

*in Perrin L. (éd.) (2006), Le sens et ses voix,
Recherches Linguistiques 28, Université P.
Verlaine, Metz, 77-104.*

POLYPHONIE ET PROSODIE : CONTRAINTES ET RENDEMENT DE L'APPROCHE MODULAIRE DU DISCOURS

Antoine AUCHLIN et Anne GROBET
Université de Genève

1. L'APPROCHE GENEVOISE DE LA POLYPHONIE (1985- 2001)

La polyphonie est une problématique majeure dans le modèle d'analyse du discours développé à Genève, dès les années 1980 (Roulet et al. 1985). Elle a suivi l'évolution du modèle, passant du statut de niveau d'organisation intégré à la structure hiérarchique du discours, à celui de dimension ou module autonome, Roulet (1991) estimant nécessaire, pour rendre compte de la complexité du discours, de développer un modèle d'analyse modulaire. Il est apparu que la dimension polyphonique ne pouvait pas être réduite à un simple « module », au sens de dimension caractérisée par son caractère élémentaire (Nølke 1994). En effet, l'étude de la polyphonie implique la prise en compte de nombreux paramètres hétérogènes : marques lexicales, syntaxiques, ponctuation, structure du discours, cadre interactionnel, etc. Dans les versions les plus récentes du modèle, la polyphonie est ainsi traitée à travers la combinaison de deux « formes d'organisation », de complexité croissante : l'organisation énonciative, chargée du repérage des séquences polyphoniques ainsi que de leur description formelle, et l'organisation polyphonique, qui rend compte de leurs fonctions (Roulet et al. 2001) ¹.

¹ Rappelons que dans ce modèle une « organisation » ou « forme d'organisation » résulte du couplage d'informations en provenance de modules (ici : syntaxique et interactionnel), ou d'organisations plus nucléaires (comme l'organisation énonciative, « en amont » de l'organisation polyphonique).

1.1. L'apport de Bakhtine et Ducrot

À l'origine des travaux genevois sur la polyphonie se trouvent les travaux de deux auteurs, à savoir Bakhtine (1977, 1978) et Ducrot (1984). Roulet reprend à Bakhtine l'idée d'une conception fondamentalement dialogique du discours : tout discours est nécessairement adressé et intégré dans une interaction, à laquelle il réagit. À cet auteur Roulet emprunte également le terme même de « polyphonie », ainsi que la description des diverses facettes qui en émergent :

- a) la diversité des voix apparaissant dans les énoncés successifs d'un discours, en particulier dans le discours romanesque ;
- b) la diversité des voix au sein d'un seul énoncé (que Bakhtine appelle *hybride*), d'une seule construction linguistique comme c'est le cas dans le style indirect libre notamment ;
- c) un cas particulier de cette pluralité de voix est constitué par « la reprise et l'intégration du discours de l'interlocuteur dans le discours du locuteur » (Roulet et al. 1985 : 70). En effet, le locuteur cherche à entrer en résonance avec le discours de son interlocuteur et commence par s'appuyer sur son point de vue avant de présenter le sien. Cette configuration doit être distinguée de la mention de la voix d'autrui, qui figure dans le discours simplement au titre d'objet de référence, et non de pôle d'interaction. Pour distinguer ces deux cas de figure, Roulet introduit le terme de *diaphonie* pour désigner la reprise de la parole du destinataire, qui permet d'enchaîner en négociant les points de vue, et celui de *polyphonie* (au sens restreint) pour mentionner la citation de la parole d'autrui, non destinataire de l'intervention (Roulet et al. 1985 : 70-71).

De la théorie polyphonique de l'énonciation (Ducrot 1984), Roulet reprend la description des énoncés ironiques, des constructions concessives en français avec *mais*, de la négation polémique, etc. (Roulet 2001 : 277). Mais alors que la conception de la polyphonie de Ducrot se situe au niveau des phrases et ne présuppose aucunement l'existence externe, contextuellement référencée, des points de vue (ou énonciateurs) qu'elle articule, celle de Roulet se situe au niveau du discours et prend en compte essentiellement des faits de discours rapporté censés exister indépendamment du discours produit qui les représente (Perrin à paraître).

Enfin, Roulet intègre les apports des approches de Bakhtine et de Ducrot dans le cadre de son modèle d'analyse du discours. Son objectif est de décrire de manière précise et systématique des fragments de discours suivis : le discours rapporté, repéré et décrit du point de vue formel (organisation énonciative), est, dans un second temps, mis en relation avec le modèle général d'analyse, afin d'être étudié d'un point de vue fonctionnel (organisation polyphonique).

1.2. L'organisation énonciative

L'analyse de l'organisation énonciative vise à distinguer, au sein des interventions, le discours *produit* du discours *représenté*. Roulet utilise à dessein le terme de « représenté » plutôt que celui de « rapporté » : le discours n'a pas nécessairement été tenu, il peut s'agir d'une invention (romanesque), ou d'une anticipation du discours d'autrui. Le discours produit se définit comme un discours *effectivement tenu* par un locuteur, qui se distingue du discours représenté qui n'est que *mentionné* dans son intervention (Perrin 1995). Par exemple, le discours produit peut être une réplique dans un dialogue, et le discours représenté les voix que le locuteur fait entendre. Ou alors le discours produit est un texte narratif romanesque, où les voix des personnages sont des discours représentés ; une lettre peut être considérée comme du discours produit, et les reprises des arguments antérieurs comme des extraits de discours représenté. Dans le cadre d'un roman, la différence entre discours *produit* et *représenté* doit être rattachée à la situation d'interaction particulière du roman, qui enchâsse (au moins) deux situations d'interaction l'une dans l'autre.

Illustrons ce repérage des séquences de discours représenté par le bref extrait d'un dialogue entre une cliente (C) et une vendeuse (V). La cliente raconte un dialogue avec son mari (M). Par convention, les séquences de discours représenté sont encadrées entre crochets, une majuscule indiquant la source juste avant. Les segments introducteurs de parole sont en italique ².

² L'analyse de cet exemple, plus complexe qu'il n'y paraît, fait intervenir de la diaphonie implicite, comme nous le montrerons ci-dessous.

- (1) C (...) alors *il me dit* M[hein::: la flèche] . *je dis* C[mais la flèche elle t'indique toujours le nord]
 V ben ouais
 C alors *il me dit* M[alors le nord il est où ?] . *je dis* C[ben tu te tournes jusqu'à ce que la flèche t'indique le nord]
 V (rire)
 C il a toujours pas compris

Les répliques des locuteurs C et V constituent du discours produit, puisqu'il ne s'agit pas d'un dialogue romanesque (où le discours des personnages est systématiquement emboîté dans celui du narrateur). Dans ces interventions, le discours représenté est clairement introduit par des marques linguistiques et son repérage ne soulève aucun problème, ce qui n'est pas toujours le cas. Il peut, en outre, y avoir différents degrés d'emboîtement de discours représenté dans du discours représenté.

