L'année psychologique

# Les structures de performance en français : caractérisation et prédiction

Pascal Monnin, F. Grosjean

**Citer ce document / Cite this document :**

Monnin Pascal, Grosjean F. Les structures de performance en français : caractérisation et prédiction. In: L'année psychologique. 1993 vol. 93, n°1. pp. 9-30;

doi : https://doi.org/10.3406/psy.1993.28679

https://www.persee.fr/doc/psy_0003-5033_1993_num_93_1_28679

Fichier pdf généré le 18/04/2018

**Résumé**

Résumé

Les recherches réalisées jusqu'à maintenant dans le but d'étudier l'évolution de la représentation gestuelle d'actions en tant qu'activité figurative avaient des difficultés à dégager des niveaux d'évolution cohérents. De là l'hypothèse, formulée par certains auteurs, d'une évolution indépendante de ce type d'activités figuratives par rapport au développement cognitif général.

La recherche présentée ici montre que ces difficultés sont liées à l'absence de prise en compte, dans l'analyse des catégories, du signifiant et du signifié des gestes.

Les modifications introduites dans le paradigme expérimental ont permis de dégager une évolution cohérente des comportements dans un échantillon de 320 enfants âgés de 4 à 12 ans ; 4 niveaux, sont en effet observés, chacun d'eux pouvant se définir comme une structure, c'est-à-dire comme composé d'un certain nombre d'éléments reliés entre eux par des rapports hiérarchiques dans un réseau représentationnel.

Mots clés : représentation gestuelle, réseau représentationnel, activité figurative, structure.

**Abstract**

Summary : Characterizing and predicting performance structures in French.

Over the past fifteen years, considerable attention has been paid to the performance structures of sentences, that is those structures based on experimental data such as pausing or parsing values. These structures are characterized by a number of properties (hierarchy, symmetry and basic units of equal size) and can best be predicted by algorithms which take into account the prosodic structure of the sentence. Research on performance structures, which has mainly concentrated on English, has recently influenced the decelopment of prosodic rules in text-to-speech synthesis. In this study, we turn to the performance structures of French. First we describe the structures obtained by means of an oral reading task and show some interesting differences with English (the prosodic status of postposed adjectives, the bundling of certain function words to heads on the left, etc.). Then we attempt to predict these structures by means of slightly modified versions of the algorithme originally developed for English and show that the predictions obtained are not jully satisfactory. Finally we propose a new algorithm, specifically designed for French, and demonstrate that it is a far better predictor of the data obtained.

Key-words : psycholinguistics, production, speech, performance structures, prediction algorithms.

BY: $ = creative commons

*L'Année psychologique*, 1993, *93*, 9-30

# MÉMOIRES ORIGINAUX

*Laboratoire de traitement du langage et de la parole*
*Université de Neuchâtel*[1]

## LES STRUCTURES DE PERFORMANCE EN FRANÇAIS : CARACTÉRISATION ET PRÉDICTION[2]

par Pascal MONNIN et François GROSJEAN

Persée creative commons BY: =

*SUMMARY : Characterizing and predicting performance structures in French.*

*Over the past fifteen years, considerable attention has been paid to the performance structures of sentences, that is those structures based on experimental data such as pausing or parsing values. These structures are characterized by a number of properties (hierarchy, symmetry and basic units of equal size) and can best be predicted by algorithms which take into account the prosodic structure of the sentence. Research on performance structures, which has mainly concentrated on English, has recently influenced the development of prosodic rules in text-to-speech synthesis. In this study, we turn to the performance structures of French. First we describe the structures obtained by means of an oral reading task and show some interesting differences with English (the prosodic status of postposed adjectives, the bundling of certain function words to heads on the left, etc.). Then we attempt to predict these structures by means of slightly modified*

1. Avenue du Premier Mars 26, CH 2000 Neuchâtel, Suisse.

2. Cette étude a pu être entreprise et menée à bien grâce à des subventions en provenance des organismes suivants : en France, l'Observatoire des industries de la langue (projet Multidif en collaboration avec les Universités de Grenoble et de Bruxelles) ; en Suisse, le Fonds national suisse de la recherche scientifique (subvention n° 12-33582.92). Les auteurs tiennent à remercier Gérard Bailly, Jean-Yves Dommergues, Lysiane Grosjean, Christian Rubattel, Juan Segui, Corinne Tschumi.

