

CHAPITRE XVII

PROSODIE ET LINGUISTIQUE

1. RAPPEL DE QUELQUES CARACTERISTIQUES DE L'ETUDE

Nous avons donc tenté, dans les limites restreintes de notre expérimentation, de dresser un bilan des relations entre les domaines linguistiques (syntaxe, sémantique et pragmatique) et les indices de la prosodie issus de la fréquence fondamentale, de l'énergie et de la durée. Les modèles ont été définis de telle manière que traitant tous les mots lexicaux, ils décrivent les mécanismes linguistiques fondamentaux mis en oeuvre dans un texte.

Précisons cependant que ce n'est pas parce que les connaissances linguistiques que nous avons formalisées sont vraisemblablement applicables à tout texte, que l'exploitation que nous en faisons pour nos énoncés nous autorise à proposer un modèle général du français lu. Tout au plus propose-t-on un modèle prosodique qui soit représentatif *de ce texte, de ces conditions expérimentales, et de nos locuteurs*. Ce modèle prosodique repose sur le principe général du traitement de l'information saillante, de nature essentiellement sémantique et/ou pragmatique, et ce, dans plusieurs de leurs dimensions.

Dans ces conditions il faut préciser le contenu que l'on accorde à l'expression "modèle prédictif". Etant donné que tout texte est interprétable de manière différente, et que le comportement cognitif du locuteur n'est pas (et ne peut pas être) entièrement prédictible, ces modèles sont en fait potentiellement prédictifs avant toute application à l'énoncé : pour chaque GM, parmi une gamme de 6 interprétations concurrentes à destination d'un auditeur, un de ces modèles rend le mieux compte des valeurs mélodiques. En ce sens on peut dire que ce modèle a "prédit" les valeurs mélodiques, parce que les valeurs qu'il soumet sont en moyenne pour 85% d'entre elles, conformes aux valeurs mélodiques. Par suite lorsque l'on retrace la suite de modèles utilisés dans le cours de l'énoncé d'un locuteur, nous avons bien affaire à des modèles prédictifs car ils expliquent, à défaut de toutes les valeurs (ce qui arrive parfois cependant), une grande majorité de celles-ci : nous réalisons ainsi une *analyse* des énoncés *par prédiction*.

2. RECAPITULATION DES PRINCIPAUX ACQUIS

Si l'on regarde globalement (cf ci-dessous le tableau 17-1) la distribution des modèles dans chaque consigne par ordre décroissant d'effectifs pour les modèles les mieux représentés, tous locuteurs confondus, toutes phrases confondues, on constate :

- que la consigne 1 au débit le plus rapide (favorisant donc une perspective moins analytique), compte le plus grand nombre de GM sous la dépendance des modèles structuraux et holistiques, EN+ER, HR,
- que la consigne 2 ne comporte plus qu'un seul modèle de ce type (EN + ER, le modèle HR étant marginalisé), alors que les modèles analytiques (CP + CM) gagnent en représentation,
- et que finalement la consigne 3 poursuit cette tendance jusqu'à la claire prédominance des modèles analytiques.

Consigne 1	EN+ER 38%	CP 28%	HR 16%	Holistique 54%	Analytique 28%
Consigne 2	EN+ER 40%	CP 30%	CM 13%	40%	43%
Consigne 3	EN+ER 35%	CP 23%	CM 23%	35%	46%

Tableau 17-1

Distribution des modèles les mieux représentés à travers les consignes.

Il en va donc dans notre expérimentation pour les phrases comme pour les consignes, les modèles holistiques et structuraux précèdent les analytiques. Ainsi la parole dans notre expérimentation, n'est pas régie de manière différente des autres processus humains, comme par exemple la perception visuelle dans le domaine de la théorie attentionnelle (Mar, 1976 ; Neisser, 1967 ; Treisman and Gelade, 1980) ou encore les mécanismes d'apprentissage, par exemple en lecture (Frith, 1980 ; Perfetti and Roth, 1981 ; Warrington and Shallice, 1980 ; Sprenger-Charolles, 1989) : *en perception visuelle, en apprentissage de lecture, en compréhension comme dans notre cas en production, la phase d'analyse avec ses traitements plus fins et spécialisés, est précédée par une phase de saisie globale de l'objet ou de la tâche.*

3. L'EXPRESSION DE LA SIGNIFICATION

3.1. PROSODIE ET COMMUNICATION

Dans notre interprétation des faits, la prosodie dans sa fonction générale, est de nature *pragmatique*, dans la mesure où elle opère un véritable *prédécodage des signifiés*, à destination d'un auditeur, dans le cadre d'une situation de communication précise, et ce, grâce à des moyens linguistiques et pragmatiques, mais aussi subjectifs, liés à un comportement.

La prosodie de lecture serait ainsi le produit d'une sorte de transaction, d'une "négociation" de la part du locuteur entre les contenus d'un texte (et son ensemble d'interprétations virtuelles), la perception des contraintes de situation, et par ailleurs, ses intentions ou motivations. A l'issue de cette transaction, processus inconscient, bien entendu, et dynamique au cours de la lecture, une sélection du sens (i.e. une restriction de ses virtualités), serait effectuée par les indices prosodiques, marqueurs de la signification.

C'est justement dans la mesure où les structures syntaxiques et sémantiques sont cooccurentes, avec leurs différents réseaux de signification, que la prosodie,— et plus particulièrement les indices de F0 —, peut restreindre le sens du discours à un type d'argument qui dans notre expérimentation, est de nature essentiellement sémantique (et pragmatique).

Plus encore, si la fonction qu'assume la mélodie se résumait à l'expression des relations syntaxiques, elle serait purement redondante et dans ce cas elle ne fournirait aucune information sur l'interprétation que le locuteur opère sur le discours. C'est d'ailleurs une possibilité réelle que le locuteur exploite lorsque le discours est sémantiquement peu saillant, ou lorsque le locuteur ne peut ou ne veut pas s'investir dans son discours¹, mais c'est loin d'être la règle universelle. A la fin des années 70, mes premiers résultats (Caelen, 1978, 1981) qui se positionnaient à rebours de la plupart des travaux s'orientant vers une explication syntaxique de la mélodie, montraient

¹ Ce qui est d'ailleurs favorisé lorsque les corpus sont constitués de phrases isolées.

l'importance des facteurs sémantiques. Depuis nombre de travaux ont peu à peu recensé tous les facteurs qui remettent en cause ce principe de congruence de la mélodie avec la syntaxe (voir à ce sujet le chapitre I).

Inversement, c'est parce que les réseaux de sens se superposent dans le discours, c'est-à-dire que les relations sémantiques sont diversifiées (et de ce fait assez imprédictibles) que la mélodie peut avoir la fonction de canaliser un sens, sans qu'il y ait pour autant de redondance avec les structures textuelles. La mélodie fournit en fait les conditions (et les indices) d'une appropriation des structures linguistiques en les réduisant à l'expression privilégiée d'une dominante², et cette restriction est rendue possible parce que les relations syntaxiques sont actualisées par ailleurs. Cette dominante est dans la plupart des cas d'ordre sémantique et pragmatique, car c'est dans ce domaine que peuvent résider les difficultés de compréhension pour un tiers en raison de l'indétermination sémantique (la composante sémantique ayant des structures plus floues que les structures syntaxiques). En ce sens elle permet alors de décoder l'interprétation que le locuteur réalise dans son discours sur le texte. En exerçant la fonction particulière d'expression de l'interprétation, la mélodie peut alors assumer un rôle essentiel et spécifique dans la communication.

Dans ces conditions, la prosodie est bien une zone tampon où la situation détermine un comportement chez le locuteur mais modulé en fonction de sa perception de la situation et de ses propres intentions de communication.

3.2. LES UNITES DE STRUCTURATION

Dans l'analyse que nous avons présentée, cinq entités ont été envisagées au niveau de la structure fonctionnelle du discours : les unités lexicales, les microstructures, les GM, les macrostructures et l'ensemble de l'énoncé.

Les unités lexicales, au niveau des indices de l'énergie et de la durée, sont dans notre analyse les éléments de structuration interne des GMI, des microstructures et des macrostructures, mais aussi les éléments de structuration externe puisque ces unités lexicales sont également en relation avec l'ensemble de l'énoncé. A ce titre elles constituent des unités de base de l'expression sur le plan de l'agencement syntagmatique et souvent syntaxique.

Ces lexèmes étant par ailleurs les éléments d'une évaluation sur les plans syntaxique, sémantique ou pragmatique constituent les unités du plan de l'expression du contenu.

Les macrostructures mélodiques constituent les espaces fondamentaux d'incorporation du sens dans la parole, définissant ainsi successivement des îlots dans lesquels les unités lexicales sont évaluées en fonction d'un principe d'analyse homogène (ou "modèle" dans notre terminologie). Ces îlots, hypothèse vraisemblable, définissent des stratégies d'interprétation à destination d'un tiers, dans lesquelles les GM constituent les plages d'articulation et d'encodage du discours, de même que vraisemblablement celles du décodage chez le destinataire.

Ces plages définissent une stratégie opportuniste de la part du locuteur dans la mesure où il semble adapter sa stratégie aux contenus du texte en fonction des impératifs de communication qui lui sont transmis par la consigne, en organisant les valeurs de F0 selon une représentation linguistique sous-jacente des signifiés (ou modèle), qui devient à ce moment-là prégnante pour lui.

Rares sont les situations où un modèle rend compte de l'ensemble du texte : ceci cependant a été constaté une fois pour le locuteur F3 en consigne 1, avec un taux de

² Cette dominante n'exclut pas d'ailleurs l'expression parallèle d'autres types de relations, syntaxiques, sémantiques ou pragmatiques, mais de manière secondaire.

prédiction d'ailleurs supérieur à la moyenne des autres locuteurs (90%). Mais généralement la variabilité est de règle. Pour la grande majorité des énoncés, la portée de l'accord entre les valeurs d'un indice mélodique et les prédictions d'un modèle, se situe aux environs de 8 mots lexicaux, ce qui correspond à 2 ou 3 GM. Et ceci trouve une confirmation dans les résultats des expérimentations psycholinguistiques : l'empan de l'effet de récence par exemple est d'environ 5 mots (entre autres, Achour & Le Ny, 1983).

En ce qui concerne la structuration "naturelle" des énoncés, ces résultats viennent en concordance par ailleurs avec ceux d'autres chercheurs et notamment ceux de Grosjean & Dommergues (1983) et leur notion de *structures de performance*, et ceux de Selkirk (1978, 1980) avec celle de *syntagme phonologique*. Ces résultats ont été confirmés récemment pour le français (Monnin & Grosjean, 1993; Keller et al., 1993; Zellner, 1996, 1998). Précisant leurs hypothèses en ce qui concerne l'aspect encodage et décodage de la parole, Grosjean & Dommergues avancent l'idée qu'à l'encodage, le traitement du langage peut s'articuler successivement sur ces syntagmes phonologiques. Au décodage, l'auditeur, après avoir localisé la tête de ce syntagme, peut ensuite décomposer ce dernier en ses éléments lexicaux dans le même temps où il entreprend une analyse syntaxique et sémantique. Si cette hypothèse est correcte, on peut ajouter alors que la fonction qu'assume la mélodie est de mettre en relief les unités lexicales qui jouent un rôle central dans ce processus.

Mais pour en revenir à cette mobilité de l'interprétation qui se manifeste dans notre étude par le changement de modèle, on constate qu'il n'existe pas nécessairement de liaison conceptuelle entre ces modèles qui se succèdent au fil du discours, si ce n'est par la trame linguistique. La liaison sur le plan prosodique est à trouver par ailleurs. Elle est en fait assurée à deux niveaux et de manière complémentaire. Au niveau de la mélodie de l'énoncé d'une part et au niveau des paramètres, énergie et durée d'autre part. Sur le plan mélodique, le discours est constitué et structuré comme un tout : comme on l'a vu, le codage qui rend le mieux compte de l'association modèle / indice est, par opposition au codage "phrase", le codage "texte" qui prend en compte l'ensemble de l'énoncé. Par ailleurs on a montré que le passage d'un modèle à un autre s'effectuait sur le plan mélodique par le moyen d'une transition douce entre les valeurs.

D'autre part la fonction de démarcation pluri-indicée de l'énergie et de la durée prend toute son importance. La cohésion entre les unités linguistiques est en effet assurée par les indices de l'énergie et de la durée grâce aux groupes minimaux qui structurent de manière interne et externe les unités lexicales, et d'autre part grâce aux relations syntaxiques et pseudo-syntaxiques qui fondent au niveau des indices de durée et de l'énergie, la plupart des ruptures et des regroupements. Il n'est pas inutile de rappeler que c'est lorsque le changement des modèles et des indices est le plus grand, lorsque les pauses distendent la trame discursive, c'est-à-dire en consigne 3, que s'intensifient les processus de globalisation et de cohésion syntagmatique ou syntaxique.

Ainsi les groupes minimaux forment le point d'intersection entre les plans de la forme et du contenu de l'expression de la signification, c'est-à-dire non seulement un point de jonction, mais un lieu où le contenu interagit avec la forme, où il s'ancre dans l'oral.

Plus précisément encore, le point privilégié de cet ancrage de la signification dans la parole réside non pas dans le contour mélodique, ni même dans l'ensemble du mot, mais dans la syllabe finale. Le processus est tout à fait éloquent. En consigne 1, malgré une lente décroissance de la phrase 1 à la phrase 3, des effectifs de la syllabe finale au profit de ceux du mot, la syllabe finale obtient les pourcentages en moyenne les plus forts.

Cette lente dérive est le signe d'un relâchement de l'attention lorsque l'intelligibilité n'est pas une contrainte essentielle pour les locuteurs.

Inversement dès la consigne 2, la cible "syllabe finale" est le point d'ancrage le plus récurrent pour l'actualisation d'un modèle linguistique par un indice mélodique, et c'est en consigne 3 que le processus se réalise avec le plus d'intensité. Dans ces deux dernières consignes, les effectifs de la syllabe finale sont en constante progression (respectivement de la consigne 1 à la consigne 3, 49% → 63% → 75%). Pour ce faire, la cible du contour a complètement disparu dès la consigne 2 (0%), et les effectifs correspondant à l'ensemble du mot ont largement diminué également (44% → 37% → 25%). Cette progression nette des effectifs de la cible "syllabe finale" s'accommode très bien en fait du ralentissement du débit opéré dès la consigne 2, et sans doute ce ralentissement est une des causes de l'augmentation de ces effectifs.

Ainsi se dessine vraisemblablement l'archétype de la précision dans le discours, modèle vers lequel tendent les locuteurs lorsque des consignes d'intelligibilité les y engagent. Cette précision en fait a deux visages, l'un en relation avec le type de modélisation linguistique, l'autre en relation avec les indices prosodiques. Il y a tout lieu de penser que pour ce corpus, et indépendamment de la consigne 3 pour laquelle l'intelligibilité est aussi matière de phonétique, l'intelligibilité est d'une manière générale synonyme de *précision*.

Sur le plan des indices prosodiques, les indices de F0, de l'énergie et de la durée, sont tous concernés : la plus grande précision est obtenue lorsque les indices démarquent de manière syntaxique les GMI et les microstructures.

Concernant F0, l'archétype de la précision se réalise grâce à $|\Delta F0|$. Mais sur le plan morphologique, cet archétype est atteint dans la syllabe finale du mot lexical, de sorte que la forme achevée de l'archétype est réalisée par l'indice $|\Delta F0|$ en syllabe finale du mot lexical. Toutes les autres formes des indices de F0 à savoir F0M, F0m, et des formes lexicales (contour, ensemble du mot), représentent diverses formes dérivées. Elles sont suffisantes pour l'expression linguistique, mais de moins en moins précises.

Dans le domaine de la modélisation linguistique, l'archétype de la précision se réalise dans notre corpus par l'emploi non seulement des modèles analytiques, mais des modèles sémantiques, à savoir ceux de la complexité lexicale (modèle CM) en phrase 1, et des informations attendues et inattendues (modèle CP) en phrases 2 et 3.

Inversement l'archétype de l'imprécision, pourrait se traduire sur le plan des indices de F0, par l'usage de F0m, ciblé au niveau morphologique sur l'ensemble du mot. En outre, sur le plan des indices de la durée et de l'énergie, les macrostructures prédomineraient jusqu'à occuper l'ensemble de la phrase. Perceptivement cela se traduirait par une énergie de plus en plus faible et par un débit de plus en plus lent au cours de la phrase lue sans pause interne. Sur le plan de la modélisation linguistique, seuls les modèles linguistiques holistiques (EN et HR) seraient utilisés, le modèle HR contribuant à dépersonnaliser l'énoncé.

4. CONCLUSION

Cette étude a tenté de donner, dans le cadre limité d'une expérimentation de lectures, une représentation des stratégies des locuteurs sur le plan prosodique (F0, durée, énergie et pauses), et sur celui du traitement syntaxique, sémantique et pragmatique.

Dans la littérature qui traite des aspects sémantiques de la prosodie, on constate généralement que seuls les éléments de focalisation sont envisagés, ce qui correspond,

certes, aux mots lexicaux jugés les plus importants pour la signification. Toutefois ils ne sont que les pics d'une structure dont on n'a pas, à ma connaissance, rendu compte jusqu'à présent. Le travail relaté dans ces pages, montre en fait qu'une telle structure existe, qu'elle est diversifiable, actualisant le temps de quelques syntagmes, parmi d'autres réseaux virtuels de signification, une dominante (complexité des unités lexicales, distribution des unités informatives, ou encore gestion des connaissances attendues et inattendues, de l'inférence...). On ne peut pas se contenter d'étudier la fonction sémantique de la prosodie par l'examen d'un ou deux items lexicaux de la phrase. Tous les mots lexicaux sont concernés en fait, et le projet de ce travail est de tenter d'en apporter la démonstration.

Dans la limite de cette expérimentation, les résultats montrent que les valeurs issues des indices de F0 sont surtout en relation avec les domaines sémantique et pragmatique, celles des indices de la durée et de l'énergie, avec les domaines syntaxique et morphosyntaxique.

Finalement, l'interprétation que nous donnons des résultats que nous tirons de cette expérimentation, est résumée en entier dans les propos des psycholinguistes, Hupet et Costermans (1981-2), propos que nous avons déjà cités à la fin du chapitre II. Ces propos insistent non seulement sur l'importance de la sémantique et de la pragmatique dans la communication parlée, mais aussi sur l'aspect de compétitivité des divers domaines d'information, compétitivité qui définit en fait toute notre méthodologie. Ces propos relus après l'analyse des résultats, résonnent de manière tout à fait appropriée :

Ne s'agit-il pas en effet, au moyen simplement d'items lexicaux, de marques morphologiques, d'ordre de mots et de contours intonatifs qui constituent notre arsenal linguistique, de transmettre non seulement une variété étonnante d'information sémantique mais aussi une large variété d'information dite pragmatique relative à l'intention communicative du parleur, aux relations existantes ou présupposées entre les interlocuteurs, à l'attitude du parleur à l'égard de ce dont il parle, au caractère supposé connu et accessible ou non de l'information alignée, au caractère thématique ou focalisé de cette information, à ses présuppositions etc. ? Comme le soulignent Slobin (1977) et Bates (1976), l'arsenal linguistique étant assez limité, c'est à une véritable compétition que l'on a affaire, une compétition que se livrent ces diverses informations sémantiques et pragmatiques dans la course à l'entrée dans la chaîne parlée. Cette compétition au droit à la parole peut donner lieu soit à une division des tâches selon laquelle différentes fonctions sont assurées par des dispositifs linguistiques distincts, soit, plus généralement, à une coexistence pacifique selon laquelle un même dispositif est dévolu à l'exercice de fonctions différentes tout en respectant certaines contraintes imposées par l'efficacité de la communication. Hupet et Costermans (1981-2).

Au terme de cette étude, on retire fondamentalement l'idée que la hauteur mélodique, et encore plus l'écart mélodique dans le mot, reproduit sur le plan acoustique une systématisation de type linguistique et pragmatique : le degré de hauteur mélodique au sein du mot lexical, ou le degré de son amplitude, est en effet proportionnel au degré d'implication de ce sens dans une des dimensions sémantique ou pragmatique, dimension dont notre modélisation est bien entendu une approximation. A notre sens, cette organisation fournit les clés objectives de l'interprétation du locuteur sur le texte.