

CHAPITRE XIV

MODÈLES LINGUISTIQUES ET INDICES DE F0 EN CONSIGNES 1, 2, 3 : RÉCAPITULATION GÉNÉRALE

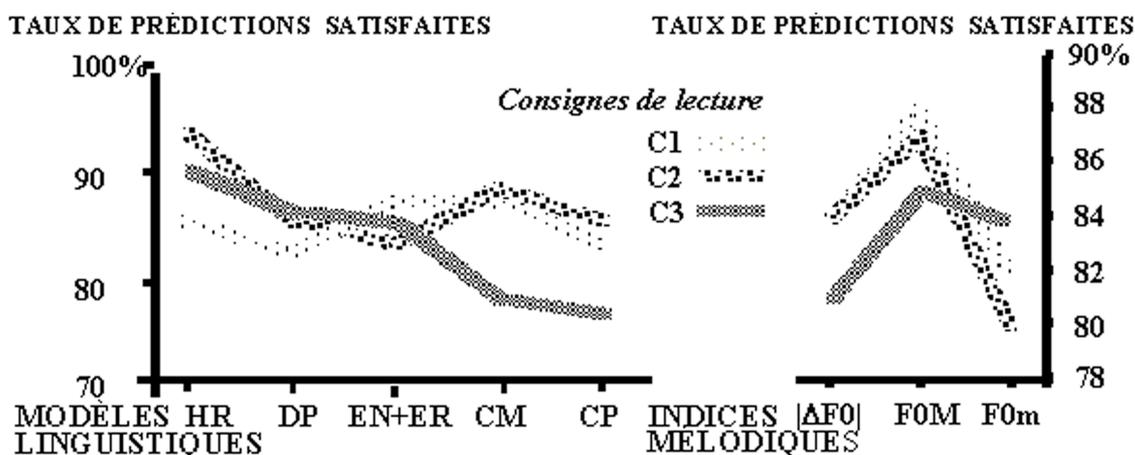
1. OBJECTIFS

Ce chapitre a pour but d'offrir une récapitulation générale des résultats concernant les trois consignes. Ces derniers portent sur les relations entre les modèles linguistiques et les indices de F0 analysées au sein des groupes minimaux, et sur les lieux ou les modes de transitions entre modèles (et entre indices). Les résultats seront présentés de manière comparative entre les 3 consignes.

2. TAUX DE PREDICTIONS SATISFAITES

2.1. MODELES LINGUISTIQUES

Le graphique 14-1 ci-dessous, présente les résultats généraux des modèles linguistiques en fonction des consignes, tous locuteurs confondus et toutes phrases confondues : les taux de prédiction des divers modèles évoluent pour la plupart entre 80 et 90%.



Graphique n° 14-1

Graphique n° 14-3

Valeurs en pourcentages, tous locuteurs confondus, toutes phrases confondues, des taux de prédictions satisfaites

des différents modèles,

Les modèles sont figurés sous forme abrégée (hiérarchie syntaxique HR, dépendance syntaxique DP, hiérarchie énonciative EN, hiérarchie énonciative et rhématisation ER, complexité sémantique CM et connaissance partagée, CP).

des différents indices mélodiques,

Les indices mélodiques sont figurés sous forme abrégée (valeur absolue de l'écart de F0 ou $|\Delta F0|$, maximum de F0 ou F0M, F0 moyen ou F0m).

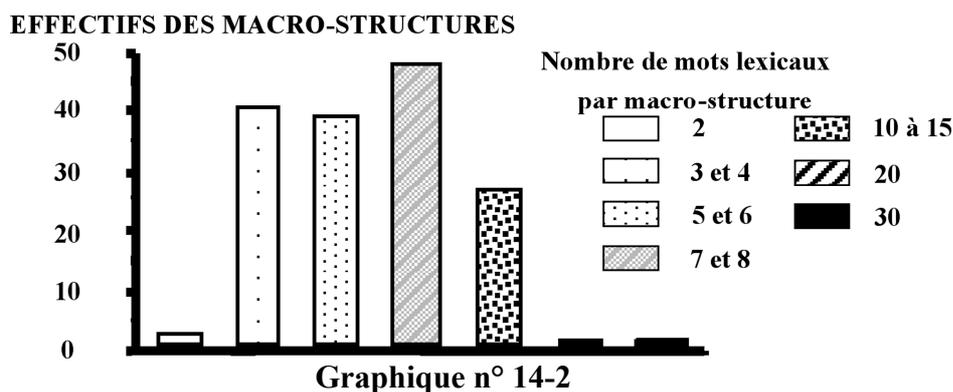
En consigne 1 les taux de prédiction satisfaites sont moins variables. Ainsi de la consigne 1 à la consigne 3, l'écart des taux de prédictions entre les modèles tend à augmenter de sorte qu'en consigne 3 ils se répartissent en 3 ensembles distincts. De ce fait cette consigne possède les scores les plus opposés pour les modèles de la hiérarchie syntaxique (HR, 90%), de la complexité sémantique et de la connaissance partagée (respectivement CM, 78% et CP, 77%). A noter toutefois que le taux global le plus élevé est réalisé par le modèle de la hiérarchie syntaxique en consigne 2 (HR, 93%). Calculés en pourcentages sur les trois consignes, les écarts-types sont respectivement de 2.1% (modèles DP, EN et ER), de 4% (HR), de 4.2% (CP), et de 5.5% (CM).

Les oppositions que l'on observe entre les scores des modèles en fonction des consignes sont toutefois à relativiser. En effet les modèles les moins représentés (par exemple HR) dans les énoncés des locuteurs, ont statistiquement les meilleures chances d'offrir les scores les plus élevés : plus les suites de valeurs sont peu nombreuses et plus il est probable que la variabilité soit réduite. D'un autre côté, il n'est pas surprenant que les modèles les plus subjectifs (les analytiques) aient comparativement les taux les moins bons, puisque l'évaluation dont ils font preuve n'est pas supportée par une structure linguistique (ce qui favorise la programmation à plus long terme et donc un ciblage plus facile), mais élaborée localement (programmation à court terme).

Nous rappelons que les pourcentages sont calculés en comptabilisant pour x groupes minimaux, le nombre moyen de mots lexicaux dont les valeurs mélodiques répondent aux prédictions des modèles, et ce, en fonction de l'ensemble des locuteurs et de chaque consigne de lecture. Ce calcul peut être également opéré sur les portions de texte relatives aux indices, les frontières des deux n'intervenant pas nécessairement au même endroit.

Dans les deux cas, les taux de prédiction pour chaque consigne correspondent au nombre de valeurs mélodiques des mots lexicaux correctement prédites par chaque modèle, ramené aux 11 groupes de l'énoncé et moyenné sur les 12 réalisations (12 locuteurs).

Concernant les modèles, ce nombre moyen varie respectivement de la consigne 1 à la consigne 3¹, de 8 mots lexicaux à 6.5 et 6.2., ce qui ramène le nombre moyen calculé sur l'ensemble des 3 consignes à 6.8 mots lexicaux. Le nombre de ces portions de texte en fonction du nombre de mots lexicaux est très variable. Le graphique 14-2 ci-dessous permet de voir notamment que si les macrostructures de 20 ou 30 mots lexicaux —soit la totalité du texte, ce qui a été effectivement réalisé par le locuteur PE en consigne 1— sont rares (respectivement 1.2% et 0.6%), celles de 2 mots lexicaux le sont tout autant (1.9%). Ces dernières sont vraisemblablement des erreurs dues à la méthode d'analyse, ou encore des défauts d'encodage de la part des locuteurs.



Effectifs des macrostructures (modèles) en fonction du nombre de mots lexicaux qu'elles recouvrent.

Calculé par ailleurs sur les indices de F0, le nombre moyen de mots lexicaux est voisin : il évolue de 6.8 mots lexicaux à 6.2, soit en moyenne pour les 3 consignes, 6.5 mots lexicaux par macro-structure.

Ces données sont intéressantes dans la mesure où elles sont en accord avec les

¹ Pour les comptages, nous avons distingué le modèle EN du modèle ER.

expérimentations des psycholinguistes. Achour & Le Ny (1983) montrent que l'effet de récence, c'est-à-dire de l'accessibilité à la représentation du mot en mémoire, «est sensible pour une distance aussi faible que cinq mots de la phrase» et qu'au-delà ce «souvenir commence immédiatement à décliner». Ainsi le codage prosodique programmé en fonction des représentations linguistiques successives opérant en moyenne sur 6-7 mots, est cognitivement fondé. Ceci laisse à penser que lorsque le modèle s'applique à un nombre de mots supérieur, c'est qu'en fait des phases d'encodage successives réitèrent le processus.

2.2. INDICES MELODIQUES

Le graphique 14-3 ci-dessus propose d'étudier la distribution des taux de prédiction du côté des indices mélodiques, tous locuteurs confondus, toutes phrases confondues et en fonction de chaque consigne de lecture. Ces taux semblent être contenus dans une fourchette plus étroite que ceux des modèles, puisqu'ils évoluent entre 80 et 88%².

Il apparaît en outre que l'indice du maximum de F0 (F0M) détient les scores les plus élevés. Les consignes 1 et 2 présentent les taux de prédiction les meilleurs pour $|\Delta F0|$ et F0M, et ces scores sont toujours supérieurs à ceux de F0m, sauf précisément en consigne 3. L'explication la plus plausible à ce processus, c'est que le rythme très ralenti de la consigne 3 diminue l'effet de récence des cibles mélodiques : de 6 à 7 mots, il tombe à quelques unités, et ce faisant la structure mélodique des cibles au sein de la phrase devient nécessairement plus approximative et plus floue.

3. DISTRIBUTION DES EFFECTIFS

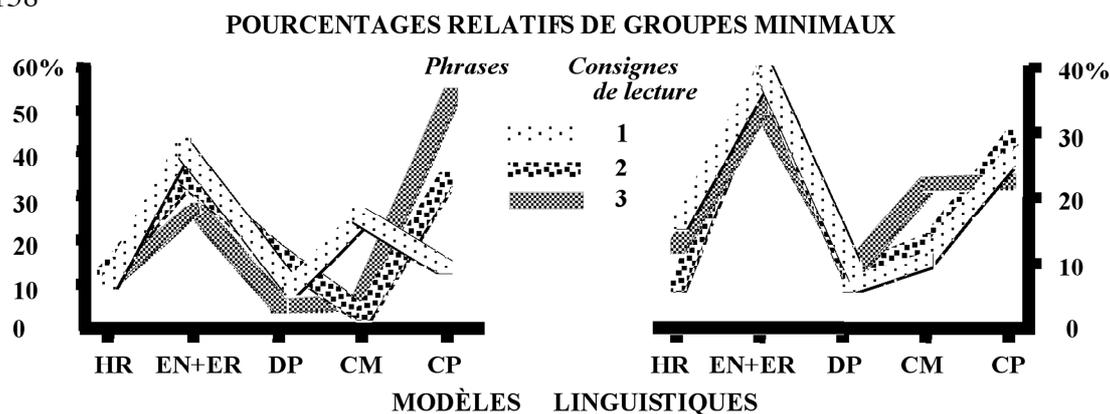
3.1. MODELES LINGUISTIQUES

L'analyse porte sur la distribution des effectifs des groupes minimaux selon les phrases et les consignes en fonction des modèles linguistiques. Le graphique 14-4 ci-dessous présente cette distribution en fonction des phrases, locuteurs confondus, consignes confondues.

Ces résultats globaux confirment les résultats des analyses précédentes. Il existe bien deux types fondamentaux de modèles, ceux qui interviennent de manière prépondérante dans la première phrase et ceux qui interviennent dans les dernières :

- les modèles les plus immédiats à mettre en oeuvre, qui offrent un recours commode lorsque les conditions d'élocution sont plus difficiles, par exemple dans les premiers moments de la prise de parole lorsque l'ancrage mélodique de la signification est encore à construire. Ces modèles, comme nous l'avons déjà précisé au début du chapitre III, offrent une structure holistique aux unités lexicales, c'est-à-dire une place dans une organisation prenant en compte l'ensemble de la phrase. Cette structure repose sur des critères de type soit syntaxique (modèles de hiérarchie syntaxique HR) soit sémantique (modèles énonciatifs EN et ER). Comme la construction linguistique et prosodique s'appuie sur la totalité de la phrase, elle est nécessairement opérée avant la phase de production. Par ailleurs cette opération faisant référence à une structure linguistique, est moins subjective que les opérations analytiques qui ajoutent une évaluation de la part des locuteurs.

² En fait les taux de prédiction sont identiques pour les modèles et les indices de F0. La différence apparente est simplement due au nombre de paramètres différents (6 modèles d'un côté, 3 indices mélodiques de l'autre), et à leur différence de contenu.



Graphique n° 14-4

Graphique n° 14-5

Distribution des différents modèles, exprimés en pourcentages relatifs, tous locuteurs confondus,

toutes consignes confondues en fonction de chacune des phrases. *toutes phrases confondues en fonction de chacune des consignes.*

Les noms des modèles sont figurés sous forme abrégée (hiérarchie syntaxique HR, dépendance syntaxique DP, hiérarchie énonciative EN, hiérarchie énonciative et rhématisation ER, complexité sémantique CM et connaissance partagée, CP).

Par ailleurs, ces modèles structuraux, en gérant l'organisation formelle d'un contenu linguistique, fondent l'évaluation sur une simple compétence de langue. De ce fait ils économisent l'effort, ce qui facilite leur utilisation en période plus délicate, lorsque par exemple le poids des contraintes augmente sensiblement. Autrement dit, ces modèles nécessitant un moindre coût cognitif permettent d'économiser l'énergie disponible pour engager éventuellement d'autres traitements (ciblage mélodique plus précis par exemple).

- avant de commenter la distribution des modèles analytiques (en particulier CM, CP), revenons sur la spécificité lexicale des phrases : la phrase 1, la plus longue, possède le vocabulaire le plus spécialisé ; la phrase 2 plus courte, présente le vocabulaire le plus simple ; la phrase 3 la plus courte, comporte des informations les plus inattendues³.

En phrase 1, à côté des modèles holistiques, on trouve également le modèle CM qui fournit une échelle de la complexité lexicale en contexte. Ce modèle, contrairement aux modèles holistiques, est dépendant du contexte local. Or il est représenté davantage en phrase 1 alors que le registre des mots est le plus spécialisé du texte, et beaucoup moins en phrases 2 et 3 (1% en phrase 2), alors que les mots sont plus facilement accessibles.

Ces divers modèles (HR, EN, ER et CM) présentent dans le graphique 14-4 ci-dessus, des effectifs de plus en plus réduits au fur et à mesure que la lecture progresse dans le texte.

Le modèle CM comme les autres (DP, CP), nécessite une évaluation plus importante et donc à notre avis, un traitement cognitif plus conséquent. Ces modèles font sans doute moins référence à la compétence acquise, mais nécessitent une interprétation et une évaluation du contenu du texte de la part du locuteur. Ils sont donc à la fois

³ 1° l'existence de sources thermales dans les bas-fonds océaniques 2° l'existence d'une température chaude dans ces régions généralement réputées froides 3° les conditions écologiques inattendues de la prospérité des vers géants (température élevée de ces sources).

dépendants de ce texte et moins formels. Rien n'interdit d'ailleurs de penser que cette évaluation plus subjective prend appui déjà sur une représentation de la structuration linguistique, ce qui augmente de toute façon la charge cognitive. Si ces modèles sont majoritairement distribués en deuxième partie de l'énoncé, c'est qu'ils nécessitent vraisemblablement que les bases conceptuelles du discours aient été posées dans une phase antérieure.

Par ailleurs, on le sait, le modèle des informations inattendues CP, comme le modèle CM, inscrit ses analyses dans le contexte local. En phrases 2 et 3, l'information devient plus inattendue, et c'est justement dans ces deux dernières phrases que la progression la plus importante des effectifs se réalise et précisément comme le montre le graphique 14-4 ci-dessus, au bénéfice de ce modèle. Comme on le voit, le contexte crée les conditions nécessaires à la sélection des modèles linguistiquement ou pragmatiquement adéquats.

Ce relais des modèles "holistiques" par les modèles "analytiques" de la phrase 1 à la phrase 3, est bien exprimé par les modèles les plus représentés qui deviennent de ce fait prototypiques : en phrase 1, les modèles EN+ER sont les plus utilisés, en phrase 2, les modèles EN+ER et CP sont à égalité, en phrase 3, le modèle CP est largement majoritaire.

Par ailleurs le graphique n° 14-5 ci-dessus présente la distribution des effectifs des GM selon les différents modèles, tous locuteurs confondus et toutes phrases confondues, et en fonction des différentes consignes. On constate que les consignes n'ont pas un effet de remise en cause radicale de la distribution, mais simplement un effet de modulation.

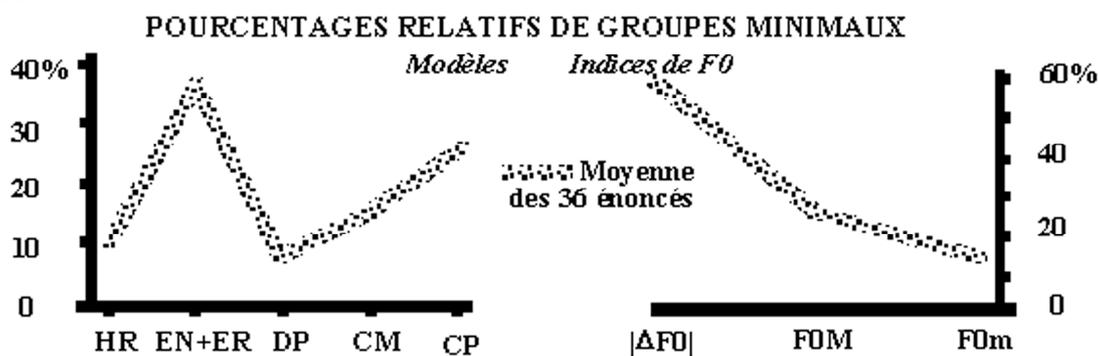
De l'examen de ce graphique, il ressort que cette modulation exercée par les différentes consignes sur la distribution est significative. En effet d'une consigne à l'autre, les modèles holistiques EN+ER⁴ et HR évoluent, toutes proportions gardées, de la même façon, mais ils perdent de plus en plus de leurs effectifs au bénéfice des modèles qui envisagent les relations locales des mots lexicaux, et spécialement au profit des modèles CM et CP. Si cependant le modèle DP est très peu bénéficiaire au niveau global quelle que soit la consigne, le modèle CP progresse le plus en consigne 2, et le modèle CM en consigne 3.

En fait les locuteurs mettent en oeuvre une stratégie nouvelle. Contrairement à la consigne 2, ce n'est plus le traitement de l'information inattendue qui est prioritaire pour la consigne 3 (les plus grands écarts mélodiques aux unités lexicales véhiculant les informations les plus inattendues), mais celui de l'information jugée comparativement la plus complexe (les plus grands écarts aux unités lexicales les plus complexes). Cette diversification des stratégies traduit bien chez certains locuteurs une souplesse d'adaptation aux exigences d'une consigne réclamant plus d'intelligibilité encore.

En tout état de cause, quelle que soit la consigne, ce sont les modèles EN+ER qui possèdent globalement les effectifs les plus nombreux, puis le modèle CP. Le graphique n° 14-6 ci-dessous donne la distribution moyenne globale des effectifs des modèles (estimés en fonction des groupes minimaux), tous locuteurs confondus, toutes phrases confondues, toutes consignes confondues, c'est-à-dire sur les 36 énoncés.

On peut ainsi résumer l'effet consigne à cette propension chez les locuteurs à passer d'une perspective holistique et structurale de présentation des items lexicaux en contexte, à une perspective locale et analytique allant de pair avec une réduction très sensible du débit.

⁴ Les modèles de l'énonciation restent majoritaires au niveau global, mais cependant pas dans les deux dernières phrases.



Graphique n° 14-6

Graphique n° 14-9

Distribution globale moyenne des groupes minimaux tous locuteurs confondus, toutes phrases confondues, toutes consignes confondues,

en fonction des modèles linguistiques, en fonction des indices mélodiques,

Les modèles sont figurés sous forme abrégée (hiérarchie syntaxique HR, dépendance syntaxique DP, hiérarchie énonciative EN, hiérarchie énonciative et rhématisation ER, complexité sémantique CM et connaissance partagée, CP). Les indices mélodiques sont figurés sous forme abrégée (valeur absolue de l'écart de F0 ou |ΔF0|, maximum de F0 ou F0M, F0 moyen ou F0m).

Dans le cadre de notre expérimentation, (les conditions de lecture établissant une communication avec un auditeur fictif mais potentialisé par des consignes de lecture, et réclamant une intelligibilité de plus en plus forte), nous pouvons établir que sur l'ensemble des 36 énoncés, les modèles sémantiques (essentiellement EN+ER) et pragmatique (CP) sont les modèles les plus représentatifs.

3.2. INDICES MELODIQUES

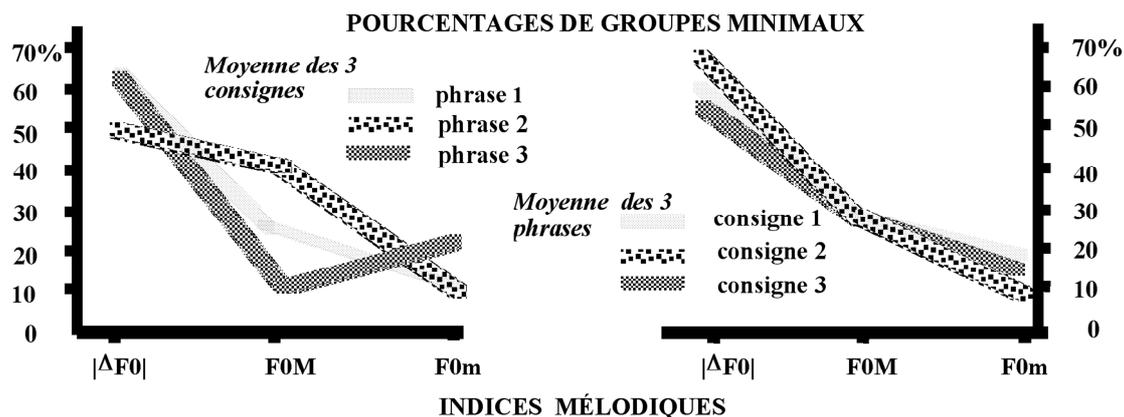
Le graphique n° 14-7 ci-dessous envisage la distribution des effectifs des 3 indices, valeur absolue de l'écart, maximum et F0 moyen, en fonction de chacune des phrases, tous locuteurs confondus, toutes consignes confondues (36 énoncés).

Les indices mélodiques ont aussi leurs caractéristiques et leurs contraintes. Des 3 indices, nous savons que |ΔF0| est certainement l'indice le plus délicat à réaliser dans la mesure où il exige de positionner au sein des pentes mélodiques croissantes et décroissantes, le temps de quelques millisecondes, 2 cibles qui sont par rapport au mot lexical des extrema absolus et par rapport à la chaîne mélodique de la phrase et du texte, des extrema relatifs.

Lorsque l'effort est trop grand, les locuteurs positionnent une des 2 cibles, le maximum de F0, en une position clé du mot lexical. Lorsque ces conditions sont encore trop difficiles, il ajuste alors au cours du mot, des valeurs mélodiques approximatives dans le registre voulu (au besoin par approximations successives grâce au feed-back). Ces 3 indices |ΔF0|, F0M et F0m semblent donc, comme nous l'avons déjà montré, se comporter comme des formes progressivement détériorées d'un même phénomène mélodique.

Si nous nous reportons maintenant au graphique 14-7 ci-dessous, nous constatons qu'en phrase 1, les effectifs de l'indice |ΔF0| ont par rapport aux autres indices, les pourcentages les plus élevés (65%). F0M vient en second (23% des effectifs), puis F0m (avec 12%).

En phrase 2, les scores de F0m restent stables, mais une balance des effectifs se produit entre les deux premiers indices, au détriment de |ΔF0| (65% → 50%) et donc au bénéfice de F0M (24% → 41%). En phrase 3, le recul extrêmement sensible de F0M profite à |ΔF0| (50% → 64%), mais aussi à F0m (9% → 23%).



Graphique n° 14-7

Graphique n° 14-8

Distribution des différents indices mélodiques, exprimés en pourcentages relatifs, tous locuteurs confondus, toutes consignes confondues en fonction de toutes phrases confondues en fonction de chaque phrase.

Les indices mélodiques sont figurés sous forme abrégée (valeur absolue de l'écart de F0 ou $|\Delta F0|$, maximum de F0 ou F0M, F0 moyen ou F0m).

Le processus en fait est très clair : la phrase 1 du fait de sa longueur et de sa relative complexité a fortement mobilisé les capacités d'attention du locuteur, de telle sorte qu'une économie des moyens (sous l'effet probable du relâchement de l'attention) se ressent en phrase 2, processus qui aboutit à une sélection moins fréquente de $|\Delta F0|$ (c'est-à-dire finalement à un mauvais ciblage du minimum de F0). Dans ces conditions, les effectifs de $|\Delta F0|$ diminuent sensiblement, tandis que ceux de F0M progressent. Mais la phrase 2, plus courte et au lexique plus simple, permet de relâcher les contraintes, ce qui crée à nouveau pour certains locuteurs, des conditions favorables pour remobiliser l'attention en phrase 3, et donc une utilisation plus fréquente de $|\Delta F0|$. Toutefois on note parfois l'émergence d'une nouvelle économie des moyens, opérée cette fois aux dépens de F0M, et qui se traduit par l'augmentation des effectifs de l'indice le moins exigeant et le moins expressif, F0m.

Le graphique 14-8 ci-dessus distribue les effectifs des groupes minimaux selon les 3 indices, en fonction des 36 énoncés. Le premier commentaire porte sur la grande homogénéité et la grande cohérence des résultats quelle que soit la consigne : la distribution globale des effectifs entre les trois types d'indices est du même ordre que précédemment.

On constate que les consignes 1 et 3 ont des effectifs extrêmement voisins. Celles-ci pourtant possèdent des débits et un régime de pauses très différents : on passe en effet d'une moyenne de 2.23 mots / seconde à celle de 1.05, et en consigne 3, l'occurrence de pauses très nombreuses vient fragmenter le discours en îlots d'une seule unité lexicale.

La consigne 2 est la consigne qui offre les qualités de précision linguistique les plus grandes grâce à $|\Delta F0|$ (66% des groupes minimaux), alors que le débit n'est pas encore extrêmement ralenti (1.82 mots / seconde) par rapport à celui de la consigne 1. Comme on le constate sur ce graphique, le surcroît de sélection de $|\Delta F0|$ s'effectue aux seuls dépens de F0m (16% → 8%). Il est remarquable que les consignes de lecture ne produisent pas d'effet sur les effectifs moyens de F0M.

Finalement le graphique 14-9 ci-dessus donne la distribution générale des effectifs en

fonction des 3 indices mélodiques, sur les 36 énoncés. Il est clair que $|\Delta F0|$ correspond au nombre de groupes minimaux le plus élevé. Les pourcentages indiquent qu'il totalise en moyenne environ les deux-tiers de l'effectif des groupes minimaux (60%) alors que F0M détient les deux-tiers des effectifs restants (26%), F0m n'en regroupant que 13%.

Comme $|\Delta F0|$ est utilisé par tous les locuteurs quelle que soit la consigne, on peut dire que cet indice est le plus adéquat à la tâche de lecture quand, à destination d'un compreneur, elle doit être intelligible ou très intelligible.

4. CONCLUSION

Les points principaux de ce chapitre récapitulatif sont les suivants :

1° Les taux de prédiction moyens des modèles sont en moyenne de 85%.

2° Les tronçons définis sur le plan des modèles correspondent vraisemblablement aux unités d'encodage homogène du discours et sont constitués d'un ensemble variable de groupes minimaux. En moyenne ces tronçons se composent d'environ 7 mots lexicaux.

3° Une lecture intelligible ou très (très) intelligible repose dans notre expérimentation sur la sélection majoritaire de deux modèles sémantique puis pragmatique, le modèle EN+ER, puis CP, le premier offrant une structure holistique, le deuxième une structure plus analytique. Cette stratégie est encore la même lorsque les consignes d'intelligibilité pèsent de plus en plus sur le discours, et alors que le débit se ralentit considérablement.

4° L'indice le plus courant est le plus délicat à mettre en oeuvre, $|\Delta F0|$. Étant donné que pour l'écart, le locuteur cible 2 seuils à la fois absolus et relatifs, en fonction d'une échelle, cet indice est sans doute comparativement, le plus fiable et le plus ajusté. Cette évaluation opérée de manière inconsciente par le locuteur, porte sur la relation entre la modélisation du contenu et le ciblage des indices mélodiques, c'est-à-dire sur les composantes fondamentales du versant langue et du versant parole.

Au prochain chapitre, nous envisageons les autres indices prosodiques, issus des paramètres de l'énergie (chapitre XV) et de la durée (chapitre XVI).