- (2) *Ils disent* I[qu'on leur a dit O[que Jean prétendait J[qu'ils avaient mal travaillé]]]

Les séquences de discours représenté peuvent ensuite être décrites du point de vue de leur forme, en précisant s'il s'agit de discours représenté a) *formulé*, b) *désigné* ou c) *implicite*.

a) *Le discours représenté formulé*

Le discours peut être représenté comme une formulation explicite, et cela en style direct, indirect ou indirect libre. L'exemple 1 illustre du discours formulé en style direct : dans le discours oral les marques de ce style direct sont les verbes de parole (*il me dit, je dis...*), les pronoms (utilisation du pronom de la deuxième personne) et la prosodie, qui joue un rôle de délimitation à certains égards comparable à celui de la ponctuation (v. *infra*). Perrin (1995 : 218) rappelle que le style direct a pour effet de produire un effet de réel, mais ne garantit pas pour autant l'authenticité du discours représenté. L'exemple 2, construit, illustre du discours représenté en style indirect, identifiable par les introducteurs (*disent que, a dit que, prétendait qu'*), les temps verbaux (*imparfait* et *plus-que-parfait*) et le passage des pronoms de première personne à la troisième personne (*leur, ils*). Le discours représenté en style direct ou indirect libre se caractérise par l'absence de séquence métadiscursive introductrice :

- (3) *J'ai tout de suite entretenu Diaghilev de ce qui vous concerne.* (...) Diaghilev était en effet un peu inquiet. D[Qu'allait-il se passer avec

ces premiers spectacles, pour lesquels il était au « pourcentage » sans garantie ?]. ... (adaptation d'un exemple proposé et analysé par Perrin 1996 : 131)

Ce passage est issu d'une lettre écrite par Ansermet, qui rapporte à Stravinsky un échange qu'il a eu avec Diaghilev. Si l'échange est globalement rappelé par la phrase initiale (suivie par un commentaire qui ne figure pas ici), la position de Diaghilev est présentée sans introduction métadiscursive : il s'agit de discours indirect libre. Les guillemets autour de « pourcentage » peuvent être interprétés soit comme renvoyant à un îlot de discours direct, soit comme marquant l'absence de prise en charge de ce terme par le locuteur. Précisons que, comme souvent dans ce type de configuration n'impliquant pas de marque explicite, la séquence peut également être interprétée comme renvoyant à un état mental et non à un discours tenu.

b) *Le discours représenté désigné*

Le discours représenté formulé, illustré par les exemples 1, 2 et 3, est la configuration la plus souvent répertoriée. Mais dans certains cas, il n'y a pas *représentation* du discours évoqué, mais simplement *désignation* de ce discours par un verbe ou un syntagme nominal :

- (4) *Il a parlé* I[] pendant des heures.
- (5) *Les critiques* J[] de Jean m'ont énervée.

Comme le discours représenté désigné n'implique pas de reprise détaillée du contenu des paroles, on le transcrit par des crochets vides, précédés de l'initiale du locuteur, après l'expression désignant le discours cité.

Le discours représenté désigné n'est que rarement pris en compte dans la littérature, mais cette omission ne se justifie pas. Roulet (1997 : 157) souligne que le discours représenté désigné permet les mêmes enchaînements que le discours représenté formulé. De plus, il n'y a pas une frontière nette entre le discours représenté formulé au style indirect et le discours représenté désigné, mais plutôt une sorte de continuum (exemples empruntés à Perrin 1995 : 217) :

- (6) *Pierre a dit* : « *J'en ai par dessus la tête de ces histoires !* ».
- (7) *Pierre a dit qu'il en avait assez de tout ça, qu'il n'en pouvait plus.*
- (8) *Pierre a exprimé sa lassitude et son impatience.*

Alors que le discours représenté formulé au style direct implique une grande autonomie syntaxique (pronoms, temps verbaux) et périodique (présence des deux points, des guillemets, d'une majuscule et d'un point d'exclamation liés au discours cité), et qu'il se caractérise par un caractère précis et évocateur (Perrin 1995), le discours représenté au style indirect perd son autonomie (intégration syntaxique et périodique) et son caractère concret ; le discours désigné par des nominalisations présente quant à lui une intégration syntaxique et périodique maximale, et un degré d'abstraction encore plus élevé.

De plus, comme l'illustrent les exemples 7a et 8a, une même unité textuelle (acte) peut combiner plusieurs modes de représentation³ :

(7a) *Pierre a dit qu'il en avait « par dessus la tête, de ces histoires »*

(8a) *Pierre a exprimé sa lassitude et son impatience de « toutes ces histoires sans fin »*

Ces différents modes de représentation s'inscrivent dans des segments modiquement séparés les uns des autres : par les guillemets, à l'écrit⁴ ou par une rupture prosodique à l'oral (cf. infra) et un accent initial ouvrant un arc accentuel sur le segment représenté, et/ou par d'autres changements de paramètres de F0 (fréquence fondamentale, hauteur de la voix), indexant une unité morpho-syntaxique à une origine énonciative⁵. Cette origine énonciative est 1) distincte du locuteur prenant en charge le discours environnant ; 2) potentiellement repérable contextuellement ; et 3) identifiable aussi bien à un type (LE snob, LE prof, LE paysan...), qu'à un individu. Dans des exemples prototypiques tels que 6 ou 7, les frontières des discours « introducteur » et « représenté » convergent avec les contraintes des segments syntaxiques qui les portent, alors que dans 7a ou 8a les éléments « en mention » sont syntaxiquement intégrés.

c) Le discours représenté implicite

3 L'exemple 3 (« pourcentage ») implique vraisemblablement aussi une configuration similaire (cf. supra).

4 En face à face, un geste mimant des guillemets peut remplir un rôle analogue – mais semble moins fréquent que les guillemets écrits.

5 Le fonctionnement de la prosodie (cf. infra) permet également d'intégrer dans un ensemble de contraintes (p.ex. le tempo) un objet distingué par un autre ensemble de contraintes (gabarit tonal – registre, par exemple).

Une configuration encore moins souvent décrite, si ce n'est dans le cadre de la théorie polyphonique de l'énonciation (Ducrot 1984), est celle du discours représenté *implicité*, qui apparaît notamment lorsqu'un connecteur tel que *mais* ou *eh bien* intervient en tête de réplique dans un dialogue. En voici un exemple, extrait de Proust, étudié par Roulet (2001) :

- (9) [contexte : Albertine vient de dire qu'elle doit partir rendre visite à une dame]
 N[...N'« A [] *Mais* on peut bien manquer une visite »]...

Le N renvoie au narrateur du roman *Sodome et Gomorrhe* de Proust, d'où est tiré cet extrait : les crochets extérieurs encadrent donc en fait la totalité du roman. Le N' renvoie au personnage du narrateur, qui donne la réplique à Albertine (A). La réplique encadrée par les guillemets est ainsi prononcée par le narrateur-personnage, et peut être décrite comme du discours représenté formulé au style direct. Cette réplique débute sur un *Mais* qui enchaîne sur une reprise implicite du discours d'Albertine, que Roulet paraphrase ainsi : « certes, vous dites que c'est impossible » (Roulet 2001 : 284).

Dans cette optique, on peut reconsidérer l'analyse de l'exemple 1 en notant que les connecteurs placés en tête de réplique impliquent eux aussi du discours représenté implicite :

- (1b) C (...) alors *il me dit* M[hein:: la flèche] . *je dis* C[M] | **mais** la flèche elle t'indique toujours le nord]
 V ben ouais
 C alors *il me dit* M[C] | **alors** le nord il est où ?] . *je dis* C[M] | **ben** tu te tournes jusqu'à ce que la flèche t'indique le nord]
 V (rire)
 C il a toujours pas compris

Dans cet exemple, *mais* n'est pas la seule marque à enchaîner sur du discours implicite : le *alors*, et le *ben* (version orale du *eh bien*) remplissent un rôle similaire, faisant « rebondir » le discours tenu sur la réplique antérieure implicite.

Le repérage énonciatif qui vient d'être décrit se fonde le plus souvent sur des marques linguistiques (syntagme verbal ou nominal renvoyant à une action langagière, ponctuation, connecteur...). Or, dans certains cas comme dans les énoncés ironiques, le discours représenté n'est pas marqué linguistiquement, ainsi que le relève Roulet (2001 : 285). Nous y revenons plus bas, car tout porte à croire que c'est justement dans le repérage de ces cas-là que la prosodie

(même mentale dans le cas de la lecture d'un roman) joue un rôle décisif.