*versions of the algorithms originally developed for English and show that the predictions obtained are not fully satisfactory. Finally we propose a new algorithm, specifically designed for French, and demonstrate that it is a far better predictor of the data obtained.*

*Key-words : psycholinguistics, production, speech, performance structures, prediction algorithms.*

## INTRODUCTION

Nous savons depuis quelque temps qu'il est possible, grâce à des tâches expérimentales diverses telles que le rappel, la lecture et la segmentation subjective, d'obtenir la structure hiérarchique d'une phrase (Grosjean, Grosjean et Lane, 1979; Dommergues et Grosjean, 1981 ; Gee et Grosjean, 1983). Cette structure, que l'on a nommé « structure de performance », est construite uniquement à partir de mesures dépendantes (probabilité de rappel, durée des pauses, indice de segmentation, etc.), et peut être caractérisée par un certain nombre de propriétés fondamentales : unités de base de longueur plus ou moins égale, organisation hiérarchique et structure symétrique (Grosjean et Dommergues, 1983). La structure de performance d'une phrase correspond moins à son organisation syntaxique qu'à sa structure prosodique (les expressions « structure de performance » et « structure prosodique » sont devenues synonymiques dernièrement), et peut être prédite de manière satisfaisante à l'aide de diverses variables telles que la nature, la longueur et le schéma accentuel des mots et des constituants, la structure syntaxique de la phrase, l'équilibre des constituants, etc. Les travaux sur les structures de performance, qui ont porté avant tout sur l'anglais, ont permis de mieux comprendre certaines étapes de la production d'un énoncé (Levelt, 1989) et ont servi de base à la prédiction de la prosodie dans la synthèse de la parole (Bachenko et Fitzpatrick, 1990 ; Quené et Kager, 1992 ; Wang et Hirschberg, 1992).

Dans cette étude, nous nous tournons vers les structures de performance du français afin de déterminer dans quelle mesure les principes et les algorithmes de l'anglais sont également pertinents pour cette langue. Bien que les travaux sur la structuration prosodique en français soient déjà fort nombreux (voir, entre autres, Aubergé, 1991 ; Bailly, 1989 ; Caelen-Haumont,

1991 ; Hirst, 1987 ; Llorca, 1984 ; Pasdeloup, 1990 ; Rossi, 1985 ; Vaissière, 1980), il nous a semblé important de mener une première étude exploratoire dans laquelle nous adapterions l'approche utilisée jusqu'à présent en anglais afin de caractériser et de prédire les structures de performance du français oral. Nos objectifs étaient, d'une part, d'étudier les structures de performance d'une langue qui possède une organisation prosodique très différente de l'anglais et, d'autre part, d'apporter un complément d'information aux travaux des chercheurs cités ci-dessus. En effet, la méthodologie utilisée dans nos travaux nous permet d'obtenir une hiérarchie prosodique très fine de la phrase, d'isoler les unités prosodiques, de quantifier l'importance des coupures qui existent entre elles et de tester des algorithmes de prédiction des structures de performance.

Dans la première partie de cette étude nous caractériserons les structures obtenues à l'aide d'une tâche de lecture orale, et dans la seconde partie, nous prédirons ces structures au moyen d'algorithmes déjà existants (mais légèrement modifiés) et d'un nouvel algorithme que nous avons développé spécifiquement pour le français.

