans son sens même, sur le ou les auteurs éventuels de l'énonciation ». On conçoit dès lors l'importance de la détermination de celui qui est à l'origine de la parole. Nous nous inspirons de la conception du sujet chez Ducrot en la simplifiant : nous distinguons en effet le locuteur, responsable de l'énonciation, des énonciateurs qui renvoient aux points de vue exprimés dans le discours. Les deux niveaux de discours, dialogal et dialogique, sont ainsi pris en compte. Plus généralement, le niveau énonciatif renseigne donc sur le degré d'implication du locuteur dans son discours, qui contribue à l'élaboration du sens : dans sa façon de mettre en scène son discours, le locuteur rend compte en effet du sens à donner aux énoncés.

En ce qui concerne les relations entre les différents niveaux de la grille, des auteurs ont déjà envisagé l'existence de certains liens tels que par exemple : comment le recours à d'autres voix permet de construire un point de vue et de développer une argumentation, ou bien encore, comment la pluralité des voix peut influencer sur le niveau subjectif des discours (BERTRAND 1999, à paraître ; VION 2001), c'est-à-dire sur les images que les locuteurs sont amenés à construire d'eux-

mêmes et de leurs interlocuteurs dans l'échange, images qui ont une incidence sur l'élaboration des discours.

D'autres lignes, qui n'apparaissent pas systématiquement dans toutes les grilles, peuvent concerner le niveau des affects, étant entendu que la prise en considération de ce dernier s'avère incontournable pour l'analyse des conversations (SELTING 1994).

2.3.4. Notation des phénomènes inter-énoncés

Les phénomènes ainsi qualifiés concernent l'établissement de la cohérence du discours par des outils de divers types. L'interprétation du discours comme un tout cohérent va s'établir en fonction d'éléments référentiels, argumentatifs et organisationnels.

Les phénomènes référentiels notés dans les grilles se fondent sur l'identification des expressions référentielles (première ligne) pour déterminer d'abord les objets de discours qu'elles construisent et les propriétés de ces objets (deuxième ligne), et ensuite leurs liens anaphoriques dans les représentations cognitives créées par le discours (ligne « **anaphores** »).

Les objets de discours (ou « référents ») désignés par les expressions référentielles sont notés : **O1**, **O2**, **O3**, etc. Il peut être utile de préciser leurs caractéristiques en termes de *type* d'objet d'abord (individu, classe, collectif) et en termes de *statut* de l'objet ensuite (**ext** = extensionnel = spécifique vs **int** = intensionnel = générique), lequel résulte d'une combinaison des déterminants employés et des prédications. Enfin, les attributs de ces objets sont issus des différentes prédications qui portent sur l'expression référentielle et, lorsqu'on a affaire à des expressions déictiques, des phénomènes énonciatifs.

L'intérêt majeur de noter ces phénomènes est de suivre l'évolution des objets de discours tout au long de la progression du texte, de déterminer comment leurs attributs se modifient, voire se contredisent (cfr la grille 2, au point 6, en ce qui concerne **O1**, qui renvoie à la locutrice BG aussi bien en tant que personnage de son discours rapporté qu'en tant que narratrice de ce même discours rapporté).

Les liens anaphoriques entre objets de discours sont établis en fonction de la référence des pronoms d'une part, et de la combinaison du type de déterminant et du signifié des substantifs dans les SN, d'autre part. Ces liens peuvent être coréférentiels (**coref**) ou associatifs (**assoc**). Ils participent à l'élaboration de la structure ou de l'organisation du discours, tout en construisant l'interprétation. Dans le cas où le lien anaphorique est coréférentiel, l'objet de discours **O** désigné par deux expressions référentielles successives est indexé de la même façon (par exemple : **O1** et **O1'** désignent un même référent, mais qui a évolué entre ses deux dénominations et qui a acquis de ce fait des attributs supplémentaires).

Les lignes concernant les phénomènes argumentatifs (**argumentat**^o) sont inspirées de la conception de l'argumentation de DUCROT, selon laquelle un énoncé donné tend vers une certaine conclusion implicite, et où les enchaînements entre énoncés sont décrits comme des liens entre un ou plusieurs arguments et une conclusion (la grille 2 sur CD-Rom en annexe) présente un extrait dans lequel les énoncés tendent vers des conclusions contradictoires).

En ce qui concerne l'organisation des unités discursives, plusieurs lignes peuvent être attribuées à la notation de la structure hiérarchique et fonctionnelle de l'extrait étudié, en appliquant par exemple le modèle modulaire de l'organisation du discours de ROULET (ROULET *et al.* 1985, 2001). Dans ce cas, la première ligne consigne la structuration hiérarchique du passage en actes, interventions et échanges et la deuxième ligne indique la structuration fonctionnelle, c'est-à-dire les liens entre ces unités discursives.

D'autres lignes servent à noter la structure rhétorique de l'échantillon considéré. Cette dernière est établie par référence au modèle de l'organisation textuelle de MANN & THOMPSON (1988, 2001). Dans cette optique, une première ligne sert à préciser la segmentation en unités qui vont constituer les segments textuels à mettre en relation. La ligne suivante explicite la hiérarchie de ces relations, leur orientation d'un segment vers un autre et les étiquettes sémantiques qui peuvent leur être attachées, selon les catégories définies par les auteurs précédemment cités. L'objectif de ces deux lignes de notation est double :

- Il s'agit, d'une part, de voir si les grilles permettent d'inventorier des indices d'ordre divers participant à l'élaboration de ces relations discursives (dites « rhétoriques ») et de mettre au jour des combinaisons entre ces indices. Les catégories proposées par MANN & THOMPSON restent, en l'état actuel du modèle, largement intuitives, et l'on ne dispose pas encore de critères de décision ni d'heuristique permettant de joindre deux segments selon telle ou telle relation. Il s'agit donc d'examiner dans quelle mesure ces relations sont « calculables » pour évaluer la pertinence du modèle concerné dans la description de l'organisation discursive.
- D'autre part, la structuration obtenue pour un passage étudié peut être mise en relation avec les résultats issus de l'application d'autres modèles discursifs, comme le modèle genevois précédemment cité, afin de dégager de cette comparaison des points de similitude et/ou de dissension.

La liste qui vient d'être présentée n'est pas exhaustive et d'autres lignes pourront compléter cette grille plurilinéaire de notation des phénomènes textuo-discursifs.

3. ANALYSE PROSODIQUE : CONCEPTION, DEMARCHE METHODOLOGIQUE ET CODAGE

3.1. Conception de la prosodie

3.1.1. Remarques liminaires

L'étude des relations de la prosodie au discours, telle que nous les envisageons, c'est-à-dire dans la double perspective de l'élaboration et de l'interprétation d'une construction hétérogène et pluriforme, consiste en premier lieu à façonner un outil de travail se présentant sous l'aspect d'une grille d'analyse multilinéaire scindée en deux volets. Nous avons présenté dans les sections précédentes les critères retenus pour édifier la partie de la grille dédiée aux notations « textuo-discursives ». Nous proposons d'exposer dans les sections suivantes les principes théoriques et méthodologiques qui président aux codages des systèmes prosodiques dont les représentations forment l'armature du second volet de la grille d'analyse.

3.1.2. De la multilinéarité des systèmes prosodiques

3.1.2.1. Multilinéarité et ordres structurels

Alors que la prosodie est souvent assimilée dans de nombreux travaux à l'intonation, nous adoptons en ce qui nous concerne une conception structurelle tripartite de la prosodie. Conformément à cette conception, les primitives et les constructions des systèmes prosodiques se distribuent selon trois axes où se déploient, respectivement, l'organisation métrique, l'organisation tonale et l'organisation temporelle du discours (DI CRISTO 2000b). Cette factorisation de la prosodie justifie par elle-même le recours à des représentations plurilinéaires.

Dans la perspective triadique que nous venons d'évoquer, nous estimons que l'organisation métrique bénéficie d'une certaine préséance au regard des deux autres ordres structurels, ce qui signifie que l'assignation des patrons tonal et temporel d'un énoncé donné semble présupposer la construction de son squelette métrique (LEVELT 1989 ; WHEELDON & LAHIRI 1997). Toutefois, la préséance du métrique n'implique nullement l'isomorphisme structurel et chaque ordre dispose d'une autonomie relative. Comme on pourra le constater par la suite, les principes d'organisation tripartite des systèmes prosodiques, de préséance métrique et d'autonomie relative, ont des retombées manifestes sur les options retenues pour le codage de la prosodie,

l'évaluation du poids relatif des indices prosodiques et la chronologie de la méthode expérimentale mise en œuvre pour accéder à ce codage.

3.1.2.2. *Multilinéarité et orthogonalité*

L'approche plurilinéaire de la prosodie que nous préconisons trouve également sa justification dans l'affirmation que chaque ordre prosodique structurel participe de deux dimensions – ou de deux vecteurs – que nous qualifions respectivement de : syntagmatique et orthogonal(e). Si l'on se réfère par exemple à l'organisation tonale, le vecteur syntagmatique concerne à la fois la spécification de balises tonales (comme les accents et les cadences mélodiques) et celle de motifs mélodiques adjacents (MERTENS *et al.* 2002). Au regard de cette même organisation, le vecteur orthogonal se rapporte à des phénomènes tels que les changements de registre, de gamme tonale et les effets dits « d'abaissement », comme l'effet de déclinaison, par exemple (CONNELL & LADD 1990).

Si l'on se réfère cette fois à l'organisation temporelle, le vecteur syntagmatique intéresse la distribution des pauses, des allongements syllabiques et la spécification de certains motifs temporels (durée syllabique croissante/décroissante dans un domaine donné), alors que le vecteur orthogonal relève plus précisément des variations de tempo qui se superposent à ces phénomènes. En définitive, la prise en considération du caractère bi-dimensionnel des trois ordres structurels prosodiques nous oriente vers une interprétation individuelle et corrélative de ces derniers, en termes de propriétés locales, globales et itératives, cette interprétation étant favorisée, une fois encore, par une représentation plurilinéaire des systèmes prosodiques.

Les difficultés majeures consécutives à l'intégration des vecteurs orthogonaux résident dans les décisions qui doivent être prises à propos de leur codage : ces vecteurs doivent-ils faire l'objet d'un codage catégoriel ou d'une notation continue ? Cette question, appliquée à la problématique de l'interprétation des phénomènes de registre notamment, a suscité des débats féconds qu'il n'est pas possible, pour des raisons d'économie, de rappeler ici (LADD 1990, 1994 ; CLEMENTS 1990 ; PATTERSON 2000). En ce qui nous concerne, nous avons délibérément opté pour un codage catégoriel des phénomènes relevant des vecteurs orthogonaux. Il est clair dans nos esprits, cependant, que cette option mérite d'être enrichie par une transcription plus fine de ces phénomènes, en faisant appel à des symboles diacritiques (analogues à ceux qui sont proposés pour le segmental par l'Alphabet Phonétique International) ou à un système de notation permettant d'introduire certaines formes de gradation. Il est probable que ces options seront introduites dans des versions ultérieures de la grille d'analyse.

3.1.2.3. *Multilinéarité et niveaux de représentation*

Les conclusions du paragraphe précédent soulèvent de façon indirecte le problème des niveaux de représentation de la prosodie. Doit-on se limiter à deux niveaux : le niveau des représentations analytiques (ou physiques) et le niveau des représentations phonologiques abstraites, comme c'est le cas dans la majorité des approches ? Ou bien peut-on postuler l'existence de niveaux intermédiaires entre ces deux niveaux extrêmes ? Est-on en droit de stipuler l'existence d'un niveau de représentation phonétique, différent de celui des représentations analytiques ? Ces questions ne sont pas oiseuses si l'on admet avec GUSSENHOVEN (2002) que les significations véhiculées par l'intonation, par exemple, ressortissent à deux composants du langage : la grammaire intonative et l'implémentation phonétique de l'intonation. Toujours selon cet auteur, l'implémentation phonétique de l'intonation contribue à exprimer des significations pour la plupart universelles (mais loin d'être négligeables, selon nous, pour ce qui concerne l'interprétation du discours), motivées par l'incidence de trois codes biologiques : le code fréquentiel, le code d'effort et le code de production. Le premier est exploité, par exemple, pour exprimer des relations de puissance (ou de dominance), le second pour marquer la force d'implication du locuteur et l'emphase, et le troisième pour signaler le début (mélodie haute) et la fin (mélodie basse) de certaines entités discursives. Il apparaît également que les figures imposées par ces codes peuvent faire l'objet de diverses formes de grammaticalisation selon les langues.

Sans nous aventurer plus loin dans ce débat, au demeurant captivant, nous proposons de mettre en œuvre un modèle d'analyse qui apporte précisément des éléments de réponse à la problématique de l'articulation phonétique/phonologie dans la représentation de la prosodie, plus précisément de l'organisation tonale. Il s'agit de l'approche INTSINT (HIRST & DI CRISTO 1998) que nous décrivons en détail dans la partie de notre exposé dévolue à la présentation des démarches méthodologiques. Nous tenterons de montrer que cette approche met en pratique les propositions de PIERREHUMBERT & BECKMAN suivant lesquelles

la division entre la phonologie et la phonétique est une question empirique qui ne peut être tranchée que par la construction de modèles complets qui explicitent leur rôle respectif dans la description de la structure sonore. (PIERREHUMBERT & BECKMAN 1988 : 5, citation traduite en français par nous)

Il apparaîtra clairement dans la suite de cet article que l'approche INTSINT procède également d'une conception plurilinéaire des niveaux de représentation.

3.1.2.4. Multilinéarité et articulation formes/fonctions

Les représentations cognitives qui président à l'élaboration et à l'interprétation du langage mettent en correspondance des formes et des fonctions. De ce point de vue, il est tentant de chercher à établir un tableau de correspondances entre les systèmes prosodiques et un panel de fonctions linguistiques et paralinguistiques, dans la mesure où il est convenu que la prosodie assiste la production et la planification de la parole, organise le matériau verbal, empaquette l'information, active les inférences, exprime l'affect, régule l'interlocution et contribue à l'identification des caractéristiques idiolectales, sociolectales et dialectales des individus, ainsi qu'à la reconnaissance des styles de discours. Pour stimulante qu'elle soit, une telle entreprise reviendrait à « mettre la charrue avant les bœufs » par rapport à notre démarche qui vise précisément à mettre en évidence la pluralité des fonctions prosodiques exhibées par le discours. Sans préjuger de cette pluralité, nous considérons cependant que la prosodie a la capacité d'assumer certaines fonctions de base que nous proposons d'inventorier par les traits suivants : [± PROMINENCE], [± NUCLEUS], [± EMPHASIS], [± BOUNDARY], [± TERMINAL], en nous référant à l'étude de HIRST (1977).

Dans notre approche, la spécification de ces fonctions fait l'objet d'un codage indépendant dans la partie de la grille d'analyse qui est réservée à la prosodie. Nous nous démarquons sur ce point d'autres systèmes de transcription qui, à l'instar du modèle ToBI (SILVERMAN *et al.* 1992), proposent de fusionner les notations formelle et fonctionnelle de la prosodie. C'est ainsi, par exemple, que dans les notations [H*] et [L H%] de ToBI, [H] et [LH] se rapportent à des formes et [*] et [%] à des fonctions, soit, respectivement, la fonction accentuelle et la fonction de frontière. Une telle façon de procéder ne convient pas à notre approche, dans la mesure où il n'est pas prouvé qu'une fonction comme celle de frontière par exemple, soit uniquement imputable à des formes prosodiques, et où la confrontation des indices prosodiques et des indices « textuo-discursifs » (voir *supra*) que permet la grille d'analyse plurilinéaire doit précisément nous conduire à évaluer la part respective de ces derniers dans l'expression de cette fonction. Pour conclure cette section, nous soulignerons que la séparation des formes et des fonctions prosodiques privilégie aussi un mode de représentation plurilinéaire.

3.2. De la matérialité des faits prosodiques : nature pluriparamétrique de la prosodie

Il est clairement établi que la substance de la prosodie est pluriparamétrique. Il apparaît cependant que de nombreux travaux se limitent à l'investigation d'un seul paramètre, le plus souvent la Fréquence fondamentale (F0) que l'on assimile parfois

à l'intonation. Nombre de ces travaux admettent aussi – implicitement ou explicitement – que l'intonation est synonyme de prosodie. Il est manifeste que ce point de vue n'est pas compatible avec notre conception pluri-systémique de la prosodie (cfr 3.1.2.1). Par ailleurs, nous réfutons l'idée qu'il est loisible d'établir des liens directs entre les paramètres prosodiques physiques et des fonctions, car nous considérons avec d'autres auteurs (LADD 1996) que le niveau intermédiaire de la représentation phonologique est indispensable. De même, nous estimons utile de rappeler que l'intonation ne se résout pas à la mélodie, ni cette dernière à la F0 ('T HART & COLLIER 1975). Il est nécessaire que nous donnions des éclaircissements sur le caractère par trop péremptoire de ces affirmations, afin de préciser les fondements de notre démarche méthodologique, exposée dans la section suivante.

La conception pluri-systémique de la prosodie a pour corrélat qu'une fonction donnée, telle que la fonction de frontière par exemple, peut être assumée par la mise en œuvre conjointe, ou complémentaire, ou alternative des trois ordres structurels définis précédemment (dans le cadre d'une coopération complexe et non systématique avec les autres systèmes de la langue comme le lexique, la syntaxe et la sémantique). L'implémentation des ordres structurels se fonde sur la mise en œuvre d'un ensemble de commandes articulatoires (organisation temporelle) et / ou vocales (organisation tonale) dont la finalité est d'émettre des signaux véhiculant certaines informations. Ces informations sont interprétées par l'auditeur selon un processus de transmutation qui consiste à construire des percepts pour reconnaître des formes. L'une des principales difficultés de l'analyse prosodique provient alors des ruptures de correspondance qui se produisent entre les divers maillons de cette chaîne.

Un premier type de rupture de correspondance se produit entre les intentions communicatives du locuteur et les signaux prosodiques dans lesquels elles se projettent. Cette rupture provient du fait que les paramètres prosodiques qui constituent ces signaux reflètent l'incidence de contraintes de bas niveau inhérentes à la production de la parole. Ces contraintes engendrent une variabilité « micro-prosodique » qu'il importe de dissocier de la variabilité « macroprosodique » que l'on suppose être directement liée à la projection des intentions communicatives (DI CRISTO 1978 ; DI CRISTO & HIRST 1986). Nous exposerons dans la section suivante la méthode retenue pour réaliser cette factorisation.

Un second type de rupture de correspondance concerne les relations qui s'instaurent entre les signaux prosodiques et leur perception. Deux raisons majeures sont à l'origine de cette rupture. D'une part, les variations physiques associées aux divers paramètres prosodiques ne sont pas perçues en l'état (ROSSI *et al.* 1981). D'autre part, la construction d'un percept prosodique (par exemple, le percept de proéminence) peut être activé par l'action d'un ou de plusieurs paramètres. Dans le cas d'une activation fusionnelle, il convient de préciser qu'elle peut être le fruit

d'une interaction complexe entre les paramètres mis en jeu. C'est ainsi, par exemple, qu'une variation intra-syllabique de F0 a un effet sur la durée perçue de cette syllabe, ou que la durée affecte la perception de la sonie (de l'intensité subjective du segment concerné), ou encore que l'intensité a des incidences sur la perception des variations de la fréquence fondamentale ou de la durée (cfr DI CRISTO 1978 ; ROSSI *et al.* 1981 ; NISHINUMA 1996, pour une synthèse de la question).

Étant donné les ruptures de correspondance dont il vient d'être question, il s'avère nécessaire de confronter les données issues de l'analyse auditive (relatives à la notation des aspects prosodiques formels, comme les allongements ou l'intonation, ou à celle des traits fonctionnels de proéminence, de frontière, etc.) avec une représentation analytique plurilinéaire du jeu des paramètres prosodiques. Cette confrontation est destinée à élucider la nature des interactions qu'entretiennent les paramètres prosodiques et à utiliser cette information pour forger un outil de codage prosodique capable d'apporter une contribution efficace à l'interprétation des relations de la prosodie au discours, en permettant de « visualiser » des correspondances et des discordances entre les différents niveaux de représentation de la grille prosodique et de la grille « textuo-discursive ». La démarche qui préside à l'élaboration de cet outil fait l'objet des développements de la section suivante.

3.3. Vers l'élaboration d'une grille multilinéaire de représentation de la prosodie.

La démarche qui conduit à édifier la partie de la grille consacrée à la prosodie vise à établir un codage plurilinéaire de cette dernière en termes de symboles catégoriels. Cette démarche est pluripartite, dans la mesure où elle procède de la mise en œuvre de trois types d'approches : l'approche auditive classique centrée sur l'écoute et la notation, le codage automatique ou semi-automatique des structures prosodiques de surface et la normalisation des paramètres prosodiques (approche statistique).

3.3.1. L'approche auditive

L'approche auditive retenue pour l'étude du corpus Fayard-Groult s'apparente à celle que préconisent des modèles influents comme ToBI (SILVERMAN *et al.* 1992) ou IViE (GRABE & POST 2002). Elle consiste à noter en regard du texte transcrit orthographiquement (ou phonétiquement) un ensemble de phénomènes perçus. C'est ainsi que neuf experts² appartenant à notre groupe de travail ont procédé par étapes au repérage et à la notation : des proéminences (2 catégories : emphatique, non emphatique), des frontières (2 catégories : terminale, non terminale), des durées

² Nous tenons à remercier les membres du groupe PRODIGE, A. COLAS et M. VION, qui ont prêté leur concours à l'analyse auditive du corpus qui a servi de base à notre étude.

syllabiques (3 catégories : réduite, longue et très longue), des variations du tempo (2 catégories : accéléré et ralenti), des pauses (2 catégories : longue et très longue), des prises de souffle et des phénomènes de laryngalisation.

Les renseignements issus de cette expérience et du traitement des données qui l'accompagne fournissent un ensemble d'informations hétérogènes relatives à des aspects formels et fonctionnels de l'organisation prosodique qui pourront être réutilisées pour informer le codage des grilles prosodiques.

3.3.2. Codage automatique ou semi-automatique des structures prosodiques de surface : l'approche INTSINT

Au stade actuel de l'avancement de nos recherches, le codage automatique (ou semi-automatique) des structures prosodiques de surface ne concerne que l'organisation tonale. Des travaux en cours que nous évoquerons brièvement dans les sections suivantes s'appliquent également à étendre cette procédure au codage de l'organisation temporelle.

L'approche INTSINT appliquée au codage automatique de l'organisation tonale de surface, se déroule en deux temps. Dans un premier temps, la courbe de F0 est modélisée au moyen de l'algorithme MOMEL (HIRST *et al.* 2000) sous la forme d'une courbe lisse et continue formée d'une séquence de points-cibles. Ce type de représentation, qui concrétise la neutralisation de certains effets microprosodiques (cfr 3.2) et qui illustre un premier niveau d'abstraction par rapport à la représentation analytique de la courbe de F0 brute, est interprété comme une représentation phonétique de l'intonation. Dans l'esprit de cette approche, les représentations phonétiques ne sont donc pas des analogues des courbes de F0 dont elles sont cependant dérivées par une procédure automatique.

Dans un second temps, les points-cibles de la courbe modélisée font l'objet d'un codage symbolique automatique fondé sur l'utilisation d'un « alphabet intonétique », dénommé INTSINT (« *an International Transcription System for INtonation* » : HIRST & DI CRISTO 1998). Ce codage est soumis ensuite à une vérification qui peut donner lieu à des corrections manuelles, ce qui justifie que nous parlions également d'approche semi-automatique lorsque nous y faisons référence. L'alphabet INTSINT se limite à huit symboles. Trois de ces derniers sont destinés à la notation des valeurs absolues qui définissent l'espace tonal des locuteurs : M(id), T(op) et B(ottom). Cinq autres symboles servent à noter les valeurs relatives de la hauteur à l'intérieur de l'espace tonal : H(igher), L(ower), S(ame), U(pstepped), D(ownstepped).

Les symboles M, T et B permettent ainsi de coder de façon discrète les hauteurs de référence que représentent le registre moyen d'un locuteur donné, son registre maximal et son registre minimal, respectivement. Tous les autres symboles encodent des niveaux de hauteur relatifs (CRYSTAL 1971). C'est ainsi que le symbole H est

retenu pour coder un point-cible dont la valeur est supérieure à celle des points-cibles adjacents. En revanche, le symbole L sera sélectionné pour coder un point-cible dont la valeur est inférieure à celle des points-cibles adjacents. Le symbole U sera choisi pour coder un point-cible plus haut que le précédent et plus bas que le suivant, et le symbole D, pour coder un point-cible plus bas que le précédent et plus haut que le suivant. Enfin, le symbole S permettra de coder un point-cible de même niveau que le précédent.

Dans cette approche, le codage des cibles de la représentation phonétique au moyen de l'alphabet INTSINT est interprété comme une représentation phonologique de surface de l'organisation tonale. Nous la caractérisons de la sorte pour désigner une représentation formée de séquences d'entités discrètes (à la différence de la représentation phonétique qui exhibe une configuration continue). La représentation phonologique de surface offre l'avantage de ne pas préjuger du système phonologique (de l'inventaire des patrons intonatifs distinctifs) de la langue faisant l'objet de cette représentation (à la différence du système ToBI, par exemple). Elle se présente plutôt comme une étape intermédiaire de la démarche qui vise à construire ces patrons distinctifs à partir des observables. Ce faisant, elle permet la mise en évidence de régularités tonales, locales et globales, qui peuvent être interprétées comme des invariants de la représentation phonologique ou comme des variantes allophoniques. Ces dernières doivent également être prises en compte dans l'analyse du discours, dans la mesure où elles ont la faculté de véhiculer de multiples informations.

L'environnement informatique de l'approche INTSINT intègre un module de resynthèse qui permet de juger auditivement de l'effet produit par des modifications manuelles apportées à la disposition des points-cibles de la représentation phonétique (modifications de l'alignement et de la hauteur). Ce dispositif s'impose comme un auxiliaire précieux pour l'étude des relations de la prosodie au discours, en ce sens qu'il apporte une contribution non négligeable à l'évaluation du poids relatif des indices prosodiques et des indices « textuo-discursifs » dans l'émergence des fonctions discursives.

La mise en œuvre de l'approche INTSINT débouche sur le codage de la composante syntagmatique de l'organisation tonale relative à l'échantillon de corpus analysé (Figure 1). Elle ne permet pas pour l'instant de procéder au codage automatique de sa composante orthogonale (cfr 3.1.2.2). Des recherches en cours visent à combler cette lacune. En attendant, les indices orthogonaux font l'objet d'un codage manuel fondé sur l'analyse auditive et l'observation des courbes modélisées. Ce codage est reporté sur les grilles d'analyse que nous présentons dans la dernière partie de cette étude.

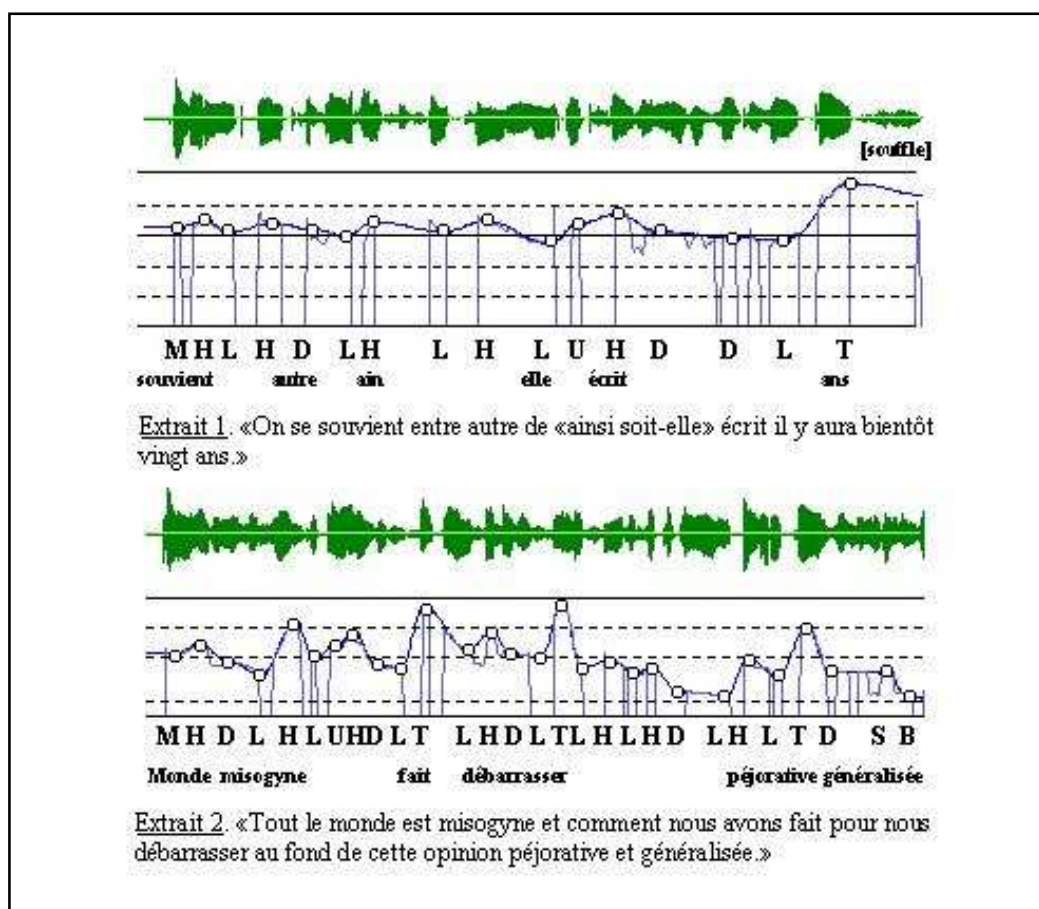


Figure 1. Illustrations du codage INTSINT pour deux extraits du corpus Fayard/Groult

3.3.3. Codage semi-automatique des structures prosodiques de surface : le cas de l'organisation temporelle

3.3.3.1. Remarques liminaires

Il a été établi dans plusieurs travaux que l'organisation des variables temporelles de la parole (durée des pauses, des syllabes, des mots, modifications du débit, etc.) a un rôle important à jouer dans l'élaboration et le traitement des énoncés et du discours (COUPER-KUHLEN & SELTING 1996 ; KOOPMANS-VAN BEINUM & VAN DONZEL 1996 ; SWERTS 1998 ; ZELLNER 1998 ; DUEZ 1987, 2001 ; AUER *et al.* 1999). Toutefois, l'analyse et le codage de l'organisation des variables temporelles s'avèrent particulièrement problématiques dans la mesure où ils présupposent une segmentation préalable des objets linguistiques (pauses, phonèmes, syllabes, etc.) à partir desquels il sera loisible de procéder à l'interprétation de cette organisation.

D'autre part, il est notoire que la durée des phonèmes formant les unités prosodiques de référence que sont les syllabes, est soumise à l'influence de multiples facteurs (DI CRISTO 1978). Nous exposons dans les lignes qui suivent les étapes de la démarche que nous adoptons pour apporter des éléments de réponse à ces problèmes. L'objectif à atteindre, en ce qui nous concerne, est de mettre au point une méthode fiable pour procéder à un codage automatique (ou semi-automatique) de l'organisation temporelle du corpus analysé, comparable au codage obtenu par l'approche INTSINT pour l'organisation tonale.

3.3.3.2. Segmentation phonémique et détection des pauses silencieuses

Nous avons procédé dans un premier temps à la segmentation manuelle et à l'étiquetage du corpus Fayard-Groult en phonèmes et en syllabes. Le choix de la segmentation phonémique manuelle est dicté par des raisons pratiques et il est tout à fait possible d'envisager ultérieurement le recours à une procédure de segmentation automatique.

Afin de détecter les pauses silencieuses du corpus (ante- et post-syllabiques), nous avons utilisé un script Praat permettant de sélectionner les parties non voisées du fichier sonore et de calculer les valeurs d'intensité correspondant à ces zones. Toute section non voisée du signal de parole d'une durée égale ou supérieure à 200 ms (valeur retenue dans la majorité des études, cfr CANDEA 2000 ; CAMPIONE 2001) et dont l'intensité moyenne est inférieure à la moyenne moins 2 écarts-types de celle mesurée sur l'ensemble du fichier est catégorisée comme pause silencieuse. Cette détection automatique est suivie d'une vérification manuelle destinée à corriger d'éventuelles erreurs.

3.3.3.3. Traitement de la durée phonémique et syllabique

L'objectif du traitement de la durée phonémique et syllabique consiste à dériver des valeurs de ce paramètre qui se rapprochent le plus possible de celles qui ont été programmées intentionnellement par les locuteurs et qui doivent ainsi faire l'objet d'un codage catégoriel sur la grille prosodique. Pour atteindre ce but il est indispensable de procéder à une normalisation des données brutes issues de la segmentation du corpus. Dans notre approche, cette normalisation se déroule en deux étapes. La première concerne la neutralisation des effets dits « co-intrinsèques », qui s'effectue en couplant un script Praat et un script en langage Perl afin de récupérer les durées phonémiques brutes et de leur appliquer un coefficient correcteur (DI CRISTO 1978) permettant d'effacer les effets du contexte consonantique sur la durée des voyelles. La seconde étape concerne la neutralisation des effets dits « intrinsèques », qui sont motivés, à l'instar des effets « co-intrinsèques », par des contraintes de bas niveau (de production, en l'occurrence). Nous avons recours dans

la mise en œuvre de cette seconde étape au *z-score* (« transformée Z », ou encore « variable centrée réduite »), inspiré des travaux de CAMPBELL (1992).

Les données issues de la double étape de normalisation font l'objet d'un transcodage automatique qui permet de classer les phonèmes et les syllabes du corpus analysé en quatre catégories : R (Réduit), N (Normal), L (Long) et XL (eXtra-Long). Ces données transcodées peuvent alors être confrontées à celles de l'analyse auditive (cfr 3.3.1), afin de juger des similitudes et des divergences engendrées par les deux approches.

3.3.3.4. *Traitement de la variable débit*

L'objectif du traitement de cette variable temporelle (orthogonale) est également de procéder à son codage automatique. Nous utilisons à cet effet un script Perl pour élaborer un algorithme dont la fonction principale conduit à coder comme

- épisode de parole ralenti, toute séquence supérieure ou égale à trois syllabes L ou XL ;
- épisode accéléré, toute séquence supérieure ou égale à trois syllabes R ;
- épisode normal, toute autre séquence.

De nouveau, les résultats de ce codage automatique peuvent être comparés à ceux obtenus à l'issue de l'analyse auditive.

3.3.3.5. *Traitement particulier du paramètre de F0 (interaction F0/durée)*

Certains travaux ont contribué à mettre en évidence le fait que les variations intra-syllabiques de F0 exercent une influence sur la durée subjective des noyaux vocaliques des syllabes (LEHISTE 1976 ; ROSEN 1977). Cet exemple d'interaction entre des paramètres prosodiques qui sont traités habituellement de manière indépendante, présente un réel intérêt dans l'optique de notre approche, car cette interférence participe de la construction d'un percept susceptible d'influencer le codage subjectif des indices prosodiques, en favorisant par exemple la notation [Le] ou [Xle] d'une syllabe donnée (cfr 3.3.1).

Dans le but de vérifier au moins partiellement cette hypothèse, nous avons procédé, au moyen d'un script Praat, au calcul de la vélocité du glissando de F0 (positif et négatif) dans le noyau des syllabes de notre corpus. Il s'agit là d'une mesure objective et nous avons renoncé momentanément à utiliser un algorithme de modélisation perceptive du glissando (ROSSI *et al.* 1981 ; 'T HART *et al.* 1991 ; D'ALESSANDRO & MERTENS 1995), afin de ne pas mêler, dans cette étape de notre travail, des données objectives et des données subjectives (même prédites). En revanche, le traitement particulier du paramètre F0 que nous avons effectué sera pris en compte dans l'estimation des variables susceptibles de contribuer à la prédiction des résultats obtenus lors de l'approche auditive.

3.3.3.6. *Mise en relation des données issues des divers traitements et des résultats de l'analyse auditive*

Dans cette étape de notre démarche méthodologique, nous cherchons à confronter les données extraites des divers traitements précédents avec celles qui découlent de la notation auditive du corpus. Cette comparaison concerne notamment la prédiction du paramètre de durée vocalique dont nous allons rendre compte ci-après.

La méthode que nous avons retenue pour mettre en évidence les relations linéaires qui peuvent s'établir entre les variables de notre analyse, est celle des régressions multiples (réf. logiciel GUIDE : *Generalized, Unbiased, Interaction Detection and Estimation*, développé par WEI-YIN LOH du Département de Statistiques de l'Université de Wisconsin-Madison). Les régressions multiples permettent de montrer de quelle manière, et dans quel ordre, les variables indépendantes interagissent pour « configurer » une variable dépendante donnée (ici le percept de durée syllabique subjective, tel qu'il ressort de l'analyse auditive). Le dispositif graphique choisi pour représenter l'interaction des variables se présente sous la forme d'un arbre dont les nœuds hiérarchiquement les plus élevés représentent les variables qui ont le rôle le plus important dans la détermination de la variable dépendante (Figure 2).

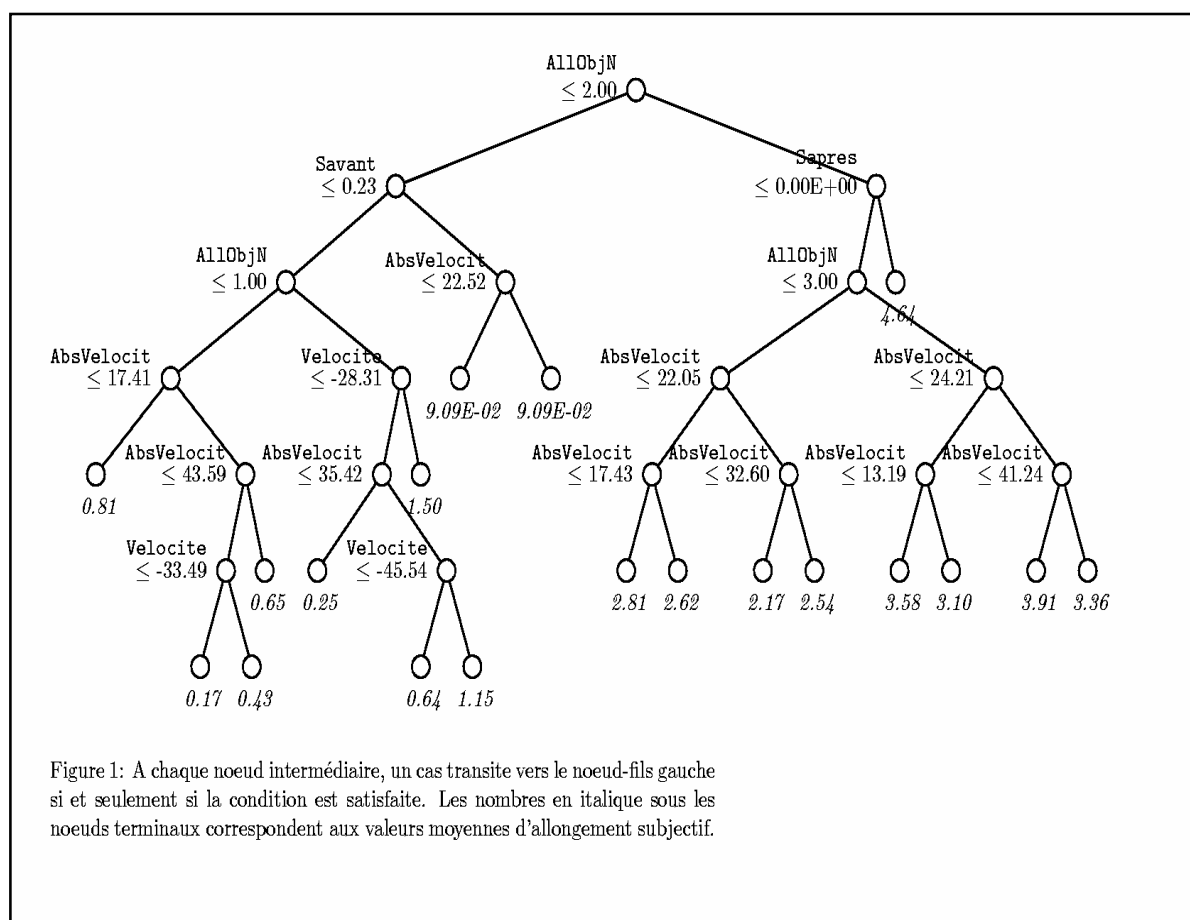


Figure 1: A chaque noeud intermédiaire, un cas transite vers le noeud-fils gauche si et seulement si la condition est satisfaite. Les nombres en italique sous les noeuds terminaux correspondent aux valeurs moyennes d'allongement subjectif.

Figure 2. Arbre de régression sur la variable dépendante « allongement subjectif »

Il n'est pas possible pour des raisons de place de décrire dans le détail les informations que l'on peut retirer de l'examen de la Figure 2. Pour résumer, nous dirons que cette analyse permet de montrer que trois variables concourent à expliquer les jugements subjectifs d'allongement formulés dans le test auditif, soit par ordre d'importance :

- la durée normalisée et transcodée (cfr 3.3.3.3),
- la pause silencieuse post-syllabique (cfr 3.3.3.2),
- la vélocité du glissando (cfr 3.3.3.5).

Ces résultats nous paraissent intéressants à plus d'un titre. D'abord, ils ouvrent la voie d'une démarche originale du codage automatique de la durée syllabique qui ne se borne pas à la mise en œuvre d'une procédure de normalisation traditionnelle. Ce faisant, ils confirment indirectement l'intérêt d'une approche multilinéaire des représentations prosodiques informée par la valeur interactive des paramètres prosodiques. Ces interactions participent de la formation de percepts prosodiques à partir desquels se construisent les représentations phonético-formelles de la prosodie. Ces dernières s'intègrent alors au réseau d'interfaces formes/fonctions d'où procède

l'interprétation du discours. Il apparaît dès lors que la multilinéarité interactive s'impose comme un outil nécessaire à tous les niveaux de l'analyse.

Enfin, il est permis d'envisager à court terme une application de la méthodologie qui vient d'être exposée à la reconnaissance et à la synthèse de la parole, dans le but d'améliorer notamment la finesse et la robustesse de certains traitements automatiques de la prosodie.

3.4. Constitution de la grille prosodique

La construction de la grille prosodique rassemble, dans son état présent, des informations en provenance de différentes sources (notations auditives, représentations analytiques et codages automatiques et semi-automatiques). Elle tend idéalement vers une représentation multilinéaire où ne figureraient en définitive, outre les représentations analytiques, que les représentations phonétiques de la prosodie et les codages catégoriels de ses vecteurs syntagmatiques et orthogonaux, issus de la mise en œuvre des procédures automatiques disponibles et en cours d'élaboration.

Nous exposons brièvement dans les lignes qui suivent la constitution de la grille prosodique que nous utilisons actuellement pour l'analyse des relations de la prosodie au discours. Pour la compréhension de cette présentation il sera utile de se reporter à la grille 1 (Figure 3 dans la section 6, *infra*) qui présente une analyse complète du premier Paraton du corpus Fayard-Groult :

- (1) On se souvient entre autres de « ainsi soit-elle » écrit il y aura bientôt vingt ans bien d'autres livres sont sortis depuis et vous n'avez cessé durant toutes ces années donc de vous pencher sur le sort des femmes et de continuer à défendre leur cause

(Nota : le texte est délibérément présenté sans ponctuation).

- **Osc.** Cette première ligne de la grille renvoie à l'oscillogramme (représentation analytique du signal physique). Toutes les lignes subséquentes sont alignées sur ce signal de référence.
- **Momel.** Cette ligne renvoie à la courbe de F0 modélisée sous la forme de l'interpolation d'une séquence de points-cibles (cfr 3.3.2).
- Ces deux premières lignes contiennent également des informations relatives au découpage en **Unités Intonatives (UI)**, selon les critères proposés par HIRST & DI CRISTO (1998). Chaque UI est numérotée en fonction de son ordre d'apparition : UI1, UI2, etc. De plus, la numérotation permet d'indiquer si une UI se trouve être enchâssée dans une autre. C'est ainsi que par suite de l'enchâssement de l'UI4 dans l'UI3, cette dernière est scindée en deux tronçons identifiés respectivement par UI3a et UI3b.
- **Ortho.** Cette ligne correspond à la transcription orthographique du texte, grossièrement alignée sur le signal. Pour faciliter la lisibilité de cette

notation, nous employons une transcription normative qui n'intègre que les diacritiques aptes à rendre compte des particularités de l'oral (troncation de phonèmes ou de morphèmes entre parenthèses, par exemple). En revanche, toute ponctuation, marque spécifique de l'écrit, est exclue.

À partir de ce point, toutes les lignes de la grille prosodique sont dévolues d'une part au codage catégoriel des niveaux de la représentation formelle-syntagmatique et orthogonale (3.1.2.2) et, d'autre part, à celui de la représentation fonctionnelle (3.1.2.4) des entités prosodiques.

- **INTSINT.** Cette ligne présente le codage catégoriel des points-cibles de la représentation phonétique (ligne Momel), au moyen de l'alphabet intonétique INTSINT (cfr 3.3.2).
- **Dd (LH), Dd (T), Dd (UI).** Cette ligne initie la notation des indices prosodiques relevant de la composante orthogonale. Sur cette ligne sont notés les effets dits d'abaissement (« *downtrends* ») qui concernent les relations de hauteur entre des valeurs-cibles ponctuelles codées par INTSINT à l'intérieur d'un domaine donné (délimité par des parenthèses) ou entre des domaines eux-mêmes. C'est ainsi que Dd (LH) fait référence à un abaissement dans une séquence de LH, dont il est possible de spécifier la pente, que Dd (T) permet de signaler un effet d'abaissement entre des maxima successifs, et que Dd (UI) servira à noter un abaissement de registre entre des UI consécutives.
- **ReSp.** Cette ligne est destinée à la notation catégorielle des variations de la gamme tonale (ReSp = « *Register Span* »), c'est-à-dire au codage des variations de l'intervalle de hauteur entre les cibles hautes et basses dans une séquence donnée. Un intervalle moyen (exprimé en demi-tons ou en valeurs ERB : HERMES & VAN GESTEL 1991) est noté N (Normal), un intervalle significativement supérieur à la moyenne est noté Ex (« *Expanded* ») et un intervalle significativement inférieur à la moyenne est noté Com (« *Compressed* »). Dans le cas d'une suite de cibles de même niveau, l'absence de variation de la gamme tonale sera notée par Ø (zéro).
- **ReLe.** Cette ligne sert à la notation du niveau global du registre (ReLe = « *Register Level* ») pour un domaine donné et à celle des réajustements locaux de ce paramètre. En ce qui concerne le premier cas, nous utilisons les symboles N (Normal), Rai (« *Raised* ») et Low (« *Lowered* »), pour noter, respectivement le registre moyen de référence du locuteur (ou sa dynamique de base), un rehaussement et un abaissement du registre par rapport à cette référence. Pour le second cas, nous employons le symbole Re (« *Reset* ») afin de signaler un phénomène abondamment décrit dans la littérature (WICHMANN 2000) qui concerne un retour au niveau de référence du locuteur indiquant l'ouverture d'une nouvelle unité prosodique majeure.

Les lignes de la grille prosodique qui vont faire dès maintenant l'objet d'une présentation concernent en premier lieu le codage des marques syntagmatiques et orthogonales de l'organisation temporelle. Nous avons également inclus dans cette dernière partie de la grille la notation des prises de souffle pour des raisons uniquement pratiques. Il est souhaitable cependant que dans une version ultérieure de la grille, nous réservions un espace spécifique à la notation des phénomènes de performance qui, outre les prises de souffle, concernent également les laryngalisations, les toux, les raclements de gorge et autres bruits audibles qui apportent des informations pouvant donner lieu à diverses interprétations.

- **Brea.** Sur cette ligne sont reportées les prises de souffle, notées par la lettre B (Brea = « *Breathing* ») et alignées avec les pauses auxquelles elles sont associées.
- **Pause.** Les pauses sont notées [#] et [##] pour signaler, respectivement, une pause simple et une pause longue. Cette notation peut être uniquement dérivée de l'analyse auditive du corpus. Il est également possible, si la précision de la description l'exige, de consigner entre crochets la durée physique de la pause en millisecondes. Dans la grille de la Figure 3 figurent également sous les symboles # et ## le nombre de juges (sur un total de sept) qui ont reconnu la présence d'une pause, « s » renvoyant à simple, et « d » à double.
- **Length.** Cette ligne recense les variations de la durée syllabique pertinentes (perçues et/ou détectées automatiquement). Quatre symboles sont utilisés à cet effet : N (*Normal*), R (« *Reduced* ») : durée réduite par rapport à la durée normale, Le (« *Lengthened* ») : syllabe longue, et Xle (« *extra-lengthened* ») : syllabe très longue. La grille 1 indique encore le nombre de réponses obtenues pour les sept juges.
- **Tempo.** La variable orthogonale de l'organisation temporelle que constitue le tempo sera notée sur cette ligne au moyen des trois symboles suivants : N (*Normal*), Fa (« *Faster* ») et Slo (« *Slower* »), en indiquant par des parenthèses l'empan du domaine sur lequel s'applique cette notation.

Dans sa version actuelle, la grille prosodique comporte deux lignes réservées à la composante métrique, qui sont identifiées par les abréviations [Proe] et [Met]. Un codage plus élaboré de la représentation de cette composante est à l'étude et sera présenté dans une version ultérieure de la grille.

- **Proe.** Sur cette ligne est consignée la distribution des proéminences perçues (P) lors de l'analyse auditive. Sur la grille 1, les chiffres souscrits et entre parenthèses indiquent encore le nombre de réponses obtenues pour les sept juges.
- **Met.** Cette ligne propose de noter une quantification des niveaux de proéminence, selon une échelle croissante allant de 1 à 3.

- **F.** La dernière ligne de la grille prosodique concerne la notation syntagmatique et paradigmatique des fonctions (prosodiques) dont l'inventaire a été présenté dans la section 3.1.2.4. Le terme « prosodique » est mis entre parenthèses pour signifier que nous n'admettons pas *a priori* que l'identification de ces fonctions soit exclusivement fondée sur des indices prosodiques. Comme nous l'avons déjà suggéré antérieurement, il est probable que cette identification procède de l'interprétation de faisceaux d'indices prosodiques et textuo-discursifs distribués de façon non uniforme. Il n'est pas indifférent, de ce point de vue, que la ligne de notation des fonctions (prosodiques) se positionne, dans la grille générale d'analyse, à une place qui permet de l'interpréter comme une interface entre la grille prosodique et la grille textuo-discursive.

4. ILLUSTRATIONS SUR DEUX EXEMPLES EXTRAITS DU CORPUS FAYARD-GROULT

Comme nous l'avons dit plus haut, l'établissement de la grille d'analyse multilinéaire que nous proposons a pour objectif principal de permettre la visualisation et la mise en évidence de correspondances et discordances significatives entre les différents niveaux de l'analyse représentés pour chaque extrait analysé.

Afin d'illustrer notre propos, nous avons choisi d'établir la grille pour deux extraits du corpus Fayard-Groult, corpus commun à tous les participants du colloque. Les deux interlocutrices de ce corpus d'interview sont représentées : la journaliste, R. Fayard, dans la grille 1, et l'interviewée, B. Groult, dans la grille 2. La première grille présente un intérêt du point de vue de la structuration discursive ; nous avons choisi l'extrait représenté dans la seconde grille pour sa richesse et son ambiguïté du point de vue polyphonique.

Seule la grille 1 sera traitée (Figure 3 dans la section 6). La grille 2 et les commentaires qui s'y rapportent sont présentés en annexe sur le CD-ROM.

4.1. Présentation générale de la grille 1

4.1.1. Extrait concerné

Nous reprenons ici l'extrait du corpus Fayard-Groult déjà cité dans la section 3.4 :

- (1) On se souvient entre autres de « ainsi soit-elle » écrit il y aura bientôt vingt ans bien d'autres livres sont sortis depuis et vous n'avez cessé durant toutes ces années donc de vous pencher sur le sort des femmes et de continuer à défendre leur cause

Afin de ne rien perdre de l'analyse des extraits choisis, il importe au préalable de les resituer précisément dans leur contexte de production. L'extrait analysé par la grille 1 correspond aux tout premiers mots prononcés par la journaliste au début de l'interview elle-même, si ce n'est de l'émission (peut-être y a-t-il eu d'abord des propos concernant l'émission elle-même, son titre, son thème, une présentation préalable de l'invitée : nous n'en savons rien et n'en disposons pas). Ce positionnement des propos par rapport aux différents moments de l'échange joue bien sûr un rôle dans l'interprétation qui en sera donnée. Il faudrait du reste ajouter à la grille une ou plusieurs lignes qui permettent d'y intégrer ces informations. Cependant, l'état que nous présentons ici de notre outil d'analyse est centré, pour cet extrait, sur un aspect particulier du fonctionnement discursif, à savoir son aspect structurel.

4.1.2. Objectifs

La prise en compte des concordances entre les différents paramètres observables notés dans la grille 1 permet avant tout de dégager deux **moments** dans l'extrait présenté, et donc de structurer cet extrait sur le plan organisationnel. Ces deux moments peuvent être mis en évidence tant au niveau prosodique qu'aux niveaux lexical et énonciatif, qui sont les deux « niveaux » textuels auxquels cette bipartition est la plus marquée.

Le premier moment est constitué de :

on se souvient entre autres de Ainsi soit-elle écrit il y aura bientôt vingt ans bien d'autres livres sont sortis depuis,

le second de :

vous n'avez cessé durant toutes ces années donc de vous pencher sur le sort des femmes et de continuer à défendre leur cause.

Les deux moments sont articulés par *et*. Cependant, la prise en compte d'autres paramètres de nature énonciative et référentielle permettrait de proposer une autre structuration de cet extrait, pour laquelle le premier moment serait constitué de :

on se souvient entre autres de Ainsi soit-elle écrit il y aura bientôt vingt ans

et le second de :

bien d'autres livres sont sortis depuis et vous n'avez cessé durant toutes ces années donc de vous pencher sur le sort des femmes et de continuer à défendre leur cause.

Cette grille a pour objectif de répertorier les indices en faveur de l'une ou l'autre organisation, et de montrer que les indices de la première structuration atteignent probablement un poids plus important que ceux de la seconde, notamment si l'on prend en compte les indices de type prosodique.

Un des intérêts méthodologiques de cette grille réside en effet dans la confrontation des informations obtenues au niveau prosodique et au niveau textuel. Il

semble que la prosodie structure le paragraphe oral représenté ici de la même façon que la majorité des paramètres textuels.

La lecture multidimensionnelle de la grille (lecture horizontale permettant le repérage d'itérations et de contrastes syntagmatiques, et lecture verticale mettant en évidence les conjonctions ou disjonctions d'indices provenant de sources analytiques différentes) permet ainsi de faire des observations importantes à différents niveaux de profondeur dans l'analyse et à différents stades d'intégration des lignes descriptives horizontales. Ces recouvrements peuvent impliquer la dimension prosodique seule, l'aspect textuel seul, ou les deux à la fois. Les recouvrements les plus intéressants pour nous mobilisent conjointement ces deux dimensions. Nous présenterons les différentes observations que permet la grille 1 en fonction des niveaux d'analyse et des dimensions qu'elles mettent en jeu.

4.2. Niveaux prosodiques

4.2.1. Observations horizontales

4.2.1.1. *Itération de cadences*

Le codage au moyen de l'alphabet INTSINT (ligne INTSINT) révèle l'existence en fin d'unités intonatives (UI, numérotées entre les lignes Osc et MOMEL pour les repères) de cadences mélodiques que l'on considère comme comparables, codées respectivement DDLT1 pour UI1 et DLT2 pour UI2.

UI1= on se souvient entre autres de Ainsi soit-elle écrit il y aura bientôt vingt ans

UI2= bien d'autres livres sont sortis depuis.

En effet, le D (comme le U) a la particularité de pouvoir être réitéré plusieurs fois sans affecter la forme globale de la structure mélodique (DI CRISTO 1998). L'écoute de l'enregistrement audio confirme la proximité perceptive des deux cadences. Elles s'avèrent très différentes par exemple des cadences de UI3 (codée HLT3) et de UI4 (codée HLT4) :

UI3= et vous n'avez cessé (...) de vous pencher sur le sort des femmes

UI4= durant toutes ces années donc

Ces cadences ont pourtant la même fonction continuative (-T sur la ligne des fonctions F) comme celle d'UI1 et UI2 et fonctionnent aussi en écho. Ces échos intonatifs contribuent à solidariser UI1 et UI2 d'une part, UI3 et UI4 d'autre part, et à opposer ainsi, par contraste, ces deux groupes d'UI. Ces informations pourront

ensuite être mises en relation avec d'autres regroupements ou d'autres contrastes, congruents ou non, issus des autres niveaux descriptifs.

4.2.1.2. *Contrastes d'abaissement*

On peut comparer ces contrastes de cadence au contraste d'abaissement représenté sur la ligne Dd(T) (Downtrend relatif aux cibles Top). Cette ligne met en évidence que les cibles T1 et T2 qui terminent respectivement l'UI1 et l'UI2 atteignent la même hauteur mélodique. Cette égalité contribue à solidariser UI1 et UI2. Elle renforce bien sûr l'impression perceptive de similarité entre les cadences mélodiques qui terminent UI1 et UI2. En outre, cette égalité annule l'effet d'abaissement sur le bloc formé par ces deux UI alors que cet effet apparaît, par contraste, sur les unités intonatives suivantes ($T2 > T3 > B$; T4 n'entre pas dans l'abaissement car UI4 a un statut particulier d'incise). Le contraste d'abaissement ainsi mis en évidence se trouve être en congruence avec le contraste mélodique décrit à partir des cadences. Les parenthèses centrales (en caractères gras sur la grille 1) figurent ainsi les limites des domaines en contraste, dont la congruence est particulièrement significative à de très nombreux niveaux d'analyse, tant textuels que prosodiques, pour l'extrait traité dans la grille 1. Nous interpréterons ultérieurement, à la lumière de l'ensemble de l'analyse, la portée de ces congruences multiples.

4.2.2. **Observations verticales**

Rares sont les indices prosodiques qui fonctionnent isolément. Dans la grande majorité des cas, c'est une conjonction d'indices prosodiques provenant des trois ordres de structuration tonal, temporel et métrique qui réalise la mise en forme prosodique des segments textuels. Cette association d'indices est clairement illustrée dans la grille 1 par la mise en forme prosodique spécifique des deux premières UI et par deux constructions formelles : la mise en forme d'un accent emphatique sur la dernière syllabe du mot *écrit* ; la mise en forme prosodique particulière de la mention du titre de l'ouvrage de B. Groult *Ainsi soit-elle*.

4.2.2.1. *Mise en forme prosodique des deux premières UI*

Reprenons le cas des cadences de UI1 et UI2. Nous avons déjà noté que leur ressemblance, outre la proximité de leur codage INTSINT, était liée à l'égalité du niveau mélodique atteint par T1 et T2. L'examen des autres niveaux d'analyse prosodique met en évidence le même allongement affectant les syllabes finales de UI1 et UI2, mais surtout, l'alignement des deux cadences avec des épisodes d'accélération locale du tempo qui mettent vraiment en parallèle la fin des deux énoncés. Cette accélération est marquée sur la ligne métrique par l'absence de

proéminence précédant la proéminence finale (de niveau 3 sur la grille métrique). Sur l'empan du domaine d'accélération (qui est aussi celui de la cadence : *il y aura bientôt vingt ans et sont sortis depuis*), il n'y a en effet que cette proéminence finale.

Les cadences finales de UI1 et UI2 présentent donc des ressemblances beaucoup plus marquées que par exemple, les cadences de UI3 et UI4 qui n'ont en commun que le codage INTSINT. Ces ressemblances impliquent non seulement des aspects tonals (ceux que l'on prend en compte le plus souvent dans la description des cadences) mais également des aspects temporels (allongements et surtout accélération du tempo) et des aspects métriques (absence de proéminence précédant l'accent final). Les trois ordres d'organisation prosodique contribuent donc ici à mettre en parallèle, de façon très nettement audible, la mise en forme prosodique des deux premières UI de l'extrait.

4.2.2.2. Construction d'un accent emphatique

L'emphase attribuée sur le plan fonctionnel à la seconde syllabe du mot *écrit* (voir la ligne F) provient de la perception d'une proéminence (ligne Proe) construite à partir de la conjonction d'un ton haut sur la ligne mélodique (ligne INTSINT), associé à un allongement important de la syllabe (XLe sur la ligne Lengthg) qui n'est pas usuel dans cette position : nous ne sommes pas à la fin d'une Unité Intonative où cette taille d'allongement est courante. Secondairement, on note l'occurrence conjointe, immédiatement après la syllabe allongée, d'une frontière de domaine défini par un contraste de débit de parole (ligne tempo, fin du premier domaine N, début du premier domaine Fa)³.

La conjonction du ton haut, de l'allongement et de la rupture de tempo construit donc la perception emphatique de la syllabe finale du mot *écrit*. La partie textuelle de la grille doit ensuite permettre de rendre compte de la valeur qu'il est possible d'accorder à la présence de cette emphase dans le processus d'interprétation.

³ On remarquera enfin la pause brève identifiée immédiatement après la syllabe accentuée. Il faut préciser ici qu'il ne s'agit pas d'une pause silencieuse mais d'une pause dite « subjective » dont la perception est liée à la présence de l'allongement. Elle figure sur la grille car sa présence a été perçue par l'analyste. Il s'avère donc indispensable de trouver un codage distinct pour les pauses subjectives et les pauses silencieuses, codage qu'il nous faudra prévoir dans la prochaine version de notre grille.

4.2.2.3. Mise en forme prosodique de la mention *Ainsi soit-elle*

Ainsi soit-elle est le titre du livre le plus connu de B. Groult, le seul titre que mentionne la journaliste. Cette mention bénéficie d'une mise en forme prosodique particulière construite à partir des indices suivants.

On identifie une proéminence pour chacune des deux syllabes frontières de la mention, c'est-à-dire sur la syllabe « ain » de *Ainsi* et sur la syllabe « elle » de *soit-elle* (voir ligne Proe). Ces proéminences sont liées à la présence de deux tons hauts dans la suite mélodique HLH alignée avec la mention. On note aussi l'allongement (simple=Le sur la ligne Lengthg) de la syllabe finale de la mention « elle ».

L'ensemble forme un « arc accentuel » qui met en valeur la particularité de la mention vis-à-vis des autres segments textuels : la mention est en effet un tout indissociable bien qu'il soit constitué de mots courants. La mise en forme prosodique au moyen de l'arc accentuel convient particulièrement au marquage de la cohérence interne d'un groupe textuel, comme c'est le cas ici.

Remarquons cependant un autre phénomène qui n'apparaît pas dans notre version de la grille mais qui mériterait d'y figurer : l'absence de liaison entre *de* et *Ainsi* marque elle aussi le traitement particulier réservé à la mention qui n'est plus analysable en morphèmes distincts comme le sont habituellement les syntagmes. Les phénomènes de liaison ont la particularité de mêler les niveaux morphologique et acoustico-phonétique. L'analyse de cette complexité s'avère particulièrement appropriée à notre approche d'où l'importance de réserver une ligne à ce genre de phénomènes dans les versions ultérieures de notre outil.

L'intérêt de la grille se justifie donc au seul niveau prosodique puisqu'elle permet d'analyser les composants indiciels d'une fonction (l'emphase) ou la conjonction d'indices qui construit une mise en forme prosodique spécifique (l'arc accentuel). Cependant, il va sans dire que la mise en forme prosodique ne peut être réellement interprétée qu'en relation avec les informations délivrées par l'analyse textuelle. Nous illustrons cette intégration dans la section suivante.

4.3. Niveaux textuels/discursifs

4.3.1. Syntaxe

Sur le plan de la syntaxe, une analyse classique en grands constituants donne trois unités :

- 1 on se souvient entre autres de *Ainsi soit-elle* écrit il y aura bientôt vingt ans
- 2 bien d'autres livres sont sortis depuis

- 3 et vous n'avez cessé durant toutes ces années donc de vous pencher sur le sort des femmes et de continuer à défendre leur cause.

La syntaxe ne donne pas d'indication sur le lien entre ces différentes unités. On peut dès lors se demander si, au niveau discursif, c'est-à-dire dans les stratégies de production et dans la construction de l'interprétation, ces unités sont davantage structurées. Trois organisations apparaissent *a priori* possibles :

- les trois constituants maximaux s'enchaînent simplement sans autre ordre de structuration (1-2-3) ;
- les deux premiers se regroupent en une macro unité plus vaste ([1-2]-3) ;
- les deux derniers se regroupent (1-[2-3]).

4.3.2. Lexique

Au niveau lexical, le type de verbes employés permet d'isoler deux phases dans l'extrait, ces deux phases s'organisant en une structuration de type ([1-2]-3) :

La première phase (*on se souvient entre autres de Ainsi soit-elle écrit il y aura bientôt vingt ans bien d'autres livres sont sortis depuis*) comporte des verbes processuels indiquant la transformation d'un objet (écrire quelque chose, sortir quelque chose). Ces verbes se retrouvent dans les deux premières unités syntaxiques maximales, c'est-à-dire dans les deux premières clauses, ce qui constitue un argument en faveur de leur regroupement.

La seconde phase (*vous n'avez cessé durant toutes ces années donc de vous pencher sur le sort des femmes et de continuer à défendre leur cause*), qui correspond à la troisième clause isolée par la syntaxe, comporte des opérateurs aspectuels duratifs (cesser de, continuer à), appliqués à des processus non résultatifs (se pencher sur, défendre).

Sur le plan des sujets employés dans chacune de ces phases, on observe que la première comporte deux sujets inanimés (*Ainsi soit-elle*, qui constitue une mention du titre d'un ouvrage, et *bien d'autres livres*), alors que la seconde comporte uniquement le sujet personnel et animé *vous*.

4.3.3. Énonciation

Ces phénomènes lexicaux prennent sens au niveau énonciatif sur le plan des instances énonciatives mises en jeu dans l'échange d'une part, et en se combinant avec les temps employés d'autre part, et confirment une organisation de type ([1-2]-3).

Dans la première phase de l'extrait, la locutrice Fayard présente son invitée B. Groult. Elle met alors en scène son discours en recourant à un énonciateur collectif représentant l'ensemble des lecteurs et auditeurs potentiels. Cet énonciateur collectif confère au discours un caractère moins intimiste (incluant les auditeurs) et

qui sied davantage au type d'interaction qui se joue, en l'occurrence une interview radio et non une interview privée. À la fin de cette présentation, la locutrice Fayard redevient l'énonciatrice principale dans sa parole. Elle se pose alors comme pleinement responsable de ce qu'elle avance et réinitie totalement sa place institutionnelle de journaliste (intervieweur) en posant explicitement (« vous ») B. Groult à sa place d'allocataire (interviewée). Cet extrait comporte donc deux types de positionnements énonciatifs spécifiques.

Concernant les temps employés, dans la première phase de l'extrait, les deux verbes processuels résultatifs (écrire quelque chose, sortir quelque chose) apparaissent sous la forme de participes passés adjectivaux (dans des formes de passé composé); au niveau aspectuel, les procès qu'ils désignent sont donc considérés comme accomplis au moment de l'énonciation M°. Dans la seconde phase de l'extrait, le verbe aspectuel duratif *cesser de* est sous la portée d'une négation et employé au passé composé : *vous n'avez cessé de* ; le résultat de l'interprétation de cette construction verbale en termes de repérages par rapport au moment de l'énonciation est que les procès *vous pencher* et *continuer à défendre la cause des femmes* sont encore en cours au moment de l'interview. On a donc affaire, sur le plan des repérages temporels, à une distinction entre ce qui s'est passé avant l'interview, et ce qui est en cours au moment de l'interview.

Au niveau énonciatif toujours, les adverbes temporels jouent un rôle dans l'établissement d'une autre structuration de l'extrait, cette fois de type (1-[2-3]). Trois adverbes (ou expressions jouant le rôle d'adverbe) participent à cette autre organisation :

il y aura [x ans]

bientôt

depuis,

auxquels on peut ajouter l'expression *durant toutes ces années*, qui sera décrite au niveau référentiel, en vertu de la référence du SN *toutes ces années* qu'elle comporte.

Les deux premiers adverbes, dont l'un est partie intégrante de l'autre (*il y aura bientôt 20 ans*), construisent un repérage temporel complexe par rapport au moment de l'énonciation M° : l'adverbe déictique *bientôt* introduit dans l'univers du discours un moment M' repéré comme postérieur à celui de l'énonciation, et proche de celui-ci ; l'expression globale *il y aura bientôt 20 ans* construit, par rapport à ce moment M' pris comme nouveau repère, un second moment M'', antérieur à celui-ci de 20 ans, et donc antérieur à celui de l'énonciation. Globalement, l'expression *il y aura bientôt 20 ans* est interprétable déictiquement, c'est-à-dire comme plus ou moins équivalente à *il y a presque 20 ans*. Ce repérage a pour but de situer

chronologiquement la parution de l'ouvrage *Ainsi soit-elle* par rapport au moment de l'interview M° (presque 20 ans avant, donc).

Avec *bien d'autres livres sont sortis depuis*, ce nouveau moment M'' est à son tour choisi comme repère temporel : l'adverbe *depuis*, qui a ici une référence relative à M'', construit un intervalle temporel ouvert à droite à partir du moment de la parution de *Ainsi soit-elle*. Cet intervalle n'est pas explicitement refermé dans la suite de l'extrait, et on peut donc dire que les deux « étapes » construites par le jeu des adverbes s'organisent de la façon suivante :

- une première étape discursive, avec *on se souvient entre autres de Ainsi soit-elle écrit il y aura bientôt vingt ans*, rappelle un événement clé dans l'historique des publications de Benoîte Groult ;
- une seconde étape discursive, avec *bien d'autres livres sont sortis depuis et vous n'avez cessé durant toutes ces années donc de vous pencher sur le sort des femmes et de continuer à défendre leur cause*, met en scène un « après » *Ainsi soit-elle*, qui peut inclure le moment de l'interview.

Ces deux étapes ne concordent donc pas avec celles isolées précédemment. Cependant, la combinaison entre les temps verbaux et les adverbes temporels invite à inférer que le repère principal intervenant dans ces repérages temporels est bien celui du moment de l'énonciation M°. La clause *bien d'autres livres sont sortis depuis*, si elle pose un repère M'' à travers l'adverbe *depuis*, renvoie implicitement au repère M° via l'emploi de l'accompli *sont sortis*. En définitive, cette clause pose donc implicitement une borne droite à l'intervalle temporel ouvert par *depuis*, qui est le moment de l'énonciation M° : elle construit l'image d'un intervalle temporel allant du moment de la parution de *Ainsi soit-elle* au moment de l'interview, image qui pourra être reprise dans la suite du discours par l'expression *toutes ces années* (voir niveau référentiel).

Cette combinaison invite donc à voir la clause 2 comme opérant une transition entre les informations introduites par la clause 1 et celles introduites par la clause 3. Elle n'est donc pas décisive en faveur de l'une ou l'autre structuration interne du passage.

4.3.4. Sémantique actantielle

La distinction entre les deux phases évoquées précédemment (structuration [1-2]-3) est par ailleurs renforcée si l'on se place au niveau d'une sémantique actantielle pour décrire les procès et les rôles des actants impliqués dans ces procès. L'extrait comporte quatre constructions verbales :

on se souvient

Ainsi soit-elle écrit

bien d'autres livres sont sortis

vous n'avez cessé de vous pencher [...] et de continuer à défendre [...]

Eu égard aux caractéristiques lexicales des verbes et de leurs sujets, on peut dire que seule la deuxième phase de l'extrait présente le sujet (unique dans ce cas, étant donné qu'il s'agit du sujet personnel animé *vous*) comme remplissant un rôle d'agent. On peut en effet dire qu'un constituant sujet renvoie à un rôle actantiel d'agent si le prédicat verbal est de type processus (i.e. non statique) et si ce sujet est animé. Dans la première phase de l'extrait, *on se souvient* comporte un verbe plutôt statique (*se souvenir*) et un sujet animé (*on*), ce qui fait qu'on peut accorder à *on* un rôle d'expérimenteur, mais pas d'agent. Par ailleurs, toujours dans cette même phase, *Ainsi soit-elle* est interprétable comme un patient du procès *écrire*, de même que *bien d'autres livres* constitue le patient du procès *sortir*, étant donné que ces deux sujets ne sont pas animés et que les prédicats processuels supposent l'existence d'un agent autre.

4.3.5. Référence et structure informationnelle

Il s'ensuit que les deux phases distinguées (structuration [1-2]-3) s'opposent également par le caractère présenté comme statique vs dynamique des objets dont on parle. Des propriétés statiques sont attribuées au référent de *on*, à l'ouvrage *Ainsi soit-elle* et aux autres livres, alors que des propriétés dynamiques sont affectées au référent de *vous* qui désigne l'allocutrice Benoîte Groult.

Le niveau catégorisé ici « sémantique actantielle » est donc fortement lié à (à des incidences sur) un niveau référentiel, auquel les objets se voient pourvus de propriétés discursivement présentées comme statiques vs comme dynamiques.

L'importance du caractère statique vs dynamique de ces propriétés ne prend sens que parce qu'au niveau de la structure informationnelle des énoncés, les objets pourvus de ces propriétés sont thématiques. Les deux moments ne se distinguent pas seulement par le fait que les sujets ont des propriétés différentes, mais par le fait que leurs thèmes ont des propriétés différentes.

Dans l'extrait présenté, *on*, *bien d'autres livres* et *vous* sont des sujets en position thématique. Le cas de *Ainsi soit-elle* est plus délicat, dans la mesure où l'on peut considérer que *écrit il y aura bientôt vingt ans* constitue soit une apposition (dont *Ainsi soit-elle* constitue un support thématique), soit une simple expansion adjectivale d'un constituant rhématique dans la clause *on se souvient entre autres de Ainsi soit-elle écrit il y aura bientôt vingt ans* (c'est d'ailleurs cette dernière option qu'a choisie l'analyse syntaxique).

Au niveau référentiel toujours, l'expression *toutes ces années* pourrait être considérée comme une anaphore associative par rapport à l'intervalle temporel précédemment construit par les adverbes temporels : *toutes ces années* désigne en

effet la classe des années séparant le moment de la parution de *Ainsi soit-elle* du moment de l'interview. Cet intervalle temporel permet d'argumenter en faveur d'une distinction dans le discours entre une première étape de présentation de l'ouvrage majeur de Benoîte Groult (*on se souvient entre autres de Ainsi soit-elle écrit il y aura bientôt vingt ans*), et une seconde étape (*bien d'autres livres sont sortis depuis et vous n'avez cessé durant toutes ces années donc de vous pencher sur le sort des femmes et de continuer à défendre leur cause*) qui retrace la continuité de l'itinéraire militant de l'écrivain depuis cette parution (structuration 1-[2-3]). Mais, dans la mesure où l'anaphore joue également un rôle résomptif, on peut dire aussi que le niveau référentiel permet de faire ressortir deux phases :

- mise en place d'un intervalle temporel antérieur à l'énonciation, par *on se souvient entre autres de Ainsi soit-elle écrit il y aura bientôt vingt ans bien d'autres livres sont sortis depuis* ;
- « résomption » de cet intervalle pour situer l'auteur au moment de l'énonciation, par *vous n'avez cessé durant toutes ces années donc de vous pencher sur le sort des femmes et de continuer à défendre leur cause*.

Ces deux phases correspondent à une organisation de type ([1-2]-3).

4.3.6. Marqueurs discursifs

Ajoutons enfin que la présence des marqueurs discursifs serait à étudier plus en détail : les connecteurs *et* et *donc* employés dans l'extrait jouent leur rôle dans la construction d'une interprétation. Intuitivement, *et* semble vouloir établir une cohérence entre les informations apportées par les segments qui précèdent son occurrence, et celles apportées par ce qui la suit. Par ailleurs, *donc* semble renforcer le caractère résomptif de l'expression *durant toutes ces années* en favorisant la lecture de l'expression comme anaphore associative, et indiquer que l'on touche dans la dernière clause à un objectif discursif considéré comme un point de pertinence saillant dans le discours.

Il semble donc que la présence de ces deux connecteurs tende à structurer l'extrait en regroupant les deux premières clauses et en les mettant en parallèle avec la troisième (structuration [1-2]-3).

4.3.7. Bilan des informations à tirer des niveaux textuels/discursifs

Cette première grille amène à se demander comment l'extrait choisi peut être structuré, et s'il peut l'être. En schématisant, on peut dire que :

- aux niveaux lexical, énonciatif, actantiel, référentiel, et au niveau de l'emploi des marqueurs discursifs, un certain nombre d'observables construits invitent à considérer que l'extrait se structure en deux moments

successifs : le premier regroupe les deux premières clauses isolées par la syntaxe, le second est constitué par la dernière clause ;

- d'autres observables aux niveaux énonciatif et référentiel permettraient d'argumenter en faveur d'une autre structure, regroupant les clauses 2 et 3, globalement mises en parallèle avec la première. Ces indices sont cependant moins nombreux que les premiers, et concernent moins de niveaux différents.

On peut donc supposer que c'est plutôt la première structuration qui l'emporte.

4.3.8. Comparaison avec le résultat de l'application de modèles de discours existants

Une application manuelle du modèle de la RST et du modèle genevois de l'organisation discursive donne des résultats similaires.

Si l'on applique le modèle de Mann et Thompson (1988), il est en effet possible de relier les deux premières clauses par une relation de type « SEQUENCE », qui les place à un même niveau de hiérarchie. L'ensemble constitue alors un satellite destiné à introduire la troisième clause qui constitue alors le noyau, selon une relation d'arrière-plan (« BACKGROUND »).

De façon similaire, la structuration en actes et interventions que l'on peut tirer du modèle genevois fait correspondre aux deux premières clauses des actes de même niveau se combinant en une même intervention, qui entrent dans une relation de préparation vis-à-vis du troisième formant à lui seul une deuxième intervention.

Il semble donc que ces deux modèles n'infirmant pas la structure que l'on peut établir à partir des combinaisons d'indices de différents niveaux relevés dans l'extrait.

En l'état actuel des grilles, ces lignes concernant les modèles de discours existants ont une double utilité :

- elles sont destinées, d'une part, à examiner dans quelle mesure les résultats obtenus par l'application de ces modèles sont concordants ou discordants par rapport aux informations que l'on peut obtenir en mettant en relation des observables de niveaux linguistiques différents ;
- elles invitent, d'autre part, à s'interroger sur la façon dont de tels modèles permettent de structurer le discours, en partant de l'hypothèse que les unités discursives ne sont pas *a priori* données, mais qu'elles sont à l'inverse construites par des faisceaux d'indices plus ou moins convergents entre eux. Un des principaux problèmes posé par ces modèles est en effet de savoir dans quelle mesure les phénomènes prosodiques interviennent implicitement dans l'attribution d'une structure à une séquence textuelle.

La structure hiérarchique que donnent AUCHLIN et SIMON (communication personnelle) de ce même extrait correspond à la structuration de type [[1-2]-3] détaillée plus haut, dans la mesure où l'extrait est scindé en deux moments correspondant d'une part au segment *on se souvient [...] sont sortis depuis*, et d'autre part au segment *et vous n'avez cessé [...] défendre leur cause*. Il est intéressant de noter que AUCHLIN et SIMON distinguent six actes dans ce passage, les segments 4 et 6 ci-dessous correspondant à un seul acte :

- 1 on se souvient entre autres de Ainsi soit-elle
- 2 écrit il y aura bientôt vingt ans (considéré comme incise)
- 3 bien d'autres livres sont sortis depuis
- 4 et vous n'avez cessé
- 5 durant toutes ces années donc (considéré comme incise)
- 6 de vous pencher sur le sort des femmes
- 7 et de continuer à défendre leur cause

Les actes dégagés correspondent, dans leur majorité, aux unités intonatives que l'on peut distinguer sur le plan prosodique, à la différence près que la première UI (*on se souvient entre autres de Ainsi soit-elle/ écrit il y aura bientôt vingt ans*) correspond à deux actes pour AUCHLIN et SIMON (sans doute en raison de l'emphase sur *écrit* ?). Les sept actes sont donc associés à des unités textuelles plus « petites » que les clauses isolées par la syntaxe, qui ont constitué notre point de départ pour l'analyse.

Cette confrontation entre différents résultats de l'application d'un même modèle pose des questions intéressantes. Elle permet d'abord de mettre en évidence le fait, pour un discours donné, qu'il n'existe pas une seule structure, mais que plusieurs structurations sont également possibles. Elle soulève ensuite le problème des procédures utilisées pour ces structurations ou, en d'autres termes, des différences entre des stratégies de prises d'indices considérés comme pertinents pour l'établissement de ces structures. Enfin, elle interroge la façon même dont ces observables sont construits, et les raisons pour lesquelles ils sont considérés comme pertinents et/ou saillants.

Nous avons pu ici comparer deux applications du modèle de ROULET *et al.* Cette comparaison n'a pas été possible en ce qui concerne le modèle de la RST de MANN et THOMPSON (1988), mais nous n'avons pas de raisons, *a priori*, de supposer qu'elle n'aurait pas fait émerger les mêmes questions.

4.4. Bilan des interactions prosodie/texte

Sur le plan des observables prosodiques, la similarité des cadences pour les clauses 1 et 2 d'une part, et la similarité des phénomènes d'abaissement pour ces mêmes clauses d'autre part, invitent à considérer que la locutrice (ici : la journaliste Roselyne Fayard) les met sur le même plan dans son discours. L'ensemble prosodique formé par ces échos entre les deux premières clauses entre donc en contraste avec la troisième. La mise en correspondance de ces observations avec les informations issues de l'analyse textuelle permet de constater que UI1 et UI2 correspondent à deux clauses (lignes RST) ou à deux actes textuels (lignes ROULET) en relation de séquence (RST) ou de succession (ROULET) l'une avec l'autre. À l'écho prosodique correspond donc l'identité de statut (2 clauses ou 2 actes) et l'égalité hiérarchique (les relations de séquence ou de succession sont établies entre des constituants textuels de même niveau hiérarchique). L'ensemble ainsi solidaire constitue donc de façon très consistante un nouveau constituant textuel de statut différent (intervention dans le modèle de Genève) qui pourra être mis en relation de dépendance hiérarchique avec le reste de l'énoncé.

Les phénomènes prosodiques permettent donc de distinguer clairement les deux moments discursifs qu'un bon nombre d'indices textuels dégagent.

En conséquence, la dernière ligne de la grille distingue ces deux moments qui semblent se dégager dans l'extrait, en qualifiant le premier de « historique des publications de Benoîte Groult », et le second de « explicitation des intentions / motivations de l'auteur ». Cette grille montre donc comment des indices prosodiques et textuels de niveaux différents peuvent concourir à structurer le discours.

5. CONCLUSION

Nous avons exposé dans cet article les fondements du programme scientifique du groupe interdisciplinaire PRODIGE dont le thème de recherche se rapporte à l'étude du rôle de la prosodie dans l'élaboration et l'interprétation du discours. Selon notre perspective, la première étape de ce programme doit s'appliquer à l'élaboration d'un outil d'analyse que nous concevons sous la forme d'une grille multilinéaire permettant de mettre en correspondance, pour un échantillon de parole donné (entretien, conversation, etc.), une notation des indices prosodiques et celle d'un ensemble d'éléments textuo-discursifs.

Il nous paraît important de préciser que cette grille, dont nous explicitons la composition au fil de cet article, ne constitue ni un modèle, ni un outil qui aurait trouvé sa forme définitive. Nous la considérons en effet comme un instrument heuristique en cours d'élaboration. De ce fait, le présent exposé ne rend compte que

de l'état présent de nos recherches et réflexions au sujet d'un outil destiné à être appliqué au plus grand nombre de données afin d'être enrichi et amélioré.

Il convient également de préciser que l'ordre vertical de présentation de la grille globale (représentations prosodiques et notations textuo-discursives) ne préjuge aucunement d'une hiérarchie particulière des niveaux d'analyse. En outre, la synthèse réalisée pour construire une interprétation ne se fait pas en une seule étape, ultime, de mise en correspondance. Elle se constitue en effet, selon nous, à différents niveaux d'abstraction, à différentes profondeurs du réseau complexe d'informations qui dégage le message linguistique de son encodage multidimensionnel dans la substance sonore.

L'un des intérêts principaux de la grille est qu'elle fait émerger des interactions d'indices inédites ou passées inaperçues. Par exemple, les deux extraits que nous avons illustrés (grille 1 ci-après, grille 2 sur le CD-ROM) présentent une mise en forme prosodique comparable de l'accent emphatique

- par des traits connus et bien décrits dans la littérature : l'association d'un ton mélodique haut et d'un allongement important (XLe) ;
- par la conjonction inédite de l'accent emphatique avec une frontière droite de domaine en contraste prosodique (de tempo dans la grille 1, de niveau de registre et de gamme tonale dans la grille 2).

Si de telles congruences s'avèrent généralisables à l'issue de l'examen de nombreux corpus, elles peuvent faire l'objet de descriptions systématiques exploitables, par exemple, en synthèse de la parole.

Comme nous l'avons longuement commenté pour l'extrait illustré par la grille 1, notre outil permet également de voir émerger de très nombreux indices, tant prosodiques que textuels, qui convergent pour rendre manifeste l'une des structurations discursives possibles d'un fragment de parole. À ce titre, la grille multilinéaire peut aussi être envisagée comme un outil d'évaluation intéressant de différents modèles discursifs, comme par exemple le modèle de Genève ou la RST (cfr les analyses discordantes de notre équipe et de celle de Genève sur le même passage).

L'outil d'analyse que nous mettons en œuvre et les résultats qu'il permet d'obtenir nous paraissent utilisables en vue d'une application à la reconnaissance et à l'interprétation automatique (application à l'analyse automatique du sens, par exemple), notamment en prenant en compte les incidences engendrées par l'usage conjugué des différents paramètres dans la segmentation automatique des unités discursives.

Dans l'optique de la synthèse de la parole, nous nous intéressons à la possibilité de configurer les paramètres orthogonaux des systèmes prosodiques, afin d'améliorer notablement la qualité de la synthèse. Nous envisageons également la possibilité de formuler des contraintes relatives à la sélection des configurations prosodiques les

plus appropriées à certaines configurations textuelles (lexico-syntaxiques), dans le paradigme notamment d'une implémentation en termes de balises (MERTENS *et al.* 2002). Enfin, il est permis d'envisager à plus long terme le recours à un modèle de synthèse par concept dans lequel les fonctions discursives que nous avons réussi à isoler et à analyser constitueraient les entrées du système qui sélectionnerait alors les configurations textuelles et prosodiques les plus appropriées.

Il nous apparaît au terme de notre travail que la multiplicité des informations auxquelles nous avons accès dans le cadre de notre approche plurilinéaire est compatible avec des techniques de codage et de représentation du type XML, dans le but de mettre ensuite en œuvre des méthodes de traitement de données hautement performantes (analyses de régressions et analyses en composantes principales) et de pouvoir « interroger » les corpus analysés au moyen de logiciels de requête.

6. ILLUSTRATION

Pour une meilleure lisibilité de ces figures, nous invitons le lecteur à consulter le CD-ROM.

Albert DI CRISTO, Cyril AURAN, Roxane BERTRAND,
Catherine CHANET, Cristel PORTES, Alain RÉGNIER
Laboratoire Parole et Langage,
UMR 6057, CNRS,
29 avenue Robert Schuman,
F-13621 Aix-en-Provence, France.

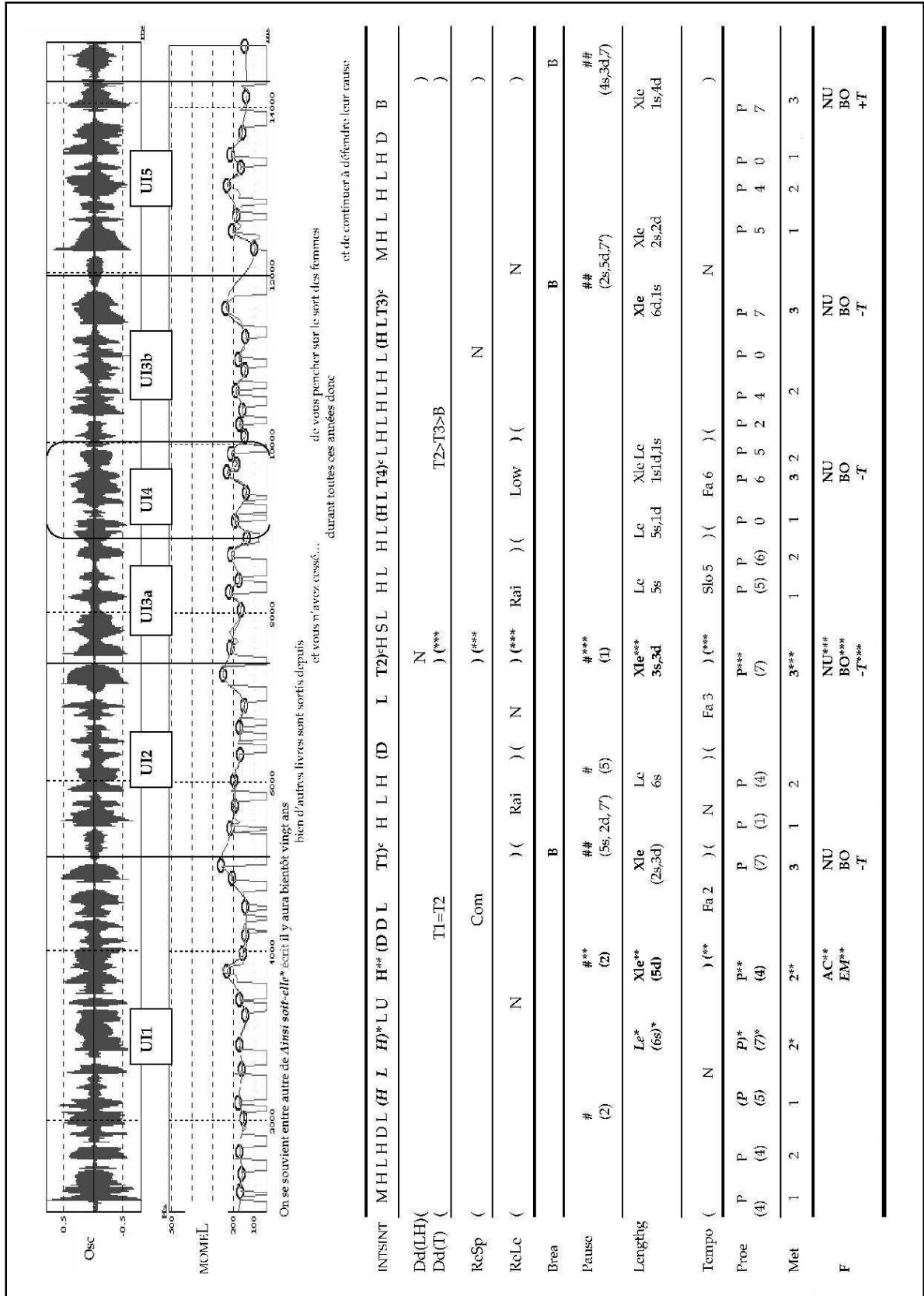


Figure 3a. Grille multilinéaire de l'extrait 1

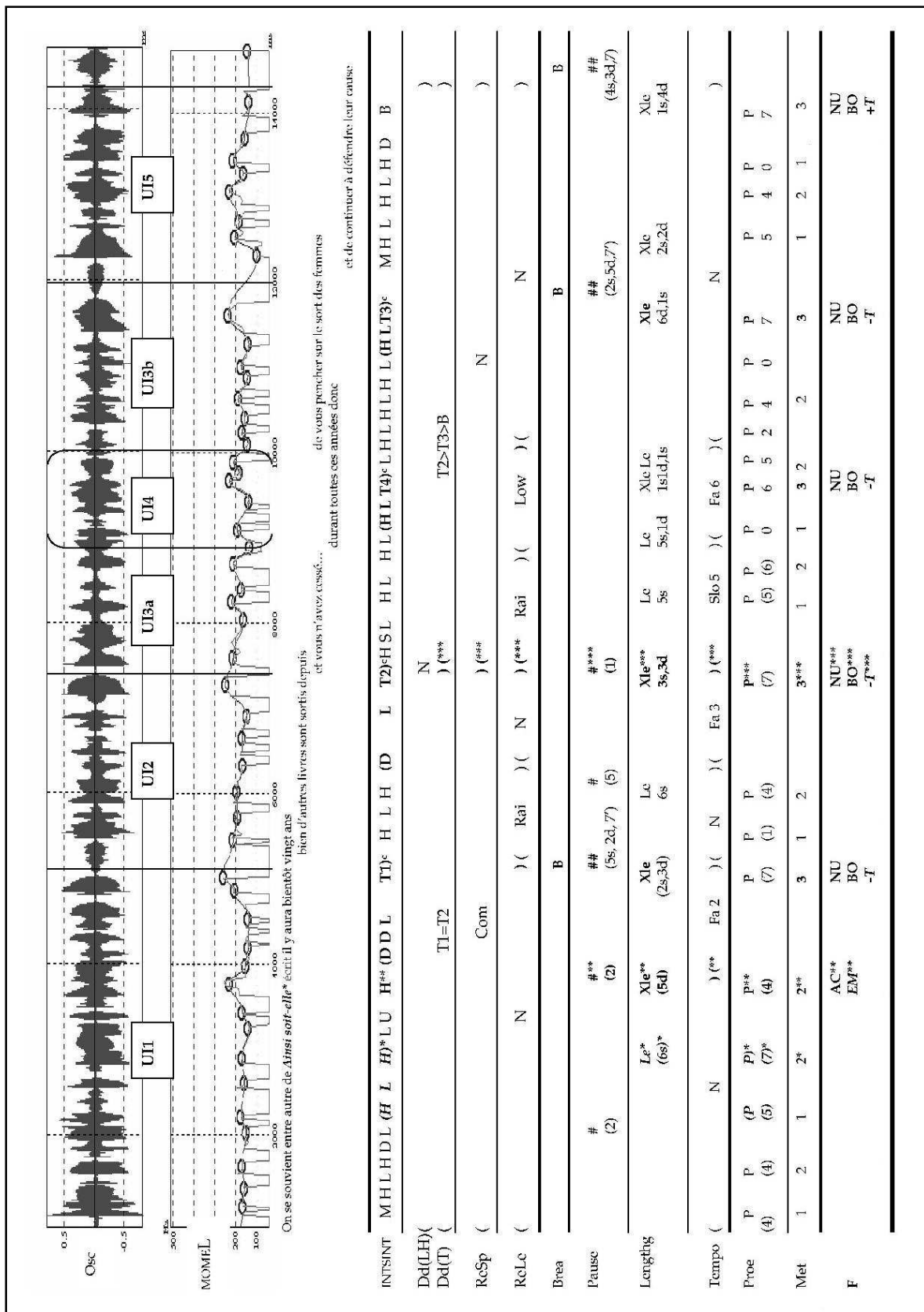


Figure 3b. Grille multilinéaire de l'extrait 1 (suite)

REFERENCES BIBLIOGRAPHIQUES

- ABEL, G. 1994. « Indeterminacy and Interpretation », *Synthèse* 37, 403-419.
- AKMAN, V. 2000. « Rethinking context as a social construct », *Journal of Pragmatics* 32, 743-759.
- ALESSANDRO, C. D' & P. MERTENS 1995. « Automatic pitch contour stylisation using a model of tonal perception », *Computers Speech and Language*, 9/3, 257-288.
- ARIEL, M. 1998. « Discourse markers and form-function correlations », in A.H. JUCKER & Y. ZIV (eds). *Discourse markers : Descriptions and Theory*, Amsterdam : John Benjamins, 223-259.
- ASHBY, W.R. 1960. *Design for a brain*, New York : John Wiley.
- ASHER, N. 1995. « From discourse macro-structure to micro-structure and back again: Discourse semantics and the focus/background distinction », *Proceedings of the Conference on Semantics in Context* (Prague).
- ASHER, N. 1999. « Discourse structure and the logic of conversation », in K. TURNER (ed.), *The Semantics/Pragmatics Interface*, vol. 1, Oxford : Elsevier, 20-48.
- AUCLIN, A. 1990. « Analyse du discours et bonheur conversationnel », *Cahiers de linguistique française* 11, 311-328.
- AUCLIN, A. 1999. « Les dimensions de l'analyse pragmatique du discours dans une approche expérientielle et systémique de la compétence discursive », in J. VERSCHUEREN (éd.), *Pragmatics in 1998*, Selected Papers from the 6th International Pragmatics Conference, vol. 2, 1-22.
- AUER, P. & A. DI LUZIO. 1992. *The Contextualization of Language*, Amsterdam : John Benjamins.
- AUER, P., E. COUPER-KUHLEN & F. MÜLLER. 1999. *Language in Time : The Rhythm and Tempo of Spoken Interaction*, Oxford : Oxford University Press.
- AUTHIER-REVUZ, J. 1995. *Ces mots qui ne vont pas de soi. Boucles réflexives et non-coïncidences du dire*, Collection Sciences du Langage, Paris : Larousse.
- AVESANI, C. & M. VAYRA. 1988. « Discorso, segmenti di discorso e un'ipotesi sull'intonazione », *Quaderni del Laboratorio di Linguistica* 2, 8-58.
- BANGE, P. 1992. *Analyse conversationnelle et théorie de l'action*, Paris : Hatier.

- BATES, E. & B. MAC WHINNEY. 1989. « Functionalism and the competition model », in B. MAC WHINNEY & E. BATES (eds), *The Crosslinguistic study of sentence processing*, Cambridge : Cambridge University Press.
- BATESON, G. 1972. *Steps to an Ecology of Mind*, New York : Ballantine Books.
- BERRENDONNER, A. 1993. « Périodes », in H. PARRET (éd.), *Temps et Discours*, Louvain : Presses Universitaires de Louvain, 47-61.
- BERRENDONNER, A. 1997. « L'auto-représentation du discours comme espace », in K. BOGACKI & T. GIERMAK-ZIELINSKA (eds), *Espace et temps dans les langues romanes et slaves*, Varsovie : Publications de l'Institut de Philologie Romane, 111-123.
- BERTRAND, R. 1999. « De l'hétérogénéité de la parole : Analyse énonciative de phénomènes prosodiques et kinésiques dans l'interaction interindividuelle ». *Thèse de doctorat*. Université de Provence.
- BERTRAND, R. (à paraître). « Être soi avec les mots d'autrui », in J. Merle (ed.), *Bibliothèque des Faits de Langues*, Paris : Ophrys.
- BLACHE, P. & A. DI CRISTO. 2002. « Variabilité et dépendance des composants linguistiques », *TALN 2002*, Nancy, 24-27 juin 2002.
- BLAKEMORE, D. 1987. *Semantic Constraints on Relevance*. Oxford : Blackwell.
- BOCK, K. 1995. « Sentence production: from mind to mouth », in J.L. MILLER & P.D. EIMAS (eds), *Speech, Language and Communication*. New York : Academic Press, 181-216.
- BRANDON, R. 1994. *Making It Explicit*. Cambridge, M.A. : Harvard University Press.
- BRUNER, J. 1990. *Acts of Meaning*. Cambridge, MA. : Harvard University Press.
- BUSQUETS, J., L. VIEU & N. ASHER. 2001. « La SDRT : une approche de la cohérence du discours dans la tradition de la sémantique dynamique », *Verbum*, XXIII (1), 73-102.
- CAMPBELL, N. 1992. « Multi-Level Timing in Speech ». Ph.D. Thesis. University of Sussex.
- CAMPIONE, E. 2001. « Étiquetage semi-automatique de la prosodie dans les corpus oraux : algorithmes et méthodologie ». *Thèse de doctorat*. Université de Provence.
- CANDEA, M. 2000. « Contribution à l'étude des pauses silencieuses et des phénomènes dits 'd'hésitation' en français oral spontané ». *Thèse de Doctorat*. Université de Paris III.

- CARNAP, R. 1956. *Meaning and Necessity*, Chicago : Chicago University Press.
- CAVÉ, C., I. GUAÏTELLA & S. SANTI. 2001. *Oralité et gestualité : Interactions et comportements multimodaux dans la communication*, Paris : L'Harmattan.
- CHAFE, W. 1976. « Givenness, contrastiveness, definiteness, subjects, topics and points of view », in C.N. LI (ed.), *Subject and Topic*, New York : Academic press, 25-57.
- CHAFE, W. 1987. « Cognitive constraints on information flow », in R.S. TIMLIN (ed.), *Coherence and Grounding in Discourse*, Amsterdam : John Benjamins, 21-51.
- CHAFE, W. 1993. « Prosodic and functional units of language », in J.A. EDWARDS & M.D. LAMPERT (eds), *Talking Data : Transcription and Coding Methods for Language Research*, Hillsdale, N.J. : Lawrence Erlbaum, 33-43.
- CHAFE, W. 1994. *Discourse, Consciousness and Time : The Flow and Displacement of Conscious Experience in Speaking and Writing*, Chicago : University of Chicago Press.
- CLARK, H.H. & B. BLY 1995. « Pragmatics and Discourse », in J.L. MILLER & P.D. EIMAS (eds), *Speech, Language and Communication*, New York : Academic Press, 371-410.
- CLEMENTS, G.N. 1990. « The status of register in intonation theory: comments on the papers by Ladd and by Inkelas and Leben », in J. KINGSTON & L. BECKMAN (eds), *Papers in Laboratory Phonology*. Cambridge : Cambridge University Press, 58-71.
- CONNELL, B. & D.R. LADD 1990. « Aspects of pitch realisation in Yoruba », *Phonology* 7, 1-30.
- COSNIER, J. 1998. *Le retour de Psyché*, Paris : Desclée de Brouwer.
- COUPER-KUHLEN, E. & M. SELTING 1996. *Prosody in Conversation : Interactional Studies*, Cambridge : Cambridge University Press.
- CRYSTAL, D. 1971. « Relative and absolute in intonation analysis », *Journal of the International Phonetic Association* 1/1, 17-28.
- CRYSTAL, D. 1980. *A Dictionary of Linguistics and Phonetics*, Oxford : Blackwell.
- DAHLGREEN, K. (1996). « Discourse coherence and segmentation », in H.E. HOVY & D.R. SCOTT (eds), *Computational and Conversational Discourse*, Berlin : Springer, 111-138.
- DAVIDSON, D. 1984. *Inquiries into Truth and Interpretation*. Oxford : Clarendon Press.

- DI CRISTO, A. 1978. « De la microprosodie à l'intonosyntaxe ». *Thèse de Doctorat d'État*, Université de Provence.
- DI CRISTO, A. & D.J. HIRST. 1986. « Modelling French Micromelody », *Phonetica*, 43/1, 11-30
- DI CRISTO, A. 1999. « Le cadre accentuel du français : essai de modélisation », 1^{ère} partie, *Langues* 2/3, 184-205 ; 2^e partie, *Langues* 2/4, 258-269.
- DI CRISTO, A. 2000a. « Une Grammaire Écologique comme cadre interprétatif de la prosodie de la parole », Communication présentée au 3rd *All-Nordic and Baltic Congress on Semiotics*, June 12-21, 2000 Imatra, Finland.
- DI CRISTO, A. 2000b. « Interpréter la prosodie », *Actes des XXIII^{èmes} Journées d'Études sur la Parole*, Aussois, juin 2000, 13-29.
- DI CRISTO, A. 2001. « La problématique de la prosodie dans l'étude de la parole dite 'spontanée' », *Parole*, 15/16, 189-249.
- DI CRISTO, A. 2002. « Modélisation et codage de la constituance prosodique », *Verbum* XXIV/1-2, 37-53.
- DIK, S. 1989. *The Theory of Functional Grammar*, Dordrecht : Foris.
- DIK, S. 1997. *The Theory of Functional Grammar*, The Hague : Mouton de Gruyter, 2 vol.
- DOMINEY, P.F., T. LELEKOV, J. VENTRE-DOMINEY & M. JEANNEROD. 1998. « Dissociable processes for learning the surface structure and abstract structure of sensorimotor sequences », *Journal of Cognitive Neurosciences* 769, 363-373.
- DUCROT, O. 1984. *Le dire et le dit*, Paris : Éditions de Minuit.
- DOUAY, C. 2000. *Éléments pour une théorie de l'interlocution*, Rennes : Presses Universitaires de Rennes.
- DUEZ, D. 1987. « Contribution à l'étude de la structuration temporelle de la parole en français ». *Thèse de Doctorat d'État*. Université de Provence.
- DUEZ, D. 2001. « Signification des hésitations dans la production et la perception de la parole spontanée », *Parole* 17/19, 113-137.
- DURANTI, A. & C. GOODWIN. 1992. *Rethinking Context : Language as an Interactive Phenomenon*. Cambridge : Cambridge University Press.
- EDWARDS, D. 1997. *Discourse and Cognition*, London : SAGE Publications.
- FILLIETTAZ, L. 2002. *La parole en action : Éléments de pragmatique psychosociale*. Québec : Éditions Nota bene.
- FRASER, B. 1999. « What are discourse markers ? », *Journal of Pragmatics* 31, 931-952.

- FORD, C.E. & S.A. THOMPSON. 1996. « Interactional units in conversation : syntactic, intonational, and pragmatic resources for the management of turns », in E. OCHS, E.A. SCHEGLOFF & S.A. THOMPSON, *Interaction and Grammar*, Cambridge : Cambridge University Press, 135-184.
- GIBBON, D. 1984. « Intonation as an adaptative process », in D. GIBBON & H. RICHTER, *Intonation, Accent and Rhythm*, Berlin : Walter de Gruyter, 165-192.
- GRABE, E. & B. POST. 2002. *The transcribed IViE corpus*, University of Oxford Phonetics Laboratory.
- GREIMAS, A.J. & J. COURTÈS. 1979. *Sémiotique : Dictionnaire raisonné de la théorie du langage*, Paris : Hachette Université.
- GRIZE, J.B. 1996. *Logique naturelle et communication*, Paris : Presses Universitaires de France.
- GROBET, A. 1997. « La ponctuation prosodique dans les dimensions périodique et informationnelle du discours », *Cahiers de Linguistique Française* 19, 83-123.
- GROSZ, B.J. & C.L. SIDNER. 1986. « Attention, intentions, and the structure of discourse », *Computational Linguistics* 12, 175-204.
- GÜLICH, E. & T. KOTSCHI. 1995. « Discourse production in oral communication: a study based on French », in U.M. QUASTHOFF (ed.). *Aspects of Oral Communication*, Berlin, New York : Walter de Gruyter, 30-66.
- GUMPERZ, J.J. 1992. « Contextualization and understanding », in A. DURANTI & C. GOODWIN (eds), *Rethinking Context. Language as an Interactive Phenomenon*, Cambridge : Cambridge University Press, 229-252.
- GUMPERZ, J.J. 1997. « Communicative competence », in N. COUPLAND & A. JAWORSK (eds), *Sociolinguistics : A Reader and Coursebook*, Houndmills, Hampshire : Macmillan, 39-48.
- GUSSENHOVEN, C. 2002. « Intonation and interpretation: Phonetics and Phonology », in B. BEL & I. MARLIEN (eds.), *Proceedings of the Speech Prosody 2002 conference*, 11-13 avril 2002, Aix-en-Provence : Laboratoire Parole et Langage, 727-730, 47-57.
- HART, J. & R. COLLIER. 1975. « Integrating different levels of intonation analysis », *Journal of Phonetics* 3, 235-255.
- HART, J., R. COLLIER & A. COHEN. 1991. *A Perceptual Study of Intonation : An Experimental-Phonetic Approach to Speech Melody*, Cambridge : Cambridge University Press.

- HERMES, D.J. & J. VAN GESTEL. 1991. « The frequency scale of speech intonation », *Journal of the Acoustical Society of America* 96, 83-92.
- HIRSCHBERG, J. & J. PIERREHUMBERT. 1986. « The intonational structuring of discourse », *Proceedings of the Twenty-fourth Annual Meeting*, New York : Association for Computational Linguistics.
- HIRSCHBERG, J. & B. GROSZ. 1992. « Intonational features of local and global discourse structure », *Proceedings of the Speech and Natural Language Workshop*, Harriman NY : DARPA, 441-446.
- HIRST, D.J. 1977. *Intonative Features*, The Hague : Mouton.
- HIRST, D.J. & A. DI CRISTO. 1998. *Intonation Systems : A Survey of Twenty Languages*, Cambridge : Cambridge University Press.
- HIRST, D.J., A. DI CRISTO & R. ESPESSER. 2000. « Levels of representation and levels of analysis for the description of intonation systems », in M. HORNE (ed.), *Prosody : Theory and Experiment*, Dordrecht : Kluwer Academic Publishers, 51-88.
- JEANNERET, T. 1999. *La coénonciation en français*, Bern : Peter Lang.
- KAMP, H. 1981. « Événements, représentations discursives et référence temporelle », *Langages* 64, 34-64.
- KAMP, H. & U. REYLE. 1993. *From Discourse to Logic*, Dordrecht : Kluwer.
- KERBRAT-ORECCHIONI, C. 1980. *L'énonciation : de la subjectivité dans le langage*, Paris : Armand Colin.
- KERBRAT-ORECCHIONI, C. 1990-1992. *Les interactions verbales*, 2 tomes, Paris : Armand Colin.
- KOOPMANS-VAN BEINUM, F.J. & M. VAN DONZEL 1996. « Discourse structure and its influence on local speech rate », *Proceedings of the Institute of Phonetic Sciences* 20, 1-11.
- LADD, D.R. 1990. « Metrical representation of pitch register », in J. KINGSTON & L. BECKMAN (eds), *Papers in Laboratory Phonology*, Cambridge : Cambridge University Press, 35-57.
- LADD, D.R. 1994. « Constraints on the gradient variability of pitch range or pitch level », in P. KEATING (ed.), *Phonological Structure and Phonetic Forms. Papers in Laboratory Phonology III*, Cambridge : Cambridge University Press, 43-63.
- LADD, D.R. 1996. *Intonational Phonology*, Cambridge : Cambridge University Press.

- LAMBRECHT, K. 1994. *Information Structure and Sentence Form: Topic, Focus and the Mental Representation of Discourse Referents*, Cambridge : Cambridge University Press.
- LASCARIDES, A. & N. ASHER. 1993. « Temporal interpretation, discourse relations and commonsense entailment », *Linguistic and Philosophy* 19, 143-186.
- LAVE, J. 1988. *Cognition in Practice*, Cambridge : Cambridge University Press.
- LEHISTE, I. 1976. « Influence of fundamental frequency pattern on the perception of duration », *Journal of Phonetics* 4, 113-117.
- LEVELT, W.J.M. 1989. *Speaking: From Intention to Articulation*, Cambridge, MA : MIT Press.
- LEVINSON, S.C. 1979. « Activity types and language », *Linguistics* 17/5-6, 365-399.
- LINDBLOM, B. 1990. « Explaining phonetic variation: a sketch of the H & H theory », in W. HARDCASTLE & A. MARCHAL (eds), *Speech production and speech modelling*, Dordrecht : Kluwer, 403-440.
- MANN, W.C. & S. A. THOMPSON. 1988. « Rhetorical Structure Theory: toward a functional theory of text organization » *Text* 8/3, 243-281.
- MANN, W.C. & S. A. THOMPSON. 2001. « Deux perspectives sur la théorie de la structure rhétorique (RST) », *Verbum* XXIII/1, 9-29.
- MCNEILL, D. 1992. *Hand and Mind: What Gestures Reveal about Thought*, Chicago : University of Chicago Press.
- MCNEILL, D. 2000. *Language and Gesture*, Cambridge : Cambridge University Press.
- MCNEILL, D., F. QUEK, K.E. MCCULLOUGH, S. DUNCAN, N. FURUYAMA, R. BRYLL, & R. ANSARI. 2001. « Catchments, prosody and discourse », *Gesture* 1/1, 9-32.
- MERTENS, P., A. AUCHLIN, J.-P. GOLDMAN, & A. GROBET. 2002. « L'intonation du discours : une implémentation par balises ; motifs et premiers résultats », in V. AUBERGÉ, A. LACHERET-DUJOUR & H. LÆVENBRUCK (éd.), *Journées Prosodie*, Actes du colloque de Grenoble, 10-11 octobre 2001, Grenoble : CNRS-GDR 34 Phonologie, Institut de la Communication Parlée, CREST Expressive Speech Project, 93-98.
- MORRIS, C. 1938. « Foundations of the Theory of Signs », in O. NEURATH, R. CARNAP & C. MORRIS (eds), *International Encyclopaedia of Unified Science*. Chicago : The University of Chicago Press, 136-147.

- NAKATANI, C. 1997. « Integrating prosodic and discourse modelling », in Y. SAISAKA, N. CAMPBELL & N. HIGUSHI (eds), *Computing Prosody*. Berlin : Springer, 67-80.
- NEWMeyer, F.J. 1998. *Language Form and Language Function*, Cambridge, MA : MIT Press.
- NICOLLE, S. 1997. « Conceptual and procedural encoding: Criteria for the identification of linguistically encoded procedural information », in M. GROEFSEMA (ed.), *Proceedings of the Hertford-shire relevance theory workshop*, Chelmsford : Peter Thoma, 45-56.
- NISHINUMA, Y. 1996. « How does the attenuation of vowel intensity affect the perception of its duration? », invited lecture by Tokyo Association of Spoken Languages, *NLRI*, Tokyo, 24 March 1996.
- NUYTS, J. 1994. « Functionalism vs. formalism », in J. VERSCHUEREN (ed.), *Handbook of Pragmatics*, Amsterdam : John Benjamins, 293-300.
- OCHS, E., E.A. SCHEGLOFF & S. THOMPSON 1996. *Interaction and Grammar*, Cambridge : Cambridge University Press.
- PASSONNEAU, R.J. & D.J. LITMAN. 1996. « Empirical analysis of three dimensions of spoken discourse: segmentation, coherence and linguistic devices », in E.H. HOVY & D.R. SCOTT (eds), *Computational and conversational discourse*, Berlin : Springer, 161-194.
- PATTERSON, D. 2000. « A Linguistic Approach to Pitch Range Modelling ». Ph.D. Thesis. University of Edinburgh.
- PEREGRIN, J. 1999. « The pragmatization of semantics », in K. TURNER (ed.), *The Semantics/Pragmatics Interface*, vol. 1, Oxford : Elsevier, 420-442.
- PIERREHUMBERT, J. & M. BECKMAN. 1988. *Japanese Tone Structure*, Cambridge, MA : MIT Press.
- PIERREHUMBERT, J. & J. HIRSCHBERG. 1990. « The meaning of intonational contours in the interpretation of discourse », in P.R. COHEN, J. MORGAN & M.E. POLLACK (eds), *Intentions in Communications*. Cambridge, MA : MIT Press, 270-311.
- PORTES, C. 2001. « Compte-rendu de E. Roulet et al. Un modèle et un instrument d'analyse de l'organisation du discours », *Marges Linguistiques 2*, 26-30, <http://www.marges-linguistiques.com>
- PRINCE, E. 1981. « Towards a taxonomy of given-new information », in P. COLE (ed.), *Radical Pragmatics*, New York : Academic Press, 223-255.

- PROVOST, J. 1996. « A Semantics of Contrast and Information Structure for Specifying Intonation in Spoken Language Generation ». Ph.D. Thesis. University of Pennsylvania.
- REBOUL, A. & J. MOESCHLER. 1998. *Pragmatique du discours*. Paris : Armand Colin.
- ROSEN, S.M. 1977. « The effect of fundamental frequency patterns on perceived duration », *Speech Transmission Lab. Quarterly Progress and Status Report* 1, 17-30.
- ROSSARI, C. 2001. *Connecteurs et relations de discours : des liens entre cognition et signification*. Nancy : Presses Universitaires de Nancy.
- ROSSI, M., A. DI CRISTO, D. HIRST, P. MARTIN & Y. NISHINUMA. 1981. *L'intonation : de l'acoustique à la sémantique*, Paris : Klincksieck.
- ROULET, E., A. AUCLIN, J. MOESCHLER, C. RUBATTEL & M. SCHELLING. 1985. *L'articulation du discours en français contemporain*, Bern : Peter Lang.
- ROULET, E. 1999. *La description et l'organisation du discours*, Paris : Didier.
- ROULET, E., L. FILLIETTAZ. & A. GROBET. 2001. *Un modèle et un instrument d'analyse de l'organisation du discours*. Bern : Peter Lang.
- SANTI, S., I. GUAÏTELLA, C. CAVÉ & G. KONOPCZYNSKI. 1998. *Oralité et gestualité : communication multimodale, interaction*, Paris : L'Harmattan.
- SCHEGLOFF, E.A. 1995. « Discourse as an interactional achievement : the relevance of action », *Research on Language and Social Interaction* 28, 201-233.
- SCHEGLOFF, E.A. 1996. « Turn organization », in E. OCHS, E.A. SCHEGLOFF & S.A. THOMPSON (eds), *Interaction and Grammar*, Cambridge : Cambridge University Press, 52-133.
- 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.
- SIMON, A.C. 2004. *La structuration prosodique du discours en français*. Bern : Peter Lang.
- SIMON, A.C. & A. GROBET. 2002. « Intégration ou autonomisation prosodique des connecteurs », in B. BEL & I. MARLIEN (eds.), *Proceedings of the Speech Prosody 2002 conference*, 11-13 avril 2002, Aix-en-Provence : Laboratoire Parole et Langage, 647-650.
- SILVERMAN, K., M.E. BECKMAN, J. PITRELLI, M. OSTENDORF, C. WIGHTMAN, P. PRICE, J. PIERREHUMBERT & J. HIRSCHBERG. 1992. « ToBI: a standard for

- labelling English prosody », *Proceedings of the Second International Conference on Spoken Language Processing 2*, 867-870.
- SPERBER, D. & D. WILSON. 1986. *Relevance: Communication and Cognition*, Cambridge : Harvard University Press.
- SWERTS, M. 1998. « Filled pauses as markers of discourse structure », *Journal of Pragmatics* 30, 485-496.
- TALMY, L. 2000. *Toward a Cognitive Semantics*, Cambridge, MA. : MIT Press.
- VALLDUVÍ, E. 1991. *The Informational Component*, New York : Garland.
- VALLDUVÍ, E. 1993. *Information packaging: A survey*, University of Edinburgh. Human Communication Research Centre, Research Paper HRCC/RP-44.
- VAN DONZEL, M.E. & F.J. KOOPMANS-VAN BEINUM. 1995. « Evaluation of discourse structure on the basis of written vs. spoken material », *Proceedings of the XIIIth International Congress of Phonetic Sciences*, vol. 3, 258-261.
- VERNANT, D. 1997. *Du discours à l'action*, Paris : PUF.
- VERSCHUEREN, J. 1994. « The pragmatic perspective », in J. VERSCHUEREN, J. O. ÖSTMAN & J. BLOMMAERT (eds), *Handbook of Pragmatics*, Amsterdam : John Benjamins, 1-19.
- VION, R. 1992. *La Communication verbale : Analyse des interactions*, Paris : Hachette.
- VION, R. 2001. « Modalités, modalisations et activités langagières », *Marges Linguistiques* 2, 209-231, <http://www.marges-linguistiques.com>.
- WHEELDON, L. & A. LAHIRI. 1997. « Prosodic units in speech production », *Journal of Memory and Language* 37, 356-381.
- WICHMANN, A. 2000. *Intonation in Text and Discourse*, London : Longman.
- WILSON, D. & D. SPERBER. 1993. « Linguistic form and relevance », *Lingua* 90, 1-25.
- WITTGENSTEIN, L. 1922. *Tractatus Logico-Philosophicus*, London : Routledge.
- ZELLNER, B. 1998. « Caractérisation et prédiction du débit de parole en français ». *Thèse de Doctorat*. Université de Lausanne.