Les séquences de discours représenté isolées peuvent ensuite être décrites plus précisément à partir de leurs caractéristiques linguistiques. On distingue ainsi le discours représenté *effectif*, présenté comme ayant vraiment été tenu (même dans un monde fictionnel), du discours représenté *potentiel*, qui imagine ou anticipe un discours possible, par exemple une objection. Ainsi, l'exemple 1 implique un discours représenté effectif, car le dialogue entre les deux conjoints est présenté comme ayant effectivement eu lieu (peut-être même plusieurs fois), notamment par l'usage de *dire* au présent et par la conclusion tirée par C : *il a toujours pas compris*. Dans l'exemple 10, le locuteur J. Chancel imagine un dialogue entre lui et F. Sagan, son interlocutrice :

- (10) JC : Alors maintenant il y a : ce nouveau roman / Le Le lit
défait / qui paraît chez Flammarion / . (0.58s.) moi je dis euh : que
c'est sans doute l'un de vos meilleurs livres \ il faut jamais dire euh : il
faut jamais dire le MEilleur livre \ .(0.3s.) peut-être /
FS : oui
JC : si je vous le demandais à vous JC[...] Françoise / peut-
être me répondriez-vous FS [Dans . Dans un mois dans un an] \
(0.27s.) je ne sais pas quel est . (0.25s.) quel est celui que vous
considérez comme le meilleur \

Si je vous le demandais à vous est du discours représenté désigné, puisque la question n'est pas formulée (mais évoquée par « le ⁶ »), et potentiel, car à ce stade la question n'a pas encore été posée : elle est introduite comme fictive par un *si* (et ne le sera qu'ultérieurement : *je ne sais pas quel est celui que vous considérez comme le meilleur*). JC imagine ensuite la réponse que pourrait lui donner FS, en l'introduisant de manière fortement modalisée (utilisation du *peut-être* et du conditionnel), ce qui indique clairement qu'il s'agit d'une réponse potentielle.

Enfin, à partir des distinctions introduites en 1985 rappelées ci-dessus, on distingue le discours représenté *polyphonique* (mettant en scène la parole d'un tiers) et le discours représenté *diaphonique* (reprenant la parole de l'interlocuteur), auxquels s'ajoute le discours

6 Le discours désigné ne matérialise généralement pas son contenu, mais il existe malgré tout différents degrés d'explicitation de ce contenu (voir la différence entre « ses propos » et « ses louanges »). Cet exemple illustre un cas intermédiaire, puisque la pronominalisation reformule en quelque sorte le contenu.

représenté *autophonique* (reprise de sa propre parole). Dans l'exemple 1, le discours représenté du mari peut être décrit comme polyphonique, de même que les discours représentés de l'exemple 2. Lorsque le locuteur représente sa propre parole, mais tenue à un moment autre que dans l'interaction présente, comme c'est le cas de la cliente dans l'exemple 1, ou dans l'intervention de JC dans l'exemple 10, on parle de discours représenté *autophonique*. Enfin, il faut distinguer de ces deux cas de figure le discours *diaphonique*, qui implique le partenaire direct de l'interaction, et qui permet donc véritablement de « rejouer la négociation ⁷ ». Ainsi, le discours représenté implicite de l'exemple 9 peut être décrit comme *diaphonique*, car il s'agit du discours de l'interlocutrice directe du narrateur-personnage. De même, lorsque JC représente la réponse que FS pourrait potentiellement lui donner, il s'agit de discours *diaphonique*. Il y a différentes manières de représenter le discours de son interlocuteur : Torck (1994) distingue les discours représentés *diaphoniques locaux* et *à distance*. Alors que les discours représentés *à distance*, qui impliquent des propos lointains de l'interlocuteur, impliquent souvent un marquage linguistique explicite (introduction du type *vous dites que...*, guillemets et/ou italiques à l'écrit, etc.), les discours représentés *locaux* concernent du discours qui vient d'être tenu (voir le discours représenté implicite de l'exemple 9). Le marquage linguistique y est souvent moindre, et le rôle de la prosodie décisif (réponses-écho, Perrin, Deshaies et Paradis 2003).

1.3. L'organisation polyphonique

L'approche modulaire sépare l'analyse des modes de représentation de discours associés à des sources distinctes du locuteur, objet de l'organisation énonciative, de l'analyse de leurs fonctions – objet de l'« organisation polyphonique ». Ces fonctions peuvent être appréhendées à différents niveaux et donc faire intervenir différents plans d'organisation du discours. Il n'y a de ce fait pas une unique démarche d'analyse, mais une adaptation au type de discours concerné. En guise d'exemple, nous étudierons deux des fonctions que

⁷ Dans l'exemple 1, la cliente mentionne les propos de son mari qui était son interlocuteur dans l'interaction représentée, mais qui n'intervient pas dans l'interaction actuelle avec la vendeuse. Il s'agit donc de discours produit polyphonique et non diaphonique, la composante diaphonique étant interne au discours représenté.

peut remplir le discours représenté, les fonctions narrative et argumentative.

L'une des fonctions majeures du discours représenté est celle que l'on peut qualifier de narrative, notamment lorsque l'on se trouve en présence de dialogues représentés. Ce peut être le cas dans le roman, lors de la reproduction de paroles de personnages (Roulet 2001), ou encore dans le récit interactif inscrit dans des interactions orales (Vincent et Dubois 1997). Ce type de discours représenté peut être illustré par le premier dialogue de Bouvard (B) et Pécuchet (P) :

- (11) N[... Pour s'essuyer le front, ils retirèrent leurs coiffures, que chacun posa près de soi ; et le petit homme aperçut écrit dans le chapeau de son voisin : Bouvard ; pendant que celui-ci distinguait aisément dans la casquette du particulier en redingote le mot : Pécuchet.
 - « B/P[Tiens !] » dit-il « B/P[nous avons eu la même idée, celle d'inscrire notre nom dans nos couvre-chefs.] »
 - « P/B[Mon Dieu, oui ! on pourrait prendre le mien à mon bureau !] »
 - « B/P[C'est comme moi, je suis employé.] »
 Alors ils se considérèrent.] (Flaubert, *Bouvard et Pécuchet*, Folio, 52)

Ce dialogue s'inscrit dans le cadre d'un récit, qui se développe à travers une série d'actions évoquées par des verbes au passé simple (*retirèrent, posa, aperçut, dit, considérèrent*) et à l'imparfait (*distinguait*). C'est donc par les actions langagières, explicitées (*dit-il*) ou implicites (dans les deux répliques suivantes), que le dialogue représenté acquiert son caractère narratif.

Plus précisément, les répliques des personnages peuvent être analysées comme des séquences de discours représenté formulé au style direct, effectif (même si fictionnel) et polyphonique, dans la mesure où il s'agit de voix tierces représentées dans la parole du narrateur (N). Il subsiste toutefois un doute quant à la distribution des paroles entre les deux personnages. En effet, le référent du pronom *il* lié à l'introduction de la première réplique est ambigu, car les critères permettant de l'identifier sont contradictoires : si l'on considère qu'il s'agit du sujet de la proposition précédente, il s'agit de Bouvard. Mais si l'on interprète ce pronom en s'appuyant sur le référent dont l'évocation est la plus proche, il renvoie à Pécuchet. Le parallélisme appuyé mis en place entre les deux personnages en ce début de roman ne donne pas d'indications permettant de lever cette ambiguïté, et il y a au contraire tout lieu de penser que cette dernière contribue à renforcer l'effet d'osmose existant entre les deux personnages. Le contenu des répliques des personnages, lié à la découverte d'un point commun, renforce bien évidemment cette impression. Le discours

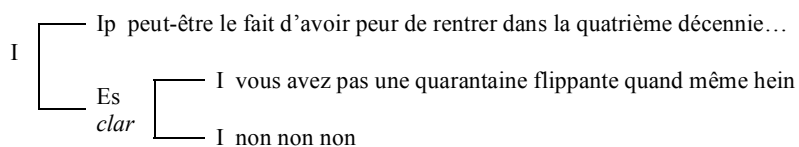
représenté s'inscrit donc dans une trame narrative à laquelle il contribue doublement, à savoir par la succession des actions langagières qu'il présente et par les points de vue des personnages qu'il met en scène.