## 1. CARACTÉRISATION DES STRUCTURES DE PERFORMANCE

L'objet de cette première partie est d'obtenir des structures de performance en français (en adaptant, lorsqu'il le faut, les mesures dépendantes utilisées jusqu'à présent en anglais) et de caractériser ces mêmes structures. Nous chercherons en particulier à évaluer si les caractéristiques de symétrie, de hiérarchie et de longueur égale des unités de base se retrouvent en français. De plus, nous observerons l'impact de certaines variables telles que la position des adjectifs (préposés ou postposés) et le statut prosodique de certains mots grammaticaux.

### MÉTHODE

#### SUJETS

Huit sujets, tous de langue maternelle française, ont participé à cette étude.

MATÉRIEL

Neuf phrases simples, d'une longueur de quinze syllabes, ont servi de matériel expérimental. Afin d'obtenir une certaine variété dans ces phrases, nous avons fait varier la longueur des groupes nominaux sujet et objet en modifiant le *nombre* de mots lexicaux (et donc la structure du groupe) dans les cinq premières phrases et la *longueur* de ces mots dans les quatre autres. Avant d'être présentées aux sujets, les phrases suivantes ont été tapées dans un ordre aléatoire sur des bandes de papier :

1) (La fille) s'est déguisée en (une jolie petite fée espiègle).
2) (La petite fille) s'est déguisée en (une jolie fée espiègle).
3) (La fille espiègle) s'est déguisée en (une jolie petite fée).
4) (La jolie petite fille) s'est déguisée en (une fée espiègle).
5) (La jolie petite fille espiègle) s'est déguisée en (une fée).
6) (La dame chic) possède (un spectaculaire chat indonésien).
7) (Le voisin âgé) possède (un magnifique chat japonais).
8) (Le professeur sympathique) possède (un joli chat tigré).
9) (Le propriétaire accommodant) possède (un beau chat blanc).

PROCÉDURE

Les sujets ont été enregistrés individuellement au moyen d'un magnétophone à cassette SONY WM-D3. Ils ont d'abord pris connaissance de chaque phrase et l'ont lue ensuite à deux débits, d'abord normal et ensuite lent.

ANALYSE DES DONNÉES

Nous avons obtenu 16 enregistrements pour chaque phrase (8 sujets, 2 lectures par sujet) que nous avons digitalisés à l'aide du système Mac Recorder rattaché à un ordinateur Macintosh SE. Les tracés oscillographiques obtenus nous ont permis de mesurer la durée des voyelles et des pauses en fin de mots. Contrairement à notre approche pour l'anglais, où nous ne mesurions que la durée de la pause silencieuse, nous avons opté ici pour une mesure plus complexe. En effet, la prédominance des syllabes ouvertes en français permet au locuteur de marquer une coupure temporelle soit par l'allongement de la dernière syllabe du mot, soit par l'insertion d'une pause silencieuse, soit par les deux approches à la fois (Duez, 1987). Nous avons donc décidé d'inclure dans ce que nous appelons ici « voyelle + pause » (dorénavant V + P), la durée du noyau (la voyelle) de la syllabe précédant une coupure entre deux mots ainsi que la durée de la pause silencieuse (sans limite inférieure) entre ces mots. Nous n'avons pas tenu compte de la consonne finale (lorsque celle-ci était présente) afin d'obtenir la même mesure à chaque coupure de mot, certains mots se terminant avec une voyelle, d'autres avec une voyelle et une consonne, d'autres encore avec une

voyelle et deux consonnes[3]. La mesure V + P est donc toujours composée au moins de la durée du noyau; la durée de la pause silencieuse vient s'y ajouter lorsqu'un arrêt est également présent dans l'onde sonore. Notons que cette mesure nous permet d'obtenir une valeur entre *chaque* mot de la phrase, même à débit normal, ce qui n'est pas normalement possible avec la seule durée des pauses silencieuses (les sujets ne faisant pas de pauses à l'intérieur des constituants ou entre constituants mineurs à débit normal). Nous avons ensuite calculé la moyenne des V + P, pour chaque phrase, sur l'ensemble des sujets et avons converti ces valeurs en pourcentages basés sur la somme totale des V + P dans la phrase[4].

