

# Interprétation des scansions rythmiques en français

*Anne Catherine Simon et Anne Grobet*

[simon@rom.ucl.ac.be](mailto:simon@rom.ucl.ac.be) / [anne.grobet@lettres.unige.ch](mailto:anne.grobet@lettres.unige.ch)

Université catholique de Louvain / Université de Genève

**Résumé:** Une description de la forme et de la fonction des scansions rythmiques régulières en français est proposée dans cet article. Cette recherche s'inscrit dans une approche perceptive (auditive) des phénomènes rythmiques, conçus comme des patterns résultant du travail perceptuel des auditeurs. Quant à l'interprétation fonctionnelle des scansions, elle repose sur la notion de *contextualization cue* (Gumperz 1992) et considère le rythme comme un des indices sur lesquels les interactants se fondent pour construire et sélectionner les aspects pertinents du contexte qui leur serviront de cadre interprétatif pour leurs énoncés. Cette étude se base sur l'analyse d'un nombre important de scansions issues de données authentiques en situation formelle (interactions en classe, interviews et émissions médiatiques, interactions de service) et informelle (conversations familiales ou entre amis).

## 1. Introduction

Le rythme est un composant fondamental de l'interaction humaine, y compris dans ses aspects langagiers. De nombreux travaux attestent de cet intérêt pour les aspects rythmiques du langage (Pasdeloup 1990, Couper-Kuhlen 1993, Barbosa 1994, Auer, Couper-Kuhlen & Müller 1999). Nous adoptons une approche du rythme qui est perceptive (subjective) et interactionnelle.

Une approche perceptive du rythme admet que l'identification de séquences rythmées (formant des groupements réguliers ou des scansions) est un produit de notre activité perceptive et que nous réorganisons le signal de parole en y projetant « des structures simplifiées et très nettes » (Fraisie 1967 cité dans Pasdeloup 1990 : 69). Ce n'est donc pas la métrique qui nous intéresse ici – en tant que virtualité d'accentuation de la langue – mais le rythme perçu dans la parole. Cette approche perceptive est soutenue par une analyse instrumentale (acoustique) des données.

Une approche interactionnelle du rythme permet d'interpréter les effets produits par les scansions rythmiques par rapport aux conduites des locuteurs dans des contextes d'interactions donnés. Nous ne nous intéressons donc pas aux aspects de production de la parole, ni au rôle du rythme dans l'impression de fluence de la parole spontanée (voir la variation du débit ou de la durée traitée par Zellner 1998). Les locuteurs impriment à certains de leurs énoncés des patrons rythmiques réguliers et identifiables, qui donnent une impression de scansion de leur discours, pour accomplir des tâches interactionnelles précises dans certains types d'interactions.

L'hypothèse centrale est que, pour interpréter les scansions rythmiques, il faut adopter un cadre qui tienne compte des macro-unités (par ex. une suite d'interactions, un genre spécifique, etc.) qui fournissent les cadres interprétatifs (Goffman 1974) utilisés par les locuteurs pour interpréter leur discours.

## 2. La perception d'une organisation rythmique régulière

La structuration rythmique est essentiellement une *construction perceptive* de l'auditeur, qui ne s'élabore pas uniquement à partir d'indices acoustiques, mais tient compte du sens du discours. Une définition très simple donne une première approximation de notre approche des structures rythmiques :

Rhythm is that property of a sequence of events in time which produces in the mind of the observer the impression of proportion between the durations of the several events or groups of events of which the sequence is composed. (Sonneschein 1925, 16, cité par Auer et al. 1999, 23)

Le rythme qui nous intéresse concerne des patrons (gestalts) prosodiques *perçus comme réguliers* dans leur organisation temporelle. Comme l'a relevé Padeloup, « la structuration rythmique ne s'élabore vraisemblablement pas à partir du seul niveau acoustique [...]. Le rythme procède d'un jeu subtil entre différents systèmes constitutifs du sens, intonatif, accentuel, syntaxique et lexical entre autres ; chacun de ces systèmes, tout en fonctionnant pour son propre compte, participe et impose des contraintes à l'organisation des autres systèmes » (1990, 49). L'interaction des différents paramètres est illustrée dans l'exemple (1), extrait de l'interview radiophonique d'une femme écrivain, qui présente un passage avec une régularité rythmique très perceptible (en gras dans la transcription) :

- (1) radAR1 en compagnie de Liliane Wouters nous avons évoqué / l'anthologie / le siècle des femmes consacrée à la poésie francophone en Belgique et au Grand Duché du Luxembourg au vingtième siècle une anthologie euh réalisée par Liliane Wouters et Yves Namur **et puis maintenant venons-en aux poèmes à proprement parler de Liliane Wouters** et c'est un / un livre un recueil de poèmes qui s'intitule ...  
[Valibel, radLW1r]

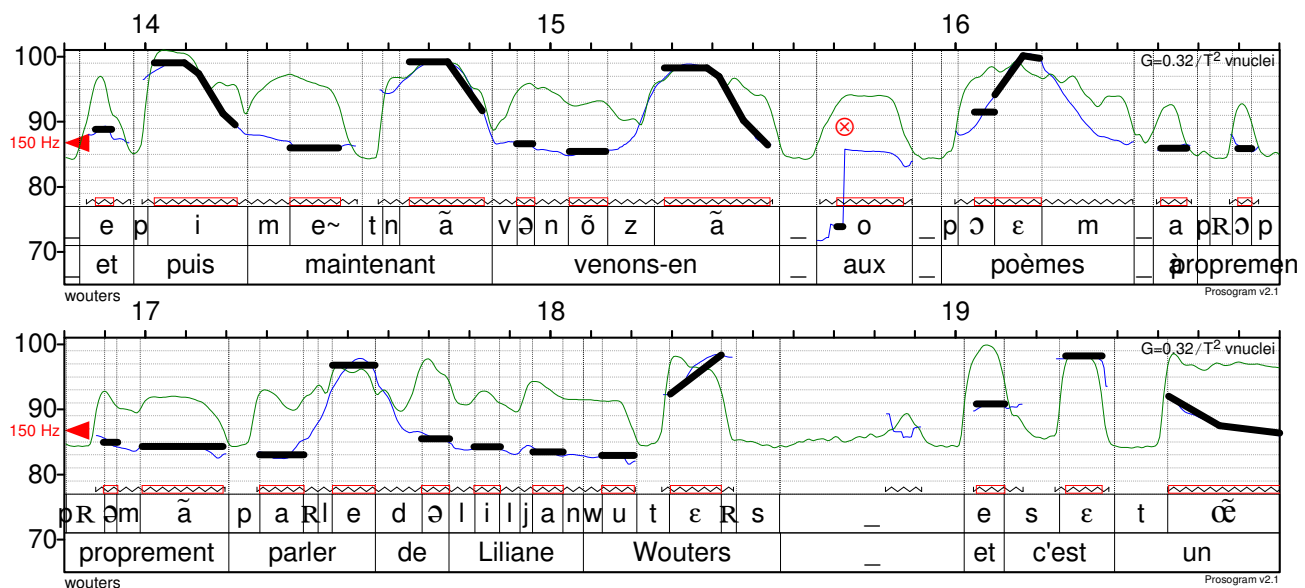


Figure 1 : Prosogramme<sup>1</sup> de l'exemple (1)

<sup>1</sup> Développé par Mertens (2004), le prosogramme est une transcription stylisée de l'intonation, basée sur un modèle de la perception et correspondant aux variations perçues de la fréquence fondamentale sur les noyaux syllabiques (traits noirs gras et graduation en demi-tons sur l'axe des ordonnées). Le prosogramme indique aussi les variations de l'intensité (trait grisé) et de la durée (graduation de l'axe des abscisses en dixièmes de secondes). Les petites croix entourées d'un cercle

Cet énoncé présente un nombre important de syllabes perçues comme proéminentes (*et puis maintenant venons-en aux po-èmes<sup>2</sup> à proprement parler de Liliane Wouters*). Perceptivement, ces proéminences sont réalisées par des mouvements intonatifs descendants (sur les 3 premières syllabes proéminentes) et par un allongement de ces syllabes, phénomènes mesurable sur le prosogramme de la figure (1). D'un point de vue intonatif, la répétition du même contour (visible elle aussi sur le prosogramme) crée une impression de régularité et organise l'énoncé selon un pattern qui se répète. Du point de vue temporel, ces mouvements mélodiques, en plus de présenter une forme similaire, sont réalisés à des intervalles temporels perçus comme identiques<sup>3</sup>. Les intervalles de durée entre les attaques des voyelles des trois premières syllabes proéminentes mesurent respectivement 631 et 638 ms (figure 2), ce qui a pour effet de créer un battement rythmique régulier sur le début de cet énoncé.

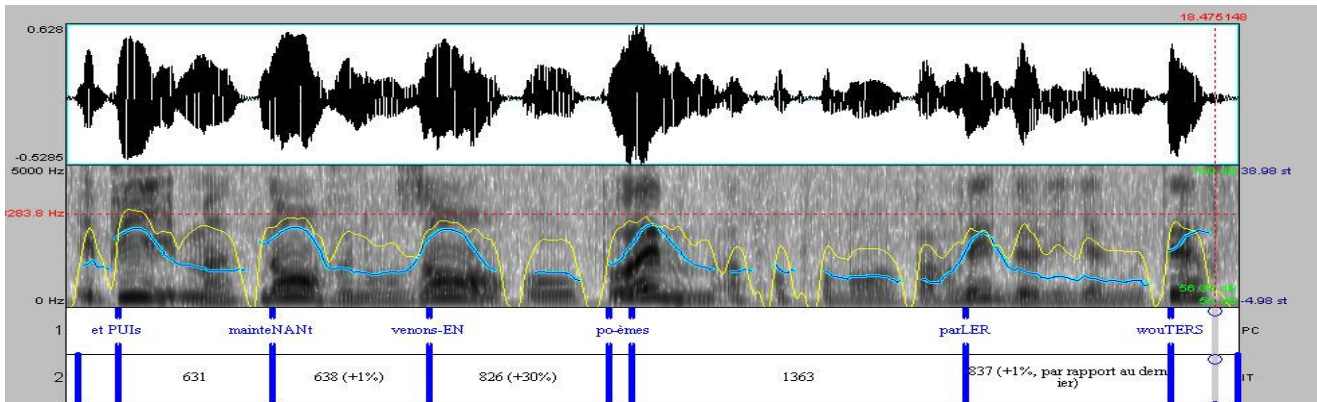


Figure 2 : Mesure des intervalles temporels (en ms) entre les attaques des voyelles (P-centres) des voyelles perçues comme proéminentes (Praat)

La syllabe constitue la plus petite unité rythmique, étant entendu qu'une organisation rythmique peut se développer sur des groupes de syllabes (un ou plusieurs éléments non proéminents étant alors regroupés autour d'un élément proéminent). Sans entrer dans une discussion sur la notion de *centre perceptuel* (P-center, voir Hoequist 1983), un accord semble exister sur le fait que la perception de l'écart temporel entre des syllabes proéminentes successives s'aligne sur l'attaque de la voyelle de la syllabe proéminente<sup>4</sup>. C'est ce point de référence que nous avons pris comme repère constant pour mesurer les intervalles rythmiques, c'est-à-dire les intervalles temporels entre les syllabes qui constituent les battements (*beats*) d'une séquence rythmiquement organisée. Selon Fraisse (1974, cité par Padeloup 1990, 62 sq.), une perception rythmique subjective<sup>5</sup> peut intervenir entre certaines limites : l'intervalle temporel minimal entre deux événements proéminents ne doit pas être inférieur à 150-200 ms (soit 5, 6 battements par seconde), ni excéder 1500-2000 ms (un battement chaque 1,5 à 2 secondes). Ainsi, « the optimal interval size for non linguistic stimuli seems to be around 400 msec » (Auer et al. 1999, 14). Enfin, les tests perceptifs de Couper-Kuhlen (1993) ont montré que ce seuil peut être exprimé non en termes absolus, mais relatifs : si la différence de durée entre deux intervalles successifs n'excède pas 20%, alors les intervalles sont *perçus* comme identiques. Si l'on récapitule ces observations par rapport à l'exemple (1), retranscrit sous (2) en adoptant des conventions de

indiquent soit un passage de *creaky voice*, soit la détection d'un saut d'octave, de sorte que l'analyse de la fréquence fondamentale n'a pas pu être effectuée correctement.

<sup>2</sup> Il est presque impossible de décider quelle syllabe du mot « poème » est proéminente, ou laquelle s'intègre dans le rythme subjectivement perçu.

<sup>3</sup> Les analyses acoustiques ont été réalisées dans le logiciel Praat (Boersma & Weeninck 2005).

<sup>4</sup> On trouve un état des discussions dans Couper-Kuhlen 1993, Barbosa 1994, Auer et al. 1999.

<sup>5</sup> « On parle de rythme subjectif lorsqu'une succession régulière de sons identiques est perçue comme composée de groupes répétitifs de sons (en général 2 ou 3, parfois 4 ; Fraisse 1974) » (Padeloup 1990, 63).

transcription spécifiques<sup>6</sup>, on note que les intervalles sont compris dans ces limites (à l'exception du battement sur à *proprement parler*) qui est réalisé à un intervalle deux fois plus long que les précédents (ce qui peut s'interpréter comme l'occurrence d'un « battement silencieux », et qui n'empêche pas la poursuite du rythme) :

- (2) une anthologie euh réalisée par Liliane Wouters et Yves Namur
- |                              |                    |  |
|------------------------------|--------------------|--|
| / et PUIS                    | /                  |  |
| / mainteNANT                 | /                  |  |
| / venons-EN                  | /                  | (+1%)                                    |
| / aux PO-                    | /                  | (+16%)                                   |
| / èmes à proprement parLER   | /                  | (+100%) [int. doublé]                    |
| de Liliane WouTERS           | /                  | (+18% par rapport à la moyenne des int.) |
| et c'est un / un livre [...] | [Valibel, radLW1r] |  |

C'est donc la perception qui a guidé notre identification de rythmes réguliers (dont on fait l'hypothèse qu'ils remplissent des fonctions interactionnelles, voir ci-dessous), et pas un modèle abstrait de l'accentuation ou de la structure métrique : les syllabes proéminentes (le plus souvent réalisées avec un fort allongement de durée et portant un mouvement intonatif) ne sont pas entièrement prédictibles à partir d'un modèle de l'accentuation<sup>7</sup>, et justement nous faisons l'hypothèse que les locuteurs peuvent faire usage de l'accentuation emphatique et non emphatique (pour reprendre une distinction proposée par Di Cristo, 2000a: 28) afin de créer des suites rythmiques régulières : accent final ou initial, nucléaire ou secondaire, contrastif ou émotif, etc. L'exemple (3) est un cas extrême où une locutrice construit un patron rythmique qui va s'accélération, en accentuant chaque syllabe accentuable dans l'énoncé *on n'avait pas ça*, avec une exhaustivité qui semble contraire aux règles de l'accentuation du français<sup>8</sup>, et qui, couplée à des mouvements mélodiques importants, produit cet effet d'emphase :

- (3) / vous saVEZ /
- |                      |   |         |
|----------------------|---|---------|
| / ça ça existait PAS | / | 847 ms  |
| / il y a trente ANS  | / | 1053 ms |
| / on n'A-            | / |         |
| / -vait PAS          | / | 615 ms  |
| / ÇA                 | / | 620 ms  |
- [corpus Filliettaz<sup>9</sup>, sport]

### 3. La question de l'emphase

Les exemples (1) et (3) se caractérisent par le fait qu'une syllabe sur deux, voire chaque syllabe, est proéminente. En d'autres termes, la proportion de syllabes proéminentes par rapport aux syllabes non proéminentes augmente sensiblement dans la partie rythmiquement régulière (pour atteindre 50% ou 100% de syllabes proéminentes<sup>10</sup>), avec pour conséquence de produire un effet de scansion qui se

<sup>6</sup> Les exemples sont transcrits selon les convention VALIBEL, qu'on peut consulter à la page : <http://valibel.fltr.ucl.ac.be> dans la section « corpus oraux ». Les passages analysés pour le rythme sont notés en police Courier et chaque retour à la ligne indique l'occurrence d'une proéminence. Les symboles / ... / en début et en fin de ligne fournissent une représentation iconique des intervalles de durée entre deux proéminences, et rendent apparents les accélérations ou ralentissements du tempo. (Repris à Auer et al. 1999, ix-xi, qui présentent encore d'autres conventions.)

<sup>7</sup> Par exemple : l'accent comme réalisation de surface de grilles métriques sous-jacentes (voir Di Cristo 2000b : 37), ou lié à la structure morpho-syntaxique des énoncés (Mertens 1993).

<sup>8</sup> La locutrice est une francophone non native, germanophone. Ceci peut expliquer l'accentuation « marquée » qu'elle réalise et l'occurrence d'un « clash accentuel » : [on n'A-vait PAS] [ÇA].

<sup>9</sup> Nous remercions L. Filliettaz pour avoir mis à notre disposition des extraits du corpus Genève-2001. Une analyse détaillée de cette interaction de service se trouve dans Auchlin et al. 2004.

<sup>10</sup> Uhmann (1992) distingue la densité 1, qui correspond au « débit » de parole hors pauses (mesuré en nombre de syllabes par seconde), de la densité 2, qui correspond au nombre de syllabes accentuées par seconde et que nous désignerons par le

double d'un effet d'emphase. Pour autant, cet effet d'emphase n'est pas nécessairement associé aux structures rythmiques perçues comme régulières, comme en attestent (4) et (5).

- (4) blaAD1 et pourquoi est-ce que tu as arrêté [de travailler] ?  
 blaJD1 / PAR- /  
 / -ce que je me maRIAIS / 1094 ms  
 / et ma belle-mère était maLADe / 1124 ms (+3%)  
 / et . j'ai resté I- / 1047 ms (-7%)  
 ci  
 [corpus Valibel, blaJD1g]

Dans (4), un effet de scansion est réalisé par la récurrence régulière de proéminences accentuelles qui se suivent à des intervalles temporels perçus comme identiques. Pourtant, comme le montre le prosogramme à la figure (3), et contrairement à l'exemple (1), les proéminences ne sont pas réalisées avec des contours intonatifs particulièrement marqués (tons dynamiques amples), et le débit de la locutrice n'est pas ralenti sur le passage scandé. En outre, chaque élément fort (syllabe proéminente) est séparé du suivant par 3-6 éléments faibles (syllabes non proéminentes), ce qui crée une densité accentuelle moindre. La conjonction de ces éléments produit pour l'exemple (4) un effet de « liste<sup>11</sup> », mais pas un effet d'emphase.

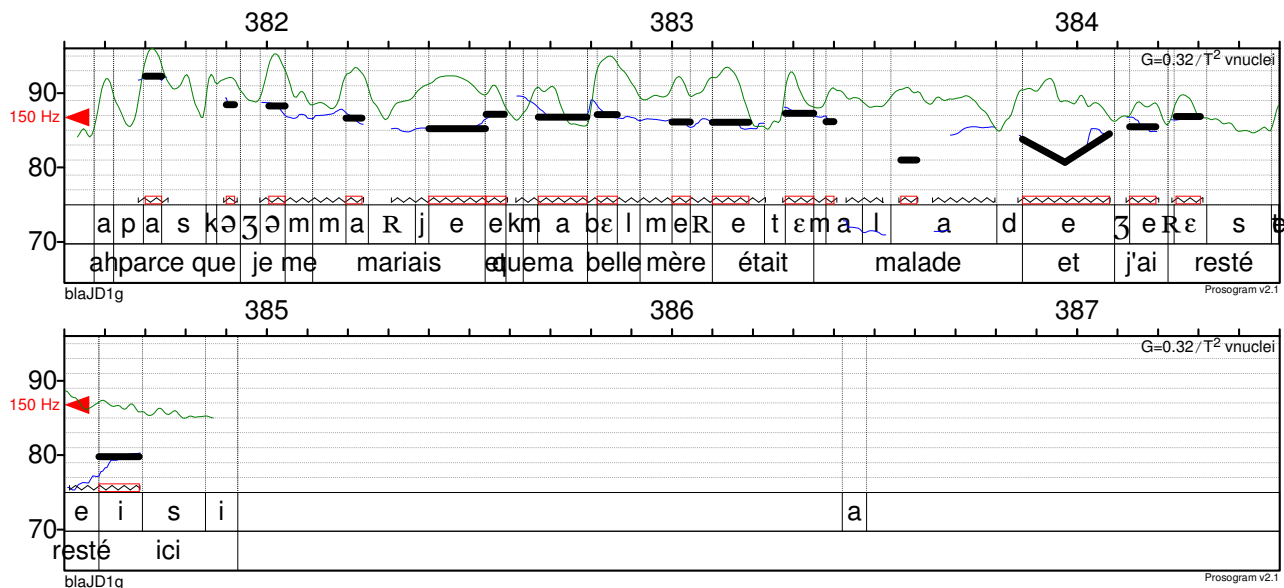


Figure 3 : Prosogramme de l'exemple (4)

Contrairement aux exemples analysés jusqu'à présent, dans (5) la structure rythmique régulière est construite par deux locuteurs, à l'occasion d'un changement de tour de parole<sup>12</sup>. Il s'agit d'un cas où on a une bonne intégration rythmique entre les tours de parole, qui peut s'interpréter comme une forme d'empathie (ou de collaboration interactionnelle réussie) entre les locuteurs en présence (voir Auchlin & Simon, 2004).

terme de « tempo ». Nous utiliserons le terme « densité (accentuelle) » pour désigner la proportion de syllabes proéminentes par rapport aux syllabes non proéminentes.

<sup>11</sup> Cet effet est bien évidemment compatible avec (et renforce) le contenu des énoncés, où la locutrice énumère une série de raisons qui l'ont poussée à abandonner son travail.

<sup>12</sup> Voir entre autres Auer 1990, Erickson 1992, Couper-Kuhlen 1993 pour des analyses de ce type de données.

(5) blaJV1	ah	oui	sinon	il	y	en	a	un	pas	loin	d'ici	.					
	/	LamberMONT	/														
	/	ils	travaillent	très	très	BIEN	/					834ms					
blaNDO	/	LAMbermont	/									802ms (-2%)					
blaJV1	/	LAMbermont	/									665ms (-17%)					
	/	c'est	une	GROSse	/							628ms (-6%)					
		société	mais	j'ai	fait	des	chantiers	avec	eux	ils	travaillent	ils	sont	très	.	très	corrects
		(silence)	[Valibel,	blaJV1,	n°4]												

Sans nier l'importance de ce phénomène dans la gestion des tours de parole, nous nous limitons aux séquences rythmiques régulières internes aux tours de parole des locuteurs, en analysant leur rôle dans la structuration des tours complexes (par exemple lors d'une narration ou du développement d'une argumentation) ou leur fonction stylistique.

Pour résumer cette première approche du phénomène, nous distinguons différents types de patrons rythmiques réguliers en fonction des critères suivants : (1) le débit moyen des énoncés (mesuré à l'intérieur des groupes intonatifs en nombre de syllabes par seconde, hors des pauses silencieuses) ; (2) le tempo (mesuré par le nombre de syllabes proéminentes par seconde) ; (3) la densité accentuelle (mesurée par la proportion de syllabes proéminentes par rapport aux syllabes non proéminentes) ; (4) le caractère statique ou dynamique, ample ou restreint, des mouvements mélodiques sur les syllabes proéminentes ; (5) le caractère prédictible ou non de la réalisation des proéminences ; (6) le lexique et la syntaxe<sup>13</sup>. Selon les critères (1) à (4), nos exemples se distribuent comme suit (Tableau 1) :

	Débit	Tempo	Densité accentuelle <sup>14</sup>	Intonation	Style perçu
ex. (3)	2.4 syll./sec.	1.4 proém./sec.	60%	tons dynamiques	emphase
ex. (1)	4.8 syll./sec.	1.5 proém./sec.	40%	tons dynamiques	emphase
ex. (5) <sup>15</sup>	6.8 syll./sec.	1.9 proém./sec	26%	tons statiques	neutre
ex. (4)	5.6 syll./sec.	0.9 proém./sec.	16%	tons statiques	neutre

Tableau 1 : Interaction des paramètres dans l'effet d'emphase<sup>16</sup>

L'effet d'emphase résulte de l'interaction complexe de différents paramètres prosodiques et linguistiques (c'est pour cette raison que Selting 1994 parle de *style emphatique*). Qui plus est, cet effet est graduel, et dépend de l'engagement plus ou moins grand que manifeste le locuteur envers son propos (Lacheret 2002). On observe ainsi un continuum entre un effet d'emphase qui semble faire primer l'engagement émotionnel (voir exemple 6) et un effet de liste plus neutre du point de vue de l'implication du locuteur. Les passages qui présentent un patron rythmique régulier seront donc

<sup>13</sup> Il faudrait pouvoir ajouter à ces critères l'intensité, qui intervient certainement dans la perception de l'emphase, mais reste difficilement mesurable dans nos corpus.

<sup>14</sup> Nombre de syllabes proéminentes par rapport aux syllabes non proéminentes.

<sup>15</sup> Les valeurs ont été calculées pour chaque locuteur, puis on a fait une moyenne.

<sup>16</sup> À quoi il faudrait ajouter une mesure de la durée des syllabes proéminentes par rapport à la durée des syllabes entre deux proéminences.



interprétés comme **emphatiques** si le débit est ralenti et la densité accentuelle élevée (> 40%, c'est-à-dire qu'au moins une syllabe sur deux ou trois est proéminente, à des intervalles perçus comme identiques), et si les tons portés par les syllabes proéminentes sont particulièrement amples ou atteignent un niveau haut. Dans les cas limites de passages emphatiques, chaque syllabe d'un mot est rendue proéminente, comme dans (6).

- (6) forAD1 je dois vous dire que il y a des discussions d'ordre politique . où tout le monde s'enguirlande . et tout le monde réclame . alors là je ferme mon appareil radio . parce que vraiment → ce n'est AB-SO-LU-MENT PAS du TOUT ce que nous attendons [corpus Forum]

Le rythme n'a plus besoin d'être isochronique (intervalles temporels réguliers) pour que la scansion soit perçue, et d'ailleurs les intervalles temporels sont presque sous le seuil de perception (dans (6), ils varient de 166 à 311 ms, l'intervalle minimal perçu entre deux événements étant 150-200 ms pour des sons purs). Les scansions de ce types sont généralement limitées à un seul mot (souvent un adverbe ayant déjà un contenu sémantique liée à une attitude, comme *absolument*, *gentiment*, dans notre corpus). L'effet d'emphase se trouve renforcé quand des syllabes qui ne sont pas en fin de groupe intonatif sont proéminentes.

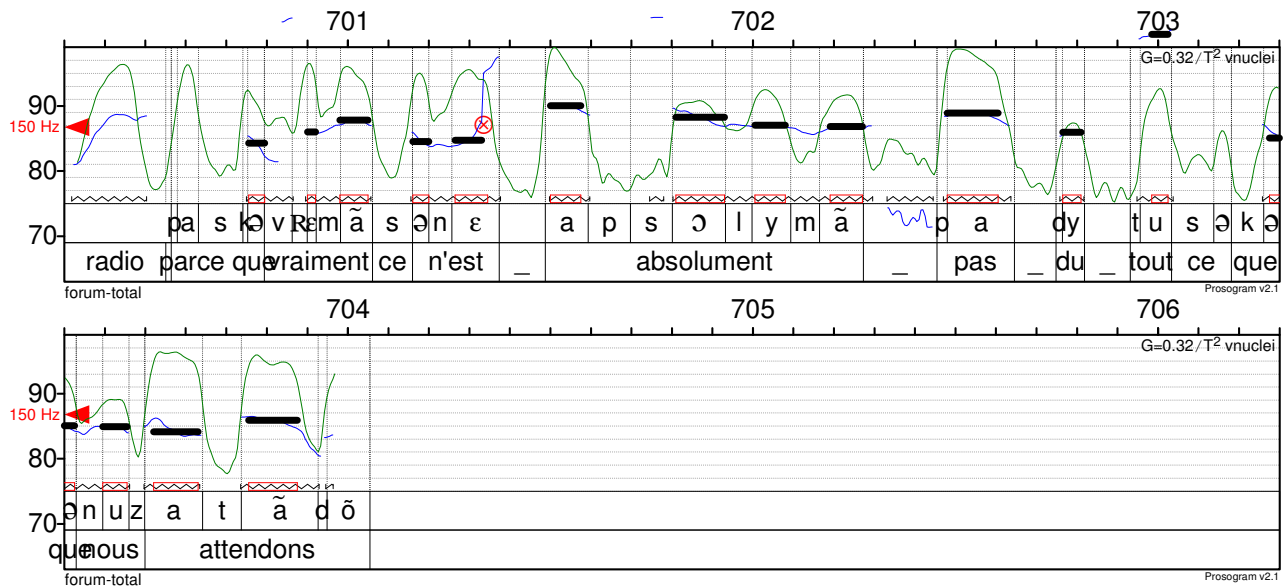


Figure 4 : Prosogramme de l'exemple (6)

Un énoncé rythmiquement régulier produit un effet de **liste** si le débit est gardé constant (sans ralentissement marqué par rapport à ce qui précède), si la densité accentuelle est moindre que dans le style emphatique (par exemple, 1 syllabe sur 5 est proéminente) et si l'intonation est moins marquée (tons statiques ou présentant des intervalles mélodiques plus réduits). Il va de soi que l'interprétation de « liste » repose nécessairement sur l'interaction du patron prosodique avec le support lexical et syntaxique, et qu'elle n'est pas compatible avec n'importe quel énoncé (voir ex. (12) infra).

Enfin, lorsqu'une régularité rythmique est réalisée conjointement par deux locuteurs à la jonction d'un tour de parole et qu'elle ne porte aucune marque particulière d'emphase (intonative, de durée), on l'interprétera comme l'indice d'une **intégration** réussie, une transition non problématique entre tours.

## 4. Cadrage et « position »

Certains contours intonatifs disposent d'un contenu sémantique minimal à partir duquel on peut les interpréter, en relation avec les structures syntaxiques qui les portent (Delattre 1966, Rossi 1999, Mertens *dans ce volume*). Il n'en va pas de même pour les phénomènes comme les variations de tempo de débit, ni pour les scansion rythmiques. Ces dernières peuvent être caractérisées par un certain degré de saillance qui les détache du discours qui précède et qui suit, installant un contraste. Mais elles ne disposent pas d'un potentiel fixe de signification qui serait invariable. Elles fonctionnent plutôt comme des indices de contextualisation (Gumperz 1982, Auer & di Luzio 1992), c'est-à-dire des signes indexicaux qui sont interprétables en fonction de la co-occurrence d'indices présents dans différents « canaux » de la communication (lexique, prosodie, gestualité) et de leur rôle dans la création d'un contexte en lien avec leur occurrence. En d'autres mots :

Du point de vue fonctionnel, l'usage de la parole spontanée met en jeu [...] un faisceau de fonctions qui participent à la fois d'une sémiotique à *géométrie variable*, dans la mesure où un même signal prosodique peut donner lieu à une interprétation sémantique différente selon son contexte d'intégration, et d'une sémiotique *synchrétique*, étant entendu que ce même signal propage simultanément des informations de nature très diverse (Di Cristo 2000b : 197-198).

Comme l'explique Gumperz (1992, 232), les indices de contextualisation « servent à mettre en évidence, à mettre à l'avant-plan ou rendre saillantes certaines séquences phonologiques ou lexicales par rapport à d'autres unités similaires, c'est-à-dire qu'ils fonctionnent de manière relationnelle (relative) et qu'on ne peut pas leur assigner des significations indépendantes du contexte, avec un noyau lexical stable. Cette (procédure de) mise à l'avant-plan, en outre, ne repose pas sur un seul indice ; au contraire, ces hypothèses dépendent de la co-occurrence des jugements qui évaluent simultanément une variété d'indices différents. Quand elle est interprétée en référence à des connaissances lexicales et grammaticales, à la position structurale dans une proposition et à la localisation séquentielle dans un fragment de discours, la mise à l'avant-plan devient un donné (input) pour faire des implicatures, produisant des interprétations situées » (notre traduction).

Dans la ligne de Gumperz, nous faisons l'**hypothèse** que les scansion s'interprètent relativement à, et contribuent à créer, le contexte d'occurrence des énoncés, et qu'elles permettent au mieux de faire des inférences interprétatives qui devront être confirmées (par la réaction de l'interlocuteur, par des indices d'autres niveaux sémiotiques). En tout état de cause, il faut sortir du cadre de l'énoncé pour interpréter les scansion, qui ne deviennent porteuses de sens que par rapport à des unités de niveau supérieur comme :

- les **topiques** discursifs globaux développés dans les grandes phases d'une rencontre interactionnelle (Grobet 2002, Chafe 1994), ou les différents **foyers** d'attention émergeant successivement ou conjointement dans une interaction (Goffman 1974, Filliettaz 2002).
- les **actions** réalisées par les locuteurs : le discours est construit, et interprété, non seulement en termes d'unités informationnelles, mais aussi en fonction des *actions* des locuteurs : « There is virtually always an issue (for the *participants* and, accordingly, for professional analysts) of what is getting done by its production in some particular here-and-now » (Schegloff 1995, 187).
- les **positions** (*footing*) adoptées par les locuteurs dans leurs discours : derrière le locuteur (ou l'animateur dans la terminologie de Goffman 1981<sup>17</sup>) qui produit les énonciations, peuvent se

---

<sup>17</sup> Notions comparables chez Ducrot 1984 (*Le dire et le dit*).



révéler différents « auteurs » (celui qui a choisi les sentiments exprimés et les mots pour les encoder, Goffman 1987, 154) ou « responsables » (le rôle social endossé en relation aux mots énoncés, *ibid.*).

- le **style** de parole adopté par un locuteur, qui peut se situer entre les pôles standard et informel de son répertoire, voire signaler une posture emphatique (Selting 1989), qui peut connoter certaines attitudes ou émotions (Fónagy 1971) ou renvoyer à certains stéréotypes (régionaux, de métiers, etc.) (Léon 1993).

Dans le reste de cet article, nous analysons plusieurs exemples de scansions pour démontrer que leur rôle dans l'interaction, loin d'être unique, est fonction des orientations des interlocuteurs et des cadres d'interprétation qu'ils convoquent pour leur interaction (voir aussi Auchlin et al. 2004, Auchlin & Simon 2004, Simon & Ronveaux 2005).

## 5. Style vocal et type d'activité

Le premier extrait « met en scène » une cliente et une vendeuse dans un magasin de bricolage. La cliente est vraisemblablement en quête d'une boussole, et elle insiste sur le fait que cette boussole doit être grande (*pas un tout petit machin*). La vendeuse lui répond qu'ils ne vendent pas ce genre d'accessoire et la renvoie vers un autre magasin (M-électronic) dans le même centre commercial. La cliente s'apprête à prendre congé, mais se ravise au dernier moment :

- (7) C euh une boussole est-ce que vous avez une boussole  
 [...]
   
V euh: allez voir à tout hasard chez M-Electronic alors en-dessous du sport
   
C ouais d'accord je vais aller voir . alors ouais
   
→ parce que mon mari il n'a toujours pas
   
/ c o m p r i s /
   
/ q u e / 891ms
   
/ LA /
   
/ bousSOLE / 426ms [tempo doublé]
   
/ ELle / 448ms (+5%)
   
/ nous DONne / 334ms (-25%)
   
/ le NORD / 406ms (+21%)
   
[...]
   
alors euh:: . faut que je lui . fasse une petite leçon alors sur le petit machin c'est pas évident
   
→ (rire) j'aimeRAIS une BELle bousSOLE

Alors qu'on s'attend à un échange de salutations qui mettrait un terme à la rencontre, la cliente débute un tour de parole avec un connecteur (*parce que*) qui projette une justification ou une cause. En fait, elle explique que son mari n'a pas compris le fonctionnement de la boussole, qu'elle devrait donc *lui faire une petite leçon*, et que pour cela elle a besoin d'une *belle boussole*. L'intérêt du passage réside entre autres dans le fait que cette intervention est réalisée sous la forme d'une scansion rythmique particulièrement saillante (par le tempo rapide, la densité accentuelle élevée et les contours intonatifs dynamiques).

D'une part, la cliente adopte un style emphatique (Selting 1994) et crée par la même un **contraste** très marqué avec le style (neutre) utilisé pour formuler sa demande de renseignement à la vendeuse. D'autre part, on se trouve face à un enchaînement non préféré au niveau textuel : la vendeuse produit un acte de conseil (*allez voir au M-électronic alors*) qui oriente la rencontre vers une clôture et semble

appeler une séquence d'acceptation de l'offre par la cliente (*OK je vais y aller*), éventuellement suivie d'un remerciement (*merci pour votre conseil*) et d'une salutation de clôture (*au revoir*). Au lieu de cela, la cliente produit une justification de sa requête initiale. Par le style qu'elle adopte, la cliente indique qu'elle s'oriente vers une autre relation que la stricte relation vendeuse-cliente, et qu'elle sélectionne pour elle-même une position plus proche de la conversation informelle. Elle modifie donc le rôle social dans lequel elle agit. Corrélativement, si elle accepte ce changement de contexte, la vendeuse devient une « oreille attentive », voire un « bon public » pour écouter l'histoire de cette dame. La suite de l'interaction montre que la vendeuse accepte ce rôle, et que la cliente poursuit la construction d'un discours dans un contexte déplacé par rapport à celui du magasin : elle va mimer les interactions avec son mari autour du fonctionnement de la boussole. Elle va donc enchâsser dans le format de production de son discours d'autres auteurs (en rapportant les paroles de son mari ou ses paroles dans un autre temps-lieu) et d'autres responsables (elle-en-tant-qu'épouse) – et ces altérations du contexte initial de l'énonciation sont signalées entre autres par les scansions qui *déplacent* effectivement les rôles et le contexte.

Un point commun entre les scansions réside dans l'impression d'effort qu'elles rendent public : il s'agit d'une forme d'hyper-articulation (Linblöm 1990) qui le plus souvent entre en contraste avec un discours plus neutre. Selon la co-occurrence d'indices (tempo, densité accentuelle, contours mélodiques, contenu lexical), le style résultant est perçu comme plus ou moins emphatique. Certaines scansions sont utilisées, dans certains contextes, avec un sens très précis, qui peut finir par faire partie du style vocal lié à une fonction ou à un métier. L'exemple (8) est extrait d'une interaction en classe<sup>18</sup>, dans laquelle les élèves sont en train de corriger un questionnaire qu'ils ont dû remplir, à propos d'une fable de La Fontaine (voir Simon & Ronveaux 2005 pour une description complète du cadre d'interaction). La routine dans laquelle l'enseignante et les élèves sont impliqués ne varie pas : (1) l'enseignante sélectionne / désigne un élève pour lire une question (dans l'ordre où elles apparaissent sur le questionnaire) et proposer sa réponse ; (2) l'élève s'exprime ensuite, et l'enseignante (3a) ratifie sa réponse, ou (3b) sélectionne un autre élève pour fournir un complément ou une nuance ; (4) si elle estime que la réponse est encore incomplète, ou semble mal comprise, elle (5) développe elle-même une explication ; finalement, (6) elle dicte ce qui selon elle constitue une réponse satisfaisante pouvant figurer par écrit sur le questionnaire de chaque élève, et (1') passe à la question suivante. L'exemple (8) montre l'enseignante proposer elle-même une explication sur un point qu'elle juge compliqué (étape 5 de la routine), puis elle dicte la « bonne » réponse (étape 6) :

- (8) Ens            donc ça c'est une manière de mettre l'interlocuteur dans/ dans sa poche de façon très habile d'accord donc ça c'est c'est euh c/ j'aimerais bien que vous le notiez ça notez que ça . ça implique l'interlocuteur
- |                    |                        |
|--------------------|------------------------|
| / ET               | /                      |
| / en quelque SORTe | / 570ms                |
| / ça le FAIT       | / 602ms                |
| / prisonnIER       | / 680ms                |
| / de votre AR      | / 631ms                |
| / gumentaTION      | / 660ms [corpus GRAFE] |

Au long des quatre heures de cours de cette enseignante que nous avons analysées, chaque passage dicté pour la prise de note est inévitablement accompagné d'une scansion, et la consigne explicite de prendre note (*j'aimerais que vous notiez ça*) est rarement présente. La scansion (qui est ici relativement peu emphatique) semble donc dédiée à une fonction précise dans le contexte de classe : elle indique aux élèves quelle activité (la prise de note) ils doivent réaliser, parmi les différentes activités possibles

<sup>18</sup> Corpus recueilli par le GRAFE, Université de Genève, et gracieusement mis à notre disposition par C. Ronveaux.

(écouter, comprendre, noter, ...). Il s'agit donc d'un dispositif précis, qui joue un rôle non ambigu dans une routine particulière, et sur laquelle les élèves peuvent se fonder pour distinguer parmi la masse des paroles prononcées par l'enseignante, celles qui doivent se retrouver fixées par écrit. La scansion facilite bien sûr le processus de prise de note, puisqu'elle réalise conjointement un ralentissement du débit (qui permet de noter au fur et à mesure les mots qu'on entend), une segmentation du flux informationnel en très petites unités informationnelles (les groupes portant une proéminence finale), et une perception aisée du discours grâce à une hyperarticulation. Il y a donc un lien iconique très fort entre la forme du discours et ce à quoi il doit servir – et c'est dans ce couplage que s'ancre un des stéréotypes vocaux les plus puissants : le style « maîtresse d'école ».

Hyperarticulation has a variety of roles in English. At the bottom line, it can sound emphatic. Emphasis, in turn, can enhance clarity, and it can communicate heightened emotion. It is also associated with the written word and a school-teachery standard – with meticulousness hence propriety. This is a bundle of meanings in which one might say that the natural blends with the conventional. Which aspect of these meanings gets attached to a particular pattern of hyperarticulation appears to depend on the style in which it is implemented. (Eckert 2005, 27 ; nous soulignons)

Ce style réalisé par une scansion non emphatique peut servir à contextualiser n'importe quelle activité qui consiste à « donner une leçon à quelqu'un », même en dehors du contexte scolaire. Un dernier exemple permettra d'illustrer ce cas de figure : à la fin d'un repas familial, qui réunit le père (P), son fils (F), et la compagne du père (C), on commence à parler des petits messages<sup>19</sup> qu'on envoie via les téléphones portables. La conversation a abouti à ce sujet parce que le père a reçu un message de sa fille (absente) et qu'il tente d'y répondre. Il s'y prend visiblement maladroitement (à cause de son manque d'expérience) et demande à son fils de lui expliquer comment on fait pour écrire en majuscules :

- (9) C: dis-moi ce que tu veux lui mettre et {on l'écrit} (rire)  
 F: non il apprend c'est bien  
 (silence)  
 → P: et en majuscules comment on fait  
 F: / majuscule /  
 / si tu as fait un POINT / 903ms  
 / tu fais un esPACE / 747ms (-20%)  
 / et il te mettra d'office une majusCULE / 1284ms  
 sinon .  
 / moi j'ai ce signe LÀ /  
 / une flèche vers le HAUT / 1043ms  
 / je sais pas si tu as la MÊME / 1059ms (+1%)  
 P: non  
 F: là en bas... [corpus Valibel, famES1r]

La réponse du fils est prononcée avec une scansion relativement peu emphatique (les proéminences ne sont pas en position d'accent initial mais final de groupe intonatif). La scansion n'est pas strictement isochronique, mais le manque de régularité dans la durée des intervalles entre deux proéminences est compensé par une syntaxe qui instaure une régularité et une prédictibilité (*si... alors..., sinon...*) et par un lexique relevant du même champ (*point, espace, majuscule...*). La co-occurrence d'indices de différents niveaux augmente la tolérance à des écarts temporels.

Du point de vue interprétatif, la scansion est à nouveau un indice qui permet aux locuteurs de recadrer l'activité en cours et de modifier légèrement les rôles sociaux des interactants. En adoptant ce

<sup>19</sup> Appelés « textos » ou « sms ».

ton qui n'est pas étranger à celui d'une maîtresse d'école, le fils se permet d'une certaine manière de doublement marquer (par le contenu de son discours, et par sa forme) qu'il donne une petite leçon à son père sur l'usage d'un objet technique que celui-ci maîtrise peu (le téléphone portable). Dans la situation de communication, où la compagne du père plaisantait en disant *dis-moi ce que tu veux lui mettre et on l'écrit* (à ta place), ce ton peut pointer vers une nuance de condescendance (bienveillante) du fils qui apprend quelque chose à son père. À nouveau, les rôles sociaux des différents interactants sont en jeu (comme dans l'interaction entre la vendeuse et la cliente) – et les indices prosodiques permettent aux locuteurs, de manière éphémère et locale, à jouer subtilement entre différents rôles sans que rien d'explicite ne soit énoncé. C'est le cadre de participation (à savoir le rôle social, le statut, de chaque locuteur, Goffman 1981) qui est remodelé, et cette modification du contexte est rendue visible entre autres via les scansions.

## 6. Scansions et formes d'organisation discursive

Les scansions rythmiques, plus ou moins emphatiques, permettent de contextualiser différentes formes d'organisations du discours (Roulet et al. 2001) : elles peuvent également marquer les étapes principales d'un récit (dénouement, chute) dans le cadre de l'organisation compositionnelle (Selting 1992), contribuer au marquage du discours rapporté dans le cadre de l'organisation polyphonique, annoncer un topique global au niveau de l'organisation topicale (voir exemple 1), ou encore, construire la structure argumentative en mettant en relief les éléments hiérarchiquement principaux. Sans prétendre à l'exhaustivité, nous examinerons ici quelques exemples représentatifs.

Le plus souvent, une scansion caractérise, au sein d'une intervention, le point sur lequel le locuteur souhaite attirer l'attention de ses interlocuteurs. Il en va ainsi dans l'exemple suivant, extrait d'un repas pris en famille (le père, le fils et la compagne du père) et où la discussion tourne autour des logements étudiants<sup>20</sup> :

- (10) F: ouais enfin c'est comme ceux qui / qui habitent ici qui étudient avec moi quoi /  
M: hm bè oui oui |- ils ont un kot oui  
F: ça c'est clair qu'ils ont un kot mais -|  
P:mais ils sont bien obligés  
F: oui oui |- mais  
M: ils vont -| pas prendre le train tous les jours hein  
P: mais mais franchement  
F: mais ils ne prennent pas du tout ça comme une contrainte hein // |- enfin certains peut-être  
→ P: moi je / je n/ -| je ne connais pas toutes les disciplines / et je ne sais pas quelle est la charge de travail dans chaque discipline / mais quelqu'un qui fait je ne sais pas moi la médecine ou ingénieur civil ou quelque chose comme ça //  
/ je crois PAS /  
/ qu'il a BEAUcoup / 439ms  
/ le TEMPS / 443ms (+1%)  
/ d'Aller / 481ms (+8%)  
/ se soûLER / 549ms (+14%)  
/ euh / 1297ms  
/ tous les SOIRS hein /  
F: viens un peu à Louvain-la-Neuve tu sais

<sup>20</sup> En Belgique, un *kot* est un logement loué par un étudiant sur le site de l'université où il étudie (soit une chambre individuelle, soit un appartement communautaire – un *kot commu*). Les *cercles* sont des associations étudiantes à vocation festive (et culturelle).

P: hein ?

F: viens un peu à Louvain-la-Neuve je dis

P: oui mais

F: les cercles et tout ça c'est uniquement ça hein [Valibel, famES1r]

Le passage scandé (*je crois pas / qu'il a beaucoup le temps / d'aller / se soûler / euh / tous les soirs hein*) se détache nettement du début de l'intervention par une importante variation de débit : le segment qui le précède commençant par *mais quelqu'un...* est prononcé avec un débit de 6,9 syllabes par seconde, tandis que le passage scandé a un débit de 3,9 syllabes par seconde ; ces deux passages sont par ailleurs séparés par une longue pause de 1,1 seconde. La scansion, régulière au début, est interrompue par une hésitation, et terminée par un groupe intonatif se terminant sur une syllabe accentuée qui le rattache au segment scandé qui précède.

La scansion marque ici le passage comme important à plusieurs niveaux : du point de vue de la syntaxe et de la structure informationnelle, on a une structure binaire n'impliquant certes pas de liens micro-syntaxiques, mais une structure informationnelle de type : cadre/topic - propos : *quelqu'un qui fait... je ne crois pas qu'il a beaucoup le temps...* Dans cette structure, le premier élément fournit un cadre, auquel s'applique le second élément, qui peut donc être considéré comme principal. Du point de vue de la prise de position du locuteur, l'ensemble de l'intervention est très fortement modalisé (voir la concession initiale, les répétitions de *je ne sais pas, je ne connais pas, je ne crois pas...*). Le passage scandé, bien que lui aussi modalisé puisque formulé par une forme négative (*je ne crois pas*), est toutefois celui qui exprime l'opinion à laquelle est attaché le locuteur. Cette opinion coïncide d'ailleurs avec la fin du tour de parole, puisque le locuteur n'a ensuite plus rien à ajouter. Le rythme contextualise ainsi l'élément scandé comme étant le plus important de l'intervention (alors que, non scandé il aurait pu être interprété comme un simple commentaire annonçant encore une suite).

Si, dans (10), la scansion va dans le même sens que les autres plans d'organisation du discours, dans certains cas elle peut contredire certaines informations inférables du matériel linguistique. Il en va ainsi dans l'exemple 6, que nous reproduisons ci-dessous :

- (6) forAD1 je dois vous dire que il y a des discussions d'ordre politique . où tout le monde s'enguirlande . et tout le monde réclame . alors là je ferme mon appareil radio . parce que vraiment → ce n'est AB-SO-LU-MENT PAS du TOUT ce que nous attendons [corpus Forum]

La scansion fortement emphatique liée à l'adverbe *absolument* a déjà été décrite ci-dessus. On peut ajouter ici qu'elle est doublée d'une redondance lexicale, qui souligne l'exaspération du locuteur : *vraiment, absolument, pas du tout*<sup>21</sup>. Cette scansion emphatique marque ici le point culminant ou « climax » (Selting 1994 : 389) de l'intervention de l'auditeur, dont la caractéristique est de porter sur un élément attitudinel (l'exaspération du locuteur) plutôt que sur un élément du contenu tel que le fait de fermer son appareil radio ou d'avoir des attentes. Du point de vue de l'organisation globale de l'intervention, ce point fort prosodique en vient à « renverser » en quelque sorte la structure prévue par le matériel lexico-syntaxique, puisqu'on admet généralement que le *parce que* introduit une cause ou un argument subordonné. Ce type d'exemple, qui est fréquent, illustre cette possibilité offerte par la prosodie : les locuteurs ne se contentent pas de (sur-)marquer quelque chose de préexistant, mais ils créent le contexte dans lequel doivent être interprétés leurs énoncés.

<sup>21</sup> Dans *absolument pas* et *pas du tout*, la particule *pas* est impliquée dans les deux constructions.

C'est particulièrement le cas chez des locuteurs experts, qui ont l'habitude de s'exprimer en public et qui exploitent peut-être plus encore que d'autres les aspects rythmiques de leur parole. Nous reprendrons en premier lieu un exemple de Lucchini (FL) étudié dans Simon (2004). Interrogé sur sa grande admiration pour l'écrivain, reconnu mais controversé, Céline, FL veut défendre l'idée que, malgré son comportement indéfendable au point de vue politique, Céline est un écrivain majeur. L'argument se résume donc à une concession (*Céline n'a pas réussi à assumer sa souffrance*) suivie d'un argument présenté comme principal (*mais il compte quand même comme grand écrivain*). La succession de rythmes très marqués et très différents les uns des autres est frappante, et contribue à instancier ce balancement entre deux pôles opposés. Le locuteur ne se contente pas d'utiliser une scansion pour mettre en relief une information importante, il alterne différents rythmes pour organiser son discours. On suit cette évolution à l'aide des lettres A-F qui en marquent les étapes :

- (11) @↑ ce qui m'embête c'est qu'il n'a pas réussi selon lui c'était pas bête  
 @ Céline n'a pas réussi à rester réellement désespéré réellement souffrant .
- [A]** / et il a quand même accRO / /  
 / -ché ## sa souffRANCe / / 2058ms  
 / disait Jean DaNIEL / / 1625ms (-21%)
- [B]** / # À une CASTe / / 810 (tempo doublé)  
 / À un TRUC / / 775ms (-4%)  
 / À une RACe / / 758ms (-2%)
- au lieu de l'assumer . ah c'est-à-dire d'assumer un truc tragique .
- [C]** / parce que quand on EST / /  
 / à cette hauteUR / / 749ms (-13%)  
 / de sensibilité / / 893ms (+19%)
- que traduit QUELque page du Voyage au bout de la NUIT
- [D]** / on compRENd / /  
 / que de toute éviDENCE / / 1206ms  
 / la VIE / / 1043ms (-14%)  
 / c'est comme RIM / / 1216ms (+17%)  
 / baud c'est pas des vies FA- / / 1134ms (-7%)  
 / -ciles hein |- <MP> ben NON -| / /
- [E]** / parce que le problÈMe / /  
 / qui est terrible chez CÉLINE / / 1386ms  
 / c'est qu'il a quand MÊMe / / 1475ms (+6%)  
 / DÉcomposé le morCEAU / quoi [arc accentuel: 1090ms] **[F]**  
 c'est-à-dire qu'il a # il a / /  
 / DÉplacé l'agenceMENT / [arc accentuel: 1090ms] **[F]**
- euh . euh . et qu/ donc il compte de toute évidence . et il a dit une chose sur Proust qui est tellement sublime... [Valibel, radFL1r]

Dans la succession et l'alternance des scansions, FL semble exploiter naturellement la plasticité du rythme de sa parole pour organiser son intervention et y mettre en valeur les éléments les plus importants (noyaux), d'un double point de vue informationnel et argumentatif. Le premier noyau informationnel [A-B] tourne autour du fait que Céline a projeté sa souffrance sur le monde extérieur (en s'en prenant directement aux Juifs) au lieu de l'assumer intérieurement. Il est réalisé par une scansion rythmique [A] au tempo lent<sup>22</sup> et à la densité accentuelle élevée, ce qui produit une impression de grandiloquence ou de gravité. Une scansion moins marquée et au tempo plus rapide (plus naturel) [B] contextualise la liste (*à une caste, à un truc, à une race*). D'un point de vue textuel, on a presque affaire à un seul acte de discours (une seule proposition syntaxique, excepté l'incise *disait Jean Daniel*) qui est excessivement découpé du point de vue rythmique. Ce premier passage scandé constitue par

<sup>22</sup> Les intervalles entre syllabes proéminentes tournent autour de 2 sec. Le débit n'est pas lent pour autant, mais les groupes intonatifs sont séparés par de longues pauses silencieuses.

ailleurs une séquence de discours rapporté (paroles d'un journaliste prises en charge par le locuteur), dont l'achèvement est réalisé par une accélération et par la frontière terminale.

Le contre-argument (*au lieu de l'assumer*) ne fait pas l'objet d'une scansion rythmique particulière – mais produit toutefois un effet de contraste avec la liste qui précède, de laquelle il est détaché par la présence d'une pause. La scansion rythmique du début de ce passage permet donc d'introduire et de mettre en évidence cette idée relativement complexe (la SOUFFRANCE NON ASSUMÉE), qui doit être formulée en plusieurs temps, avant de fonctionner comme **topique global** pour la suite de la séquence.

Le tour de parole se poursuit par une nouvelle alternance entre deux rythmes [**C-D**], qui produit un effet de contraste. [**C**] se développe sur un tempo rapide (débit accéléré) qui devient plus emphatique ensuite [**D**], avec un ralentissement du tempo et une augmentation de la densité accentuelle dans le second patron rythmique. On notera que même le régulateur produit par le journaliste (MP : *non*) respecte le rythme installé par FL. Cette séquence, introduisant l'idée de la sensibilité, constitue une intervention qui, globalement (voir la présence du connecteur *parce que*), enchaîne sur le topique global SOUFFRANCE NON ASSUMÉE (verbalisée par : *au lieu de l'assumer*). Le tempo rapide et à faible densité [**C**] contextualise l'ouverture d'une forme de digression, tandis que le passage emphatique [**D**], met à nouveau en évidence l'idée de SOUFFRANCE, commentée avec une comparaison des écrivains qui n'ont pas eu des vies faciles. Une comparaison des deux passages décrits comme emphatiques [**A**] et [**D**] peut suggérer l'impression que [**D**] est plus léger (peut-être à cause des effets de timbres et de l'absence de longues pauses silencieuses). Cela ne serait pas incompatible avec son statut textuel de « digression », qui vient commenter par le biais le topique global de SOUFFRANCE (NON ASSUMÉE).

La fin de l'extrait est encore scandée [**E-F**] : [**E**] annonce un **nouveau topique** (un nouveau PROBLÈME – qui résulte de l'importance de Céline au point de vue littéraire malgré ses idées politiques), puis il y a enchaînement linéaire puis à topique constant dans les actes suivants. Cette progression topicale constante est soutenue par deux arcs accentuels [**F**]. À l'intérieur du domaine focal, les deux syntagmes *décomposé le morceau* et *déplacé l'agencement* sont mis en évidence par ces arcs accentuels – une manière « d'insérer un (ou plusieurs) élément focalisé au sein d'un focus plus large » comme le décrit Di Cristo (2000a, 41). D'une certaine manière, ces deux nouveaux topiques sont poursuivis puisque FL va donner « en acte » un exemple du génie littéraire de Céline en récitant un extrait d'un de ses textes (c'est la suite de la séquence).

Le discours de FL exploite abondamment la scansion d'unités textuelles relativement importantes en alternant les tempos. Il fait preuve d'une réelle persévérance (Müller 1999, 158 in Auer et al. 1999) pour imprimer au matériel morpho-syntaxique des patterns rythmiques réguliers sur une certaine durée. Mais les scansions peuvent également intervenir au sein d'unités beaucoup plus brèves et leur effet est alors d'autant plus important. Il en va ainsi dans l'exemple suivant, produit par Georges Simenon :

(12) GS : [...] je vous répondrai **très** franchement / que / j'aimais mieux/

/ j'aime MIEUX	/	
/ être critiquÉ	/	1593ms
/ même être déteSTÉ	/	1800ms (+12%)
/ POUR	/	1393ms (-27%)
/ ce que je suis vraIMENT	/	1341ms (-4%)
/ que d'être aiMÉ	/	1957ms (battement retardé)
/ ou admirÉ	/	1480ms (+9%) (rythme récupéré)
/ pour ce qu(e) je n(e) suis PAS	/	1792ms (+21%)

[corpus Apostrophe]



Cet exemple est relativement atypique par rapport aux configurations étudiées jusqu'à présent. En effet, l'accentuation n'y est pas particulièrement perçue comme emphatique (pas de forte variation de F0 ou de l'intensité). Par contre, le tempo est marqué : les intervalles temporels entre proéminences successives sont étonnamment longs (ils atteignent même la limite supérieure du seuil jusqu'auquel on perçoit une régularité), et l'effet d'isochronie est renforcé par la construction syntaxique qui instaure elle aussi un parallélisme. Du point de vue de l'organisation textuelle, cette segmentation prosodique a pour effet de scinder le flux discursif en unités minimales plus petites que ce qu'elles auraient été sans la présence de la scansion. Par là, les propos qu'elles véhiculent acquièrent une certaine autonomie (voir la règle de « one new idea per intonation unit » établie par Chafe 1994), et donc un poids informationnel supérieur à celui qu'ils auraient eu si le locuteur avait formulé sa pensée d'un trait. Un tel ralentissement du débit produit donc lui aussi un effet de mise en relief de l'information, bien que différent de celui de l'accent emphatique. Il a également des effets du point de vue du marquage de l'attitude du locuteur, puisqu'il témoigne de l'importance que ce dernier accorde à chacun de ses mots ainsi que du caractère hautement « surveillé » de son discours. De ce point de vue, cet exemple illustre en quelque sorte « l'opposé » de (6), dans lequel le locuteur se laissait guider par son exaspération, au détriment de la rectitude de la construction syntaxique.

## 7. Conclusion

Dans la première partie de cet article (sections 1, 2 et 3) nous avons décrit les formes que peuvent prendre les scansions rythmiques, définies comme des séquences saillantes du point de vue de leur organisation rythmique régulière, et caractérisées par :

- la récurrence de syllabes proéminentes à des intervalles perçus comme identiques – ces syllabes pouvant être en position d'accent final ou non final (isochronie définie par Auer et al. 1999) ;
- l'augmentation de la densité accentuelle (nombre de syllabes proéminentes par rapport aux syllabes non proéminentes) (voir Uhmman 1992) ;
- souvent, mais pas toujours, un ralentissement du débit de parole.

Ces scansions développent un caractère emphatique quand certains indices sont présents, le caractère (non) emphatique d'une scansion étant une affaire de degré :

- un tempo (nombre de syllabes proéminentes par seconde) relativement lent ;
- des contours intonatifs marqués (dynamiques, avec des valeurs de F0 relativement hautes par rapport aux valeurs qui précèdent) ;
- un contenu lexical dénotant des attitudes ou des émotions du locuteur (voir Selting 1994).

En tant qu'indices de contextualisation (Gumperz 1992), les scansions ne disposent pas a priori d'un « potentiel de signification » stable qui se réalise identiquement dans n'importe quelle situation. Au minimum, une scansion permet de signaler le caractère « marqué » de la séquence, ou d'indiquer que quelque chose change, comme par exemple un aspect du contexte invoqué pour interpréter correctement les énoncés. Le contexte n'est en aucune manière donné – et les scansions, parmi d'autres indices, permettent aux locuteurs de rendre accessible aux interlocuteurs certains aspects du contexte qui deviennent pertinents comme : le rôle avec lequel ils prennent la parole, l'activité dans laquelle ils sont engagés, les aspects de leurs discours qui sont au premier vs. à l'arrière-plan, etc. L'étude du contexte ne précède pas, mais commence avec, l'étude des indices de contextualisation :

If there are indefinitely many potentially relevant aspects of context and of personal or categorical identity which could have a bearing on some facet of, or occurrence in, interaction, and if the analyst must be

concerned with what is relevant to the parties at the moment at which what is being analyzed occurred, and is procedurally consequential for what is being analyzed, then **the search for context properly begins with the talk or other conduct being analyzed**. That talk or conduct, or what immediately surrounds it, may be understood as displaying which out of that potential infinity of contexts and identities should be treated as relevant and consequential (both by co-participants and by professional analysts). (Schegloff 1992, 197)

À travers l'analyse d'exemples issus d'un large éventail de situations de communication formelles et informelles (interactions de service, interview médiatiques, interactions en classes, conversations informelles), nous avons montré une série de fonctions des scansions dans la communication pour : souligner l'argument principal dans une longue séquence argumentative, signaler l'engagement émotif d'un locuteur dans son discours, indiquer un changement d'activité en classe, modifier le rôle social à partir duquel on interagit avec quelqu'un, etc. Les scansions constituent en tout cas une manière non directive, et non explicite, de réaliser ces fonctions – et elles semblent plus fréquentes dans les discours des experts de la langue (enseignants, écrivains, journalistes) que dans les interactions quotidiennes.

## Références bibliographiques

AUCHLIN A. & SIMON A.C. (2004), Gabarits prosodiques, empathie(s) et attitudes, *Cahiers de l'Institut de Linguistique de Louvain* 30/1-3, 181-206.

AUCHLIN A., FILLIETTAZ L., GROBET A. & SIMON A.C. (2004), Énaction, expérience du discours et prosodie, *Cahiers de Linguistique française* 26, 217-249.

AUER P. (1990), Rhythm in telephone closings, *Human Studies* 13, 361-392.

AUER P. & DI LUZIO A. (eds) (1992), *The Contextualization of Language*, Amsterdam, Benjamins.

AUER P., COUPER-KULHEN E. & MÜLLER F. (1999), *Language in Time. The Rhythm and Tempo of Spoken Interaction*, New York, Oxford, Oxford University Press.

BARBOSA P. (1994), Caractérisation et génération automatique de la structuration rythmique du français. Thèse de Doctorat Spécialité Sciences Cognitives, Grenoble (France), Institut National Polytechnique de Grenoble, 194 p. (Disponible à l'adresse : <http://www.icp.inpg.fr/~bailly/theses.html>)

BOERSMA P. & WEENINK D. (2005), Praat: doing phonetics by computer (Version 4.3) [Computer program]. Retrieved July 1, 2005, from <http://www.praat.org/>

CHAFE W. (1994), *Discourse, consciousness and time*, Chicago, The University of Chicago Press.

COUPER-KUHLEN E. (1993) *English Speech Rhythm: Form and function in everyday verbal interaction*, Amsterdam, Benjamins.

DI CRISTO A. (1999), Vers une modélisation de l'accentuation du français (première partie), *French Language Studies* 9, 143-179.

DI CRISTO A. (2000a), Vers une modélisation de l'accentuation du français (deuxième partie), *French Language Studies* 10, 27-44.

DI CRISTO A. (2000b), La problématique de la prosodie dans l'étude de la parole spontanée, *Revue PArole* 15/16, 189-250.

- DI CRISTO A., AURAN C., BERTRAND R., CHANET C., PORTES C. & REGNIER A. (2004), Outils prosodiques et analyse du discours, *Cahiers de l'Institut de Linguistique de Louvain* 30/1-3, 27-84.
- DELATTRE P. (1966), Les dix intonations de base du français, *French Review* 40, 1-14.
- DURANTI A. & GOODWIN C. (1992), *Rethinking context. Language as an interactive phenomenon*, Cambridge, Cambridge University Press.
- ECKERT P. (2005), Variation, convention and social meaning. Paper presented at the Annual Meeting of the Linguistic Society of America. Oakland CA. Jan. 7, 2005. 33 p.
- ERICKSON F. (1992) They know all the lines, in Auer P. & Di Luzio A. (eds), *The contextualization of language*, Amsterdam, Philadelphia: John Benjamin, 365-397.
- FILLIETTAZ L. (2002), *La parole en action*, Québec, Nota bene.
- FÓNAGY I. (1971), The functions of vocal style, in Chatman S. (ed), *Literary style: a symposium*, London, New York, Oxford University Press, 159-174.
- GOFFMAN E. (1974). *Frame analysis. An Essay on the Organization of Experience*. Reprint 1986: York, Northeastern University Press.
- GOFFMAN E. (1987) [1981], La position, in *Façons de parler*, Paris, Éditions de Minuit, 133-166.
- GOODWIN C. (2000), Action and embodiment within situated human interaction, *Journal of Pragmatics* 32, 1489-1522.
- GROBET A. (2002), *L'identification des topiques dans les dialogues*, Bruxelles, De Boeck-Duculot.
- GUMPERZ J.J. (1982), *Discourse Strategies*, Cambridge Mass., Cambridge University Press.
- GUMPERZ J.J. (1992), Contextualization and understanding, in Duranti A. & Goodwin C. (eds), *Rethinking context. Language as an interactive phenomenon*, Cambridge, Cambridge University Press, 229-252.
- HOEQUIST C. (1983), The perceptual center and rhythm categories, *Language and Speech* 26/4, 367-376.
- LABOV W. (1972), *Language in the Inner City*, Philadelphia, U. of Pennsylvania Press. French translation: (1978), *Le Parler Ordinaire*, Paris, Éditions de Minuit.
- LACHERET A. (2002), *Modélisation prosodique du français parlé. Analyse de la substance, Représentation formelle, Interprétation symbolique*, Rapport d'habilitation, Paris X, Nanterre, cd-rom.
- LEON P. (1993), *Précis de phonostylistique. Parole et expressivité*, Paris, Nathan.
- LINDBLÖM B. (1990), Explaining phonetic variation: A sketch of the H&H theory, in Hardcastle W.J. and Marchal A. (eds), *Speech Production and Speech Modelling*, The Netherlands, Kluwer Academic Publishers, 403-439.

- MERTENS P. (1993), Accentuation, intonation et morphosyntaxe, *Travaux de Linguistique* 26, 21-69.
- MERTENS P. (2004), Le prosogramme : une transcription semi-automatique de la prosodie, *Cahiers de l'Institut de Linguistique de Louvain* 30/1-3, 7-25.
- MERTENS P. (à paraître), Syntaxe, prosodie et structure informationnelle, une approche prédictive pour l'analyse de l'intonation dans le discours, *dans ce volume*.
- PASDELOUP V. (1990), Modèles de règles rythmiques du français appliqué à la synthèse de la parole, Thèse de doctorat nouveau régime, Université de Provence, Aix-Marseille I, 386 p.
- ROSSI M. (1999), *L'intonation, le système du français : description et modélisation*, Paris, Gap, Ophrys.
- ROULET E., FILLIETTAZ L., GROBET A. & BURGER M. (2001), *Un modèle et un instrument d'analyse de l'organisation du discours*, Bern, Peter Lang.
- SCHEGLOFF E. (1992), In another context, in Duranti A. & Goodwin C. (eds), *Rethinking context. Language as an interactive phenomenon*, Cambridge, Cambridge University Press, 191-227.
- SCHEGLOFF E. (1995), Discourse as an Interactional Achievement III: The Omnirelevance of Action, *Research on Language and Social Interaction* 28/3, 185-211.
- SELTING M. (1989), Speech styles in conversation as an interactive achievement, in Hickey L. (ed), *The pragmatics of style*, London, New York, Routledge, 106-130.
- SELTING M. (1992), Intonation as a Contextualization Device: Case Studies on the Role of Prosody, Especially Intonation, in Contextualizing Story Telling in Conversation, in Auer P. & Di Luzio A. (eds), *The contextualization of language*, Amsterdam, Philadelphia: John Benjamin, 233-258.
- 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. & RONVEAUX C. (2005). La gestualité prosodique au service de l'objet enseigné. Communication présentée au colloque *Interacting Bodies*, École nationale supérieure, Lyon, 17 juin 2005.
- UHMANN S. (1992), Contextualizing relevance: on some forms and functions of speech rate changes in everyday conversation, in Auer P. & Di Luzio A. (eds), *The contextualization of language*, Amsterdam, Philadelphia: John Benjamin, 297-336.
- ZELLNER B. (1998), *Caractérisation et Prédiction du Débit de Parole en Français. Une étude de cas*, Thèse de Doctorat. Faculté des Lettres, Université de Lausanne.