Une autre fonction du discours représenté est celle que l'on peut qualifier d'argumentative (au sens large, regroupant l'argument d'autorité, la concession, etc.). Ici, le but n'est pas tant de reproduire des paroles qui auraient été tenues, que d'exploiter le discours représenté dans le but d'étayer une argumentation. Il en va ainsi dans cet extrait d'un entretien entre C. Defaye et G. Depardieu (qui sera commenté ultérieurement du point de vue de son organisation prosodique) :

- (12) GD (...) peut-être le fait d'avoir peur de rentrer dans la quatrième décennie heu - bon ça fait que - ça fait heu que j'suis p'tête moins patient - et en même temps j'commence à [!] l'être plus mais enfin
 CD vous avez pas GD[une quarantaine flippante] quand même hein
 GD non - non non - mais on a t| bon après des films comme *Trop belle pour toi* - (...)

Le passage qui nous intéresse s'inscrit dans l'intervention de CD, qui interrompt GD pour lui poser une question qui repose sur une diaphonie explicite : CD résume par « quarantaine flippante » les paroles de GD, qui avoue notamment « le fait d'avoir peur de rentrer dans la quatrième décennie ». Il s'agit de discours représenté formulé par une expression nominale axiologiquement marquée (le terme de « flippante » étant porteur d'une orientation plus péjorative que l'expression « d'avoir peur »).

La structure hiérarchique-relationnelle permet d'appréhender plus précisément encore la fonction argumentative du discours représenté. Sans entrer dans les détails, on constate ici que CD ouvre un échange subordonné (Es) à la suite de l'intervention de GD (Ip) :



Comme l'échange subordonné suit une intervention dont il pointe un élément, on considère qu'il porte la relation générique de clarification

(*clar*). Plus précisément, par la présence de la négation (*vous avez pas...*) et du connecteur *quand même*, la question de CD ne se présente pas comme une simple demande d'informations complémentaires, mais comme une rectification qui contredit au moins partiellement les propos de GD et qui, par son orientation argumentative marquée, n'appelle pas tant une réponse qu'une confirmation par dénégation (fournie par GD qui répond *non non non*).

La séquence diaphonique peut encore être interprétée au niveau de l'organisation stratégique et des relations de faces mises en jeu dans ce dialogue. Quelques indications sur le contexte de ce passage sont utiles à ce niveau : GD est interviewé par CD et se présente, au long de cet entretien, comme quelqu'un qui fait des erreurs, des « conneries » et qui a peur. Ce qui ne correspond pas exactement aux attentes de CD, qui attend de son invité notamment une « belle déclaration en conclusion » (peu avant l'extrait cité). Dans un tel contexte, la séquence diaphonique « quarantaine flippante », axiologiquement marquée, permet à CD d'exagérer les faiblesses de GD d'une manière inacceptable pour l'image que l'acteur veut donner de lui à son public, afin de l'inviter à nier ces faiblesses, et du même coup à restaurer l'image qu'il donne de lui au public.

2. POLYPHONIE ET PROSODIE : REPÉRAGE ET QUALIFICATION DU DISCOURS REPRÉSENTÉ

Contrairement à la problématique de la polyphonie, la prosodie n'a été intégrée que récemment dans le modèle genevois⁸ ; les contributions prosodiques aux organisations énonciative et polyphonique n'ont pas encore été systématiquement explorées⁹. Il semble pourtant exister une relation privilégiée entre ces deux formes d'organisation, quoique complexe et difficile à cerner.

D'après Couper-Kuhlen (1998), la prosodie guide l'interprétation du discours rapporté, et cela en particulier lorsque les autres indices

8 Voir Simon 2004 ; Perrin 1999 ; Auchlin & Ferrari 1994 ; *Cahiers de linguistique française* 23, 2001.

9 Voir cependant Simon 2004 : 276-285 ; Grobet à paraître a ; Simon & Auchlin 2001.

(p. ex. syntaxiques ou lexicaux) font défaut. Toutefois, ce lien n'a rien d'obligatoire, ni même de figé. Demers (1996) insiste sur la grande variabilité de ce marquage ; pour Klewitz & Couper-Kuhlen (1999) le marquage prosodique peut caractériser aussi bien des séquences de style direct que des séquences de style indirect. Ajoutons que, comme le relève Simon (2004 :112), « (...) le statut des objets prosodiques est loin d'être stable et homogène ; si certains morphèmes prosodiques, identifiés pour le français, véhiculent une signification relativement conventionnelle, bien des observables prosodiques ne se laissent pas réduire à une forme stable à laquelle on peut attacher, de manière récurrente, une fonction précise, associée à un niveau de fonctionnement du discours (morphème de frontière *ou* morphème énonciatif *ou* autre chose...). »

Les travaux des conversationnalistes – Couper-Kühlen, Selting, Auer etc, montrent qu'une grande partie des indicateurs prosodiques procèdent par contraste local, dans une ou plusieurs des dimensions où peut se manifester un changement prosodique : les régularités de F0 (hauteur et amplitude des variations de hauteur), du débit, et de l'intensité, auxquels il faut ajouter le timbre (Fónagy 1983 ; Grosjean 1995 ; Léon 1993), fournissent autant d'« indices de contextualisation ». Dans la mesure où l'établissement d'un contraste met en rapport une manifestation « passée » avec une manifestation « présente », tout contraste opère une forme de mention, minimale, implicite, du trait en contraste, en une sorte d'« auto-phonie endémique » ; pour se convaincre de ce que ce n'est pas une vue de l'esprit, on peut rappeler l'anecdote fameuse de l'échange entre une dame et un petit garçon :

« regarde Mamie, un TVG
– un téGEvé
– un té VEgé (...),

où l'on voit l'importance de l'élément implicitement rappelé par le contraste, le petit garçon modifiant *l'accentuation* sur la syllabe mais pas sa *consonne*. Néanmoins, dans la mesure où cette « auto-phonie endémique » relève de l'emploi ordinaire, on ne la considérera pas dans le cadre de l'approche modulaire.

Dès lors que se manifeste un *changement* dans l'une ou l'autre des dimensions qui viennent d'être évoquées, et pour autant qu'il survienne en concomitance – en phase – avec une frontière d'une autre dimension discursive, notamment morpho-syntaxique, il peut être associé à un changement affectant l'origine énonciative,

modification d'état (intérêt ou attitude), ou changement de source. Ce fonctionnement postule une certaine « origine énonciative par défaut », et un régime, par défaut, d'*emploi* (Vs de *mention* ou de *représentation*) des unités; c'est relativement à ces repères, propres au discours *produit*, que se manifestent les changements.

Insistons sur le fait que l' « origine énonciative par défaut » et les traits prosodiques qui lui sont propres sont définis en proportions variables par les propriétés du cadre interactionnel, et par l'interprétation qu'en font les sujets parlants producteurs empiriques ; ainsi par exemple la prosodie de la journaliste du Journal télévisé de TF1 diffèrera-t-elle de la prosodie propre au sujet parlant Claire Chazal. En sens inverse, l' « effet de vivant » accordé au discours représenté au style direct peut être vu comme un relatif *déplacement* du cadre interactionnel, dans la mesure où son interprétation, tant par les changements de cadre déictique (interprétation des pronoms et autres expressions déictiques) que par sa prosodie, abolit localement le cadre extérieur du discours produit.

Les différents « canaux » prosodiques (propriétés mélodiques, rythmiques-temporelles, dynamiques d'intensité, de timbre), par ailleurs, agissent comme les « voix » dans un chœur, potentiellement indépendantes¹⁰, dans le marquage du discours représenté. Bertrand et Espesser (1999) par exemple montrent que l'intensité et ses variations, le tempo et ses variations, ne sont pas corrélés significativement à l'alternance « discours direct-discours représenté », contrairement à l'amplitude et à la dynamique tonale (variations de hauteur et placement de la moyenne dans la tessiture du locuteur), qui présentent des corrélations statistiquement significatives avec les alternances discours produit – discours représenté.

Plus récemment, Bertrand & Espesser 2002 montrent que si les variations de F0 permettent de distinguer des segments de discours produit (« direct ») de segments de discours direct rapporté, le

10 Même si elles sont nécessairement co-présentes ; on pourrait, comme pour l'« auto-phonie endémique » évoquée ci-dessus, être tenté de voir là un fait pertinent pour l'examen des faits de polyphonie dans le discours. En son état actuel, le modèle modulaire ne cherche cependant pas à capter ces phénomènes « micro-polyphoniques », s'en tenant à la description de manifestations discursives « ordinaires », discours produits ou représentés de diverses façons, mais pas à la description 'polyphonique' interne des objets et processus sémiotiques qui les construisent. Nous nous en tiendrons ici à ce « seuil » imposé par le modèle. Simon & Auchlin 2001 proposent une conception « gricéenne » de coopérativité entre ces diverses « voix » et les différents plans d'organisation du discours.

discours rapporté d'une source autre que le locuteur peut se distinguer du discours rapporté « autophonique » sur la base non plus d'une « saillance mélodique » (Caelen-Haumont 1991) mais d'une « saillance temporelle », montrant un usage différencié complexe de ces deux dimensions, paramètres tonaux – mélodiques, et temporels.

Toutefois, les auteurs notent que ces changements, attestés dans leur étude de cas, peuvent être dus aux emplois particuliers des discours rapportés, auto-phonique et polyphonique (discours d'une source autre), visant à construire une image spécifique du locuteur. Autrement dit, un phénomène observé peut n'être qu'un épiphénomène d'un autre ; une « fonction » supposée peut n'être qu'une fonction adjacente, subordonnée ou subordonnante. Telle modulation prosodique – augmentation de la hauteur et accroissement des écarts mélodiques, par exemple, pourra donner accès à un « effet de vivant », à une « théâtralisation », visant des fins expressives, et ne servir qu'indirectement à l'identification du statut énonciatif du « discours représenté ».

Malgré ces réserves, nous tenterons de préciser les contributions de la prosodie aux organisations énonciatives et polyphoniques en distinguant plusieurs fonctions distinctes, au moins du point de vue théorique :

- à un premier niveau, minimal, une fonction de « bornage » d'unités en fonction de leur statut (ou régime) énonciatif, comme discours produit (texte *employé*) ou discours représenté (*mentionné*) ; les marques prosodiques qui sont comme des guillemets oraux ¹¹ sont une rupture de phonation suivie d'un accent initial sur l'unité suivant la pause, qui, accompagné d'un changement de débit (accélération, en général) construit un arc accentuel sur l'unité concernée par le changement (§2.1.)
- à un second niveau en termes de complexité intervient un marquage attitudinal du producteur à l'égard du discours représenté, présentant notamment adhésion ou rejet ; les reprises diaphoniques manifestent ainsi des degrés variables de prise en charge et d'adhésion au discours représenté. (§2.2.)

11 Il est difficile de tenir durablement l'hypothèse que les guillemets typographiques sont une traduction visuelle de marques initialement prosodiques, fût-elle dotée de fondements historiques ; la pratique et les habitudes de lecture à haute voix imposent bien davantage de traduire oralement des guillemets écrits que l'inverse (la même chose vaut d'autres parallélismes envisagés entre prosodie et ponctuation – voir Ferrari et Auchlin 1997 p. ex.).

- enfin, à un troisième niveau, il y a les différenciations qui permettent de construire l'origine du discours représenté et de le « caractériser ». Les effets de théâtralisation et d'emphase souvent associés au discours représenté direct, l'effet de « vivant » qui peut lui être lié, relèvent de ce niveau. (§2.3)

2.1. Marques de début et de fin : bornage des segments de discours représenté

2.1.1 Bornes initiales

La prosodie, notamment en lecture à haute voix, permet de repérer, par une borne initiale, un segment de discours affecté d'un changement de régime énonciatif, en mention : la borne est constituée par un accent initial après une micro-pause, associé à un léger décrochement mélodique contre en haut, et une accélération perceptible ouvrant un « arc accentuel » (Di Cristo 1999) dont la portée détermine celle du régime modifié. On peut s'en convaincre en oralisant les exemples 7a et 8a ci-dessus.

Cet effet interprétatif, dans ces exemples, dépend en grande partie de la position syntaxiquement intégrée des termes contrastés ; de telles attaques hautes, en début d'énoncé ou autre « frontière discursive », peuvent être associées à d'autres fonctions (Grobet & Auchlin 2001). De plus, l'accent initial et l'arc accentuel marquent davantage le changement de régime énonciatif que la mention en tant que telle, dans la mesure où ils fonctionnent également comme marqueurs d'insistance (Mertens 1987 : 137) ; le décrochement mélodique ascendant peut être associé à des effets de sens variés, qui vont de l'ouverture à la co-énonciation, chez Morel & Danon-Boileau (1998), à la présentation de discours représenté ; il peut être ascendant, mais également descendant, en un effet de parenthèse basse, ou d'incise, qui peut être contraint par l'apparition de discours représenté (Simon 2004 : 247).

2.1.2 Bornes finales

Les abaissements ou rehaussements de registre ont des débuts nets correspondant à des frontières bien marquées, mais des achèvements souvent graduels et insensibles, sinon par l'effet de « remise à zéro » des paramètres à l'ouverture d'une nouvelle unité. Dans l'exemple

suivant emprunté à Grobet (à paraître a), l'auteure relève un tel effet de clôture par diminution progressive d'activité :

- (13) FS : ras-le-bol \ (0.4s.) vraiment (XXX) vingt ans qu'j'en parle / (0.4s.) qu'on me pose les mêmes questions → **vo**tre **pe**tit **mi**lieu **tra**la**la** **ar**tificiel **et**cetera / (0.7s.)
 JC : parce qu'on disait petit milieu \
 FS : oui c'est → **pe**tit **mi**lieu **ar**tificiel **et** **do**ré **la**la / **to**u**jo**urs **le** **m**ême (**la**la) / . (0.6s.) et : je m'obstinais à expliquer que non / que c'était pas ça / (...)

Les deux cas de discours rapporté polyphoniques fléchés se caractérisent par plusieurs indices signalant la non prise en charge des propos par la locutrice : vocabulaire connoté péjorativement (*petit milieu*) et qualifications telles que *artificiel*, *doré* ; abréviations par *etcetera* (qui peut être considéré comme une marque de discours rapporté, Authier-Revuz 1985) et *lalala* ; du point de vue prosodique : mélodie plate, aux intervalles réduits. La fin de ces séquences est marquée de manière parallèle aux niveaux lexical et prosodique : pour Grobet (à paraître a), les marques telles que *etcetera*, *lala*, couplées à une diminution de l'intensité générale, « témoignent d'une diminution de l'activité articulatoire, comme si la locutrice était tellement désintéressée par le propos qu'elle ne se donnait pas la peine de l'articuler jusqu'au bout ». Cette baisse d'activité contraste fortement avec les attaques hautes qui marquent le début de la séquence suivante, présentée, par là, comme prise en charge par la locutrice (comme discours produit). C'est ce contraste qui marque ainsi la borne finale du discours représenté (Grobet à paraître a et b).

Au demeurant, une part essentielle du marquage prosodique fonctionnel en français affecte les dernières syllabes des unités (groupes ou paquets intonatifs, à différents niveaux) ; Mertens (1987 : 108-109) relève l'existence d'une séquence spécifique de « morphèmes intonatifs », combinant les tons *...h ... \HH*, et *...h...B-B-*¹² « souvent employée pour évoquer des opinions générales, des lieux-communs, ou pour indiquer l'acception figurée ou générique des mots. Le contour apparaît fréquemment dans la citation et dans les énumérations d'éléments présentés comme allant de soi. *Dans les*

12 B= bas, H = haut, /H « rehaussement » ou \H « abaissement » du ton, inférieur à une tierce musicale ou 5 demi-tons ; B-B- désigne l'infra-grave qui marque la frontière prosodique maximale (Mertens 1987 ; Simon 2004). Les minuscules désignent des syllabes non accentuées ; les majuscules des syllabes accentuées, redoublées lorsqu'il s'agit d'un accent final.

deux cas, le locuteur présente ses paroles comme si c'étaient celles d'un autre ou celles des autres » (c'est nous qui soulignons).

Simon quant à elle (2004 : 280-282) observe que la fin de séquences de discours représenté peut mobiliser le ton B-B- (infra-grave), frontière prosodique maximale, à des endroits qui, tant du point de vue praxéologique que textuel, ne pourraient être considérés comme des frontières discursives maximales (fin d'énoncé produit). Ainsi dans l'exemple suivant, emprunté à Mertens (1987) ¹³ :

- (14) ^bvous avez dit^{BB} un jour^{H/H} **RB**^b que la littérature^{H/H} allait à sa perte^{B-B-}
^bc'était un mot^{BB} RB[] ^bou bien vous l' RB[] pensez vraiment^{B-B-}

Le segment de discours représenté (en gras), attribué explicitement à RB, est borné par le ton B-B-, qui prend en charge le marquage d'une fonction de clôture du discours représenté, au détriment de sa fonction de structuration du discours produit – dans la mesure où il marque ici comme maximal un regroupement textuel qui n'est manifestement pas maximal dans le discours produit, entraînant de ce fait un effet d'« enjambement » de contraintes à un certain niveau (structuration prosodique de l'intervention citante) par un autre ensemble de contraintes (ici : praxéologiques ; on mentionne le discours de l'interlocuteur pour en faire quelque chose).

Quant à la prise en charge du morphème B-B-, on peut se demander si elle appartient au discours représenté, dont elle marquerait l'achèvement, ou au discours produit, où elle marquerait l'achèvement d'une activité de représentation de discours, ou de préparation. Mais on peut, au contraire, considérer qu'il n'y a pas d'ambiguïté, et que la marque appartient simultanément aux deux discours.

2.2. Discours représenté « commenté »

Le *bornage* prosodique des segments de discours représentés paraît donc sujet à une grande variation. Au delà des marques initiales et finales repérant les bornes de segments représentés, les segments eux-même peuvent être affectés de divers effets de coloration prosodique, destinés à manifester une certaine « attitude » du locuteur vis-à-vis du discours représenté ; *prise en charge* ou *rejet*, accord ou désaccord

¹³ Entretien entre Jacques Cancel et Roland Barthes. Voir la note 12 pour la transcription de la prosodie.

avec la position citée, sont des ‘modalités polaires’ ; il y a des positions intermédiaires (*doute*) et d’autres modalités (*désir* ; *dégoût* ; etc.). Elles sont rendues accessibles entre autres par le timbre et le jeu des mécanismes de mimique vocale (Fónagy 1983) propres au marquage vocal des états affectifs (Scherer 1989), qui se combinent aux autres dimensions de variation prosodique.

Ainsi dans l’exemple 13 la locutrice manifeste une attitude de dédain et de désintérêt, de « désengagement » énonciatif, par une baisse générale de contrastes (articulatoires, mélodiques et d’énergie) affectant les segments de discours représenté. C’est un effet interprétatif des ‘codes biologiques’ (Gussenhoven 2002), ici particulièrement les codes « de l’effort » et « de la production », qui traitent de l’intégration des paramètres dynamiques (énergie) et de l’empan de production – ici, une fin de segment de production.

Certaines reprises diaphoniques « écho » recourent au même type de mécanisme, particulièrement celles qui assument la fonction « minimale » des reprises diaphoniques (Perrin 1999), la ‘prise en compte’, ou accusé de réception. Dans l’exemple suivant :

- (15) (contexte : C client entre dans la librairie et s’adresse à V vendeur)
 C : bonjour - est-ce que vous avez Contes et nouvelles d’Oscar Wilde
 V : →^BOscar Wilde^{BB} – ^ben^H français^{H/H}

le segment fléché est prononcé avec une mélodie plate sur un ton bas, avec un léger accent initial ; cette prise en compte, prototypique, tient ses caractéristiques à la fois de son amont (caractère diaphonique) et de son aval (statut de topique, pour la demande de précision qui suit). L’attaque basse joue un rôle : une attaque haute indiquerait, de la part de V, une reprise du segment hors de son contexte initial, comme si, par exemple, V disposait d’une information particulière à propos de cet auteur, alors que l’attaque basse donne l’instruction de « raccrocher le constituant au constituant déjà commencé » (Grobet & Auchlin 2001 : 172) ; le ton *BB* en finale indique quant à lui le statut non terminal du constituant.

Le ton de la syllabe pénultième est pertinent lui aussi : si l’on remplace le ton *b* que prend dans cet environnement l’avant-dernière syllabe¹⁴ par un ton *h* comme dans :

14 C’est bien sa place de pénultième qui est porteuse de ce ton ; si l’on augmente le nombre de syllabes de la chaîne le ton ne sera pas porté par la deuxième, ni celle du milieu, mais par la pénultième.

(16) ^BOs^hcar Wilde^{BB}

l'effet obtenu peut être glosé par « vous avez bien dit *Oscar Wilde* ? », et manifeste un étonnement du locuteur à l'égard de la séquence. Cette dernière forme n'est pas sans rapport avec le morphème *...h...HH* noté ci-dessus, dont elle est une variante abaissée, voire « forcée contre en bas » (indications de timbre); elle semble disposer d'un certain statut conventionnalisé de 'morphème intonatif'; et elle a un contenu instructionnel spécifique, qui contient les informations suivantes : - la chaîne syllabique qui le porte n'est pas prise en charge par le locuteur, qui n'assume qu'une répétition; la séquence est un écho d'un élément proche dans l'amont discursif¹⁵.

L'exemple 12 repris ci-dessous présente aussi ce morphème de « reprise dubitative » :

(17) GD (...) peut-être le fait d'avoir peur de rentrer dans la quatrième décennie heu - bon ça fait que - ça fait heu que j'suis p'tête moins patient - et en même temps j'commence à [l] l'être plus mais enfin
 CD ^bvous avez pas^{BB} une quarantaine^{BB} /b flippante^{HB}
^bquand même hein^{b/b} (...)

La configuration discursive dans laquelle intervient cette forme prosodique est celle de l'ouverture d'un échange subordonné dans une intervention de l'interlocuteur, qui introduit un commentaire correctif. La reprise interprétative (non littérale) des propos de l'interlocuteur (*quarantaine flippante*) est présentée, par cet effet d'**écrasement de l'espace tonal**, comme mise en doute par le locuteur; le morphème de focalisation HB sur « flippante^{HB} » clôt prosodiquement l'énoncé complet de CD (*quand même hein* est un 'appendice'); dans la mesure où l'on peut paraphraser cet énoncé par *ne me dites pas que vous avez une quarantaine flippante*, sa portée semble plus large que le simple adjectif focalisé, et s'étend à la proposition enchâssante (*ne me dites pas que*). L'effet d'**écrasement de l'espace tonal** procède ici de la combinaison de paramètres de F0 (minimisation des contrastes dans la partie basse de l'espace tonal), et du timbre, plus sombre et plus sourd, lié à une posture phonatoire « en retrait » et à une mimique vocale de bouche fermée et arrondie.

15 Autrement dit son caractère diaphonique est contextuel : on peut imaginer la séquence par exemple en écho à une information radiophonique, adressée à quelqu'un, dans un contexte interactionnel distinct : « tu as bien entendu comme moi – ils ont dit *Oscar Wilde* ? ».

D'autres faits exemplifient la fonction « commentative » de la prosodie. Perrin (1999 : 453) distingue, à la suite de De Fornel (1997) les réponses écho « positives », marquées par une « intonation descendante », des simples « échos », marqués par un « chantonement ». Le ton B-B-, « descendant », à la place du BB dans l'exemple 16 produirait bien un effet de prise en compte avec accord ; et le « chantonement » pourrait être lié à la combinaison de l'attaque basse, du ton haut non accentué de la pénultième, et du ton final « non conclusif », s'opposant au ton B-B-, et marquant un appel à continuation.

Le commentaire prosodique peut également dépasser les frontières du discours cité : l'exemple suivant implique une diaphonie « à distance » (Vs « locale » ; Torck 1994). De telles mentions sont généralement marquées verbalement, leur caractère de reprise n'étant pas accessible dans l'environnement immédiat. La séquence d'où est tiré l'exemple 12 ci-dessus commence par :

- (18) CD bon ben heu : une belle déclaration en conclusion – que vous allez quand même un peu m'expliquer - ↓ vous dites^{BB} j'ai l'âge^{BB} des conneries^{BB} magnifiques^{B-B-} ↑ Helles ont commencé^{BH}

L'abaissement de registre (↓ ... ↑) encadre non pas le discours représenté mais toute la séquence comportant ce discours ; l'abaissement de registre (et d'intensité, corrélé, ici) peut servir à indiquer et préparer l'enchaînement dont le segment abaissé constitue le topique ; un autre exemple de discours diaphonique analogue ne présente pas cet abaissement :

- (19) CD (...) ben justement heu heu ça [en]chaîne bien ça parce que - heu vous dites^{HH} GD [la nature me rassure^{HH} parce qu'il n'y a aucun^{HH} pouvoir à exercer^{B-B-}] - finalement vous détestez le pouvoir

Cet exemple présente, plutôt, un effet rythmico-temporel : le début est prononcé avec un débit rapide, jusqu'à un grand allongement et une coupure de *au - cun* destinée à placer un accent fort sur la dernière syllabe ; cette 'emphase accentuelle' est intéressante, en ce qu'elle est attribuable simultanément et indistinctement aux locuteurs cité GD et producteur CD, qui paraît marquer là son adhésion au contenu qu'il représente.

2.3. Figuration de la source, théâtralisation, mise en scène

De nombreux travaux lient le marquage prosodique du discours rapporté à un effet de *théâtralisation* (Jacob 1987) ou d'*emphatisation* (Bertrand 1999 : 281) ; mélodie, intensité, rythme, timbre, marquent la prise en charge des séquences de discours rapporté – ou leur absence de prise en charge (ex. Sagan ci-dessus). L'effort voco-prosodique engagé peut également servir à la *caractérisation*, d'un *individu* (parodie, caricature vocale), ou d'un *type*, entité générique, classe sociale, provenance régionale. Les techniques de caractérisation incluent l'« émailage » (Bally, rappelé par Léon 1993), et ses matériaux incluent notamment les clichés, mélodiques (*Il était une fois, ...*), ou prosodiques, qu'ils affectent tout un discours, comme les phonostyles (Léon 1993), ou agissent localement, comme le faisceau de traits présentant une parole comme émanant de la sagesse populaire, et 'faisant sentence'. Caractérisation, figuration d'une origine de parole, théâtralisation, disposent d'une fonction de 'présentation' d'entités, de type narratif ; mais toute narration peut être intégrée dans une structure de type argumentatif (Roulet & al 2001).

La fin de l'exemple 1 illustre comment différents effets de théâtralisation peuvent se combiner :

- (20) (...)
 C alors *il me dit* M[alors le nord il est où ?] .
 je dis C[**ben tu te tournes jusqu'à ce que** - ^Bla **flèche**^{BH} -
^Bt'**indique**^{HH} - /^Hle **nord**^{HB}]
 V (rire)
 C il a toujours pas compris

Comme le montrent Auchlin, Filliettaz, Grobet & Simon (2004 : 235), les segments de discours représenté (analysés ci-dessus) sont marqués par une expansion et un rehaussement de l'espace tonal (de 5 demitons) ; les auteurs montrent également que la fin de la réplique rapportée de C fait l'objet d'une forme de scansion, liée à un très net ralentissement, à une segmentation par des pauses et des accents marqués sur chaque paire de syllabe. Cette sur-segmentation scandée convoque, dans le discours représenté, le phono-style de la maîtresse d'école ; C montre en s'en parant qu'elle occupe telle place et tel rôle communicationnel face à M son interlocuteur.

Ajoutons que ce récit conversationnel vise à légitimer la demande de C, qui veut donner à M une « petite leçon » ; l'ostension du phonostyle, comme ressource prosodique polyphonique, dans le discours représenté, joue donc un rôle argumentatif clair.

La fonction de présentation voco-prosodique d'entités à des fins descriptives ou narratives entretient cependant des liens fort complexes avec les unités linguistiques qui les portent, comme le montre l'exemple suivant :

(21) « Oskawawa est un énorme hippopotame ; Iskiwiwi est un tout petit indien ; (...) »

Tout naturellement, la fillette qui « dit l'histoire » emprunte une « voix d'hippopotame », la plus « impressionnante », massive », « lourde », qu'elle puisse pour prononcer, de façon très lente, la première partie. Ensuite elle prend une voix très aiguë et un tempo rapide. Le changement vocal ici est consacré à l'évocation des personnages, un gros lourd, lent, un petit, agile, rapide.

On voit à l'oeuvre : la hauteur vocale (F0), évidemment plus basse pour le gros animal (cf. Gussenhoven 2002, i.a.), le débit, plus lent pour le plus gros, et le timbre, paramètre à la fois hautement manifeste, et parmi les plus complexes à décrire – bien que relativement aisément modifiable, en tout cas dans ses aspects liés à la posture phono-articulatoire.

La composition de ces indices avec le sens du discours mérite aussi un commentaire : il s'agit d'une « voix d'emprunt », mais elle est destinée à porter des mots que ne prononcerait certainement *pas* la source discursive évoquée.

Pour autant, et bien qu'il ne s'agisse pas ici de discours « représenté » au sens où les mots prononcés pourraient, auraient été, ceux de la source dépeinte, ce discours représente sous deux formes différentes, l'une textuelle, l'autre prosodique, le(s) même(s) personnage(s), dans un « récit » qui se place comme récit par le décalage, manifeste, entre les voix affichées et ce que l'auditeur reconstruit – ou connaît – de la « voix standard » du locuteur. On a donc ici deux « représentations » de personnages, ou, plus exactement, une *représentation* (verbale) et une *présentation* (vocale), qui s'assemblent, en une figure comparable aux « blends expérientiels » (Auchlin 2003).

3. CONCLUSION

L'approche modulaire appréhende le discours représenté dans son contexte discursif et l'analyse à plusieurs niveaux impliquant différents degrés de complexité (organisation énonciative, organisation polyphonique). Elle permet d'analyser les apports des manifestations voco-prosodiques aux faits de polyphonie en distinguant différents niveaux de contribution, et en rendant compte de faits complexes par différents couplages entre dimensions (comme par exemple la marque structurelle de continuation, avec l'abaissement forcé, produit un effet de mise en doute d'un segment diaphonique).

Face aux données prosodiques et à leur contribution au façonnage du discours, l'approche modulaire désigne l'un de ses horizons de développement : la prise en compte des polyphonies internes aux différents systèmes sémiotiques coopérant au discours, et la description de leurs différents *jeux*.

BIBLIOGRAPHIE

- AUCLIN, A., 2003, « Compétence discursive et co-occurrence d'affects : “blends expérimentiels” ou (con)fusion d'émotions ? », in COLLETTA J.-M. & TCHERKASSOF A. (dir.), *Les émotions. Cognition, langage et développement*, Hayen, Mardaga, p. 137-152.
- AUCLIN, A. & FERRARI A., 1994, « Structuration prosodique, syntaxe, discours : évidences et problèmes », *Cahiers de linguistique française* 15, p. 187-216.
- AUCLIN, A., FILLIETTAZ L., GROBET A. & SIMON A. C., 2004, « (En)action, expérimentation du discours et prosodie », *Cahiers de linguistique française* 26, p. 217-249.
- AUTHIER-REVUZ J., 1985, « La Représentation de la parole dans un débat radiophonique : Figures de dialogue et de dialogisme », *Langue Française* 65, p. 92-102.
- BAKHTINE, M., 1977, *Le marxisme et la philosophie du langage*, Minit, Paris.
- BAKHTINE, M., 1978, *Esthétique et théorie du roman*, Gallimard, Paris.
- BERTRAND, R., 1999, *De l'Hétérogénéité de la Parole. Analyse énonciative de phénomènes prosodiques et kinésiques dans l'interaction interindividuelle*, Thèse de doctorat, Université d'Aix-Marseille 1.
- BERTRAND, R. & ESPESSER, R., 1998, « Prosodie et discours rapporté: la mise en scène des voix », in VERSCHUEREN J. (ed.), *Pragmatics in 1998: Selected papers from the 6th International Pragmatics Conference*, vol. 2, Anvers, IPrA, p. 45-56.
- BERTRAND, R. & ESPESSER, R., 2002 « Voice diversity in conversation : a case study », in BEL B. & MARLIEN I. (éds), *Speech Prosody 2002*, Aix-en-Provence, p. 171-174.

- Cahiers de linguistique française* 23, 2001, *Prosodie : carrefour entre syntaxe, analyse du discours, psychologie des émotions et interprétation simultanée*, Département de linguistique, Université de Genève.
- COUPER-KUHLEN, E., 1996, « The prosody of repetition : on quoting and mimicry », in COUPER-KUHLEN E. & SELTING M. (éds) *Prosody in conversation*, Cambridge, Cambridge University Press, p. 366-405.
- COUPER-KUHLEN, E., 1998, « Coherent Voicing. On Prosody in Conversational Reported Speech », *Inlist* 1, p. 1-25.
- DEMERS, M., 1998, *La prosodie du discours rapporté*, Québec, Ciral.
- DI CRISTO A., 1999a, « Vers une modélisation de l'accentuation en français (première partie) », *Journal of French Language Studies* 9/II, p. 143-163.
- DI CRISTO, A., 1999b, « Vers une modélisation de l'accentuation en français (seconde partie) », *Journal of French Language Studies* 10/I, p. 27-45.
- DUCROT, O., 1984, « Esquisse d'une théorie polyphonique de l'énonciation », in *Le dire et le dit*, Paris, Minuit, p. 171-233.
- FIANDINO, CH., G. FITZGERALD, & VAISSIÈRE, J., 2001, « Is there a specific intonation for doubt ? Preliminary study », *Journées Prosodie Grenoble*, octobre 2001.
- FÓNAGY, I., 1991[1983], *La vive voix, essais de psycho-phonétique*, Paris, Payot.
- GROBET, A., 2002, *L'identification des topiques dans les dialogues*, Bruxelles, De-Boeck-Duculot.
- GROBET, A., à paraître a, « La fin des séquences de discours rapporté est-elle marquée dans le discours oral ? », in LOPEZ MUÑOZ J. M., MARNETTE S. & ROSIER L. (éds.), *Le Discours rapporté dans tous ses états*, Paris, L'Harmattan.
- GROBET, A., à paraître b, « Echanges de propos. Du discours représenté à la co-construction du discours », in LAGORGETTE D. (éd.) *Actes du colloque de Chambéry*, cd-rom.
- GROBET, A. & AUCHLIN, A. 2001, « À l'attaque ! Vers une typologie des différentes prises d'élan dans le discours », *Cahiers de linguistique française* 23, p. 165-187.

- GUSSENHOVEN, C. 2002, « Intonation and interpretation : phonetics and phonology », *Speech Prosody* 2002.
- JACOB, L., 1987, « Discours rapporté et intonation : illusion ou réalité de la polyphonie ? », *Études de linguistique appliquée* 66, p. 71-87.
- JEANNERET, T., 1999, *La coénonciation en français. Approches discursive, conversationnelle et syntaxique*, Berne, Lang.
- JOHANSSON, M., 2000, *Recontextualisation du discours d'autrui. Discours représenté dans l'interview politique médiatique*, Turku, Turun Yliopisto.
- KLEWITZ, G. & COUPER-KUHLEN, E., 1999, « Quote-Unquote? The Role of Prosody in the Contextualization of Reported Speech Sequences », *Pragmatics* 9, p. 459-485.
- KUYUMCUYAN A., 2002, *Diction et mention*, Berne, Peter Lang.
- MERTENS P., 1987, *L'intonation du français. De la description linguistique à la reconnaissance automatique*, mimeo, thèse de doctorat, Université catholique de Louvain.
- MOREL, M.-A. & DANON-BOILEAU, L., 1998, *Grammaire de l'intonation : l'exemple du français*, Paris, Ophrys.
- NØLKE, H., 1994, *Linguistique modulaire : de la forme au sens*, Louvain, Peeters.
- PERRIN, L., 1994, « Mots et énoncés mentionnés dans le discours », *Cahiers de linguistique française* 15, p. 217-248.
- PERRIN, L., 1995, « Du dialogue rapporté aux reprises diaphoniques », *Cahiers de linguistique française* 16, p. 211-240.
- PERRIN, L., 1996, « De la structure énonciative et de l'organisation polyphonique d'un échange épistolaire », *Cahiers de linguistique française*, 18, p. 129-156.
- PERRIN, L. 1999. « La fonction des reprises diaphoniques locales dans le dialogue », in Verschueren J. (éd.) *Pragmatics in 1998. Selected papers from the 6th international pragmatics conference* (vol. 2), Anvers, International Pragmatics Association, p. 448-461.
- PERRIN, L., 2000, « L'argument d'autorité comme forme de modalité "allusive" dans la conversation », in MARTEL G. (éd.), *Autour de l'argumentation*, Québec, Nota Bene, p. 81-106.

- PERRIN, L., à paraître, « Polyphonie et séquence écho », in Bres J., Haillet P.P., Mellet S., Nølke H., Rosier L., *Dialogisme, polyphonie : approches linguistiques*, Duculot (Actes du colloque de Cerisy de septembre 2004).
- PERRIN, L. & VINCENT D., 1997, « L'interprétation du verbe dire en contexte conversationnel : des clauses performatives aux clauses descriptives », *Revue de sémantique et pragmatique* 2, p. 201-217.
- PERRIN, L., DESHAIES, D. & PARADIS C., 2003, « Pragmatic functions of local diaphonic repetitions in conversation », *Journal of Pragmatics*, Vol. n° 35, p. 1843-1860.
- ROULET, E., 1991, « Vers une approche modulaire de l'analyse du discours », *Cahiers de linguistique française* 12, p. 53-81.
- ROULET, E., 1993, « De la structure diaphonique du discours épistolaire. À propos d'une lettre d'Aurore Dupin à sa mère », in *Mélanges offerts à Jean Peytard. Annales littéraires de l'Université de Besançon* 502, p. 85-99.
- ROULET, E., 1997, « L'organisation polyphonique et l'organisation inférentielle d'un dialogue épistolaire », *Cahiers de linguistique française* 19, p. 149-179.
- ROULET, E., 1998, « Polyphony », in VERSCHUEREN J., ÖSTMAN J.-O. & BLOMMAERT J. (éds), *Handbook of pragmatics*, Amsterdam, John Benjamins.
- ROULET, E. *et al.*, 1985, *L'articulation du discours en français contemporain*, Berne, Lang.
- ROULET, E., GROBET A., FILLIETTAZ L. & BURGER M., 2001, *Un modèle et un instrument d'analyse de l'organisation du discours*, Berne, Lang.
- SCHERER, K. R., 1989, « Les émotions : fonctions et composantes », in RIMÉ B. & SCHERER K. (éds), 1989, *Les émotions*, Neuchâtel, Delachaux & Niestlé, p. 97-134.
- SELTING, M., 1996, « Prosody as an activity-type distinctive cue in conversation : the case of so-called astonished questions in repair initiation », in COUPER-KUHLEN E. & SELTING M. (éds), *Prosody in conversation*, Cambridge, Cambridge University Press, p. 366-405.
- SIMON, A.C., 2004, *La structuration prosodique du discours en français*, Berne, Lang.

- SIMON, A.C. & AUCLIN A., 2001, « Multimodal, multifocal ? Les *hors-phase* de la prosodie », in CAVÉ C., GUAÏTELLA I. & SANTI S. (éds), *Oralité et gestualité. Interactions et comportements multimodaux dans la communication*, Paris, L'Harmattan, p. 629-633.
- TORCK, D., 1994, « Diaphonie et interprétation », *Littérature*, 93, p. 15-30.
- VINCENT, D. & DUBOIS, S., 1997, *Le discours rapporté au quotidien*, Québec, Nuit Blanche.