Avant même de construire les structures de performances de chacune de nos phrases à l'aide des valeurs obtenues, nous avons dû choisir entre les V + P obtenues à débit normal, à débit lent ou aux deux débits à la fois. Rappelons que dans les études portant sur l'anglais, une moyenne basée sur l'ensemble des débits était utilisée car au seul débit normal les pauses ne sont pas assez nombreuses et ne permettent donc pas la construction d'arbres hiérarchiques complets. Nous avons calculé, à l'aide des valeurs V + P, les corrélations entre les deux débits pour les neufs phrases et avons obtenu une corrélation moyenne de 0,86. Devant cette valeur élevée, qui indique une forte ressemblance des structures de performance à débit normal et à débit lent, et étant donné que notre objectif était avant tout d'étudier la structure prosodique d'une phrase à débit normal, nous avons décidé de ne pas tenir compte des durées à débit lent. Nous avons donc construit pour chacune des neuf phrases, et à l'aide des valeurs V + P obtenues à débit normal, une structure de performance en suivant la procédure itérative suivante (Grosjean *et al.*, 1979) :

*a)* Trouver la/les V + P la/les plus courte(s) de la phrase;
*b)* Grouper les éléments (mots ou groupes de mots) séparés par cette/ces V + P en les rattachant à un nœud commun et ne plus tenir compte de cette/ces V + P;
*c)* Passer à la/les V + P la/les plus courte(s) suivante(s) et rattacher les mots (ou ensembles de mots) séparés par cette/ces V + P;
*d)* Continuer le processus jusqu'au regroupement de tous les mots.

3. Dans une étude pilote, nous avons remarqué que si la hauteur des nœuds les plus bas des structures de performance pouvait changer légèrement lorsqu'on inclut la consonne dans les mesures, les constituants prosodiques majeurs restaient les mêmes.

4. Les données ont été converties en pourcentages afin de suivre au plus près la procédure utilisée dans les études précédentes et ne pas donner trop d'importance aux valeurs d'un seul sujet qui aurait pu avoir, grâce à un débit très lent (ou très rapide), des durées V + P beaucoup plus longues (ou plus courtes), que les autres sujets.

Cette procédure, qui peut également être utilisée avec d'autres variables dépendantes (probabilité de rappel, indice de segmentation, etc.), permet de représenter visuellement les structures de performances et ainsi de les décrire plus facilement.

## RÉSULTATS ET DISCUSSION

La série de structures obtenue montre que les trois propriétés fondamentales des structures de performance qui ressortent en anglais (Grosjean et Dommergues, 1983 ; Gee et Grosjean, 1983) se retrouvent dans les structures du français, à savoir la symétrie, la hiérarchie et la longueur plus ou moins égale des unités de base[5]. Concernant la symétrie, nous observons que la frontière prosodique principale est située vers le milieu de la phrase et ne correspond donc pas toujours à la coupure syntaxique principale. Dans la figure 1, par exemple, le syntagme nominal sujet « la fille » est court et contraint le locuteur à marquer une coupure principale après le verbe « s'est déguisée » pour des raisons de symétrie. Il s'ensuivra, comme nous le verrons ci-dessous, que la seule structure syntaxique d'une phrase n'est pas forcément un bon prédicteur de la structure de performance. En ce qui concerne la deuxième caractéristique, la hiérarchie, nous remarquons dans nos phrases que les unités de base se regroupent en unités plus grandes qui, elles-mêmes, sont regroupées en unités encore plus larges. Dans la figure 1 par exemple, « la fille » est rattaché à « s'est déguisée » tandis que « en une jolie » est rattaché à « petite fée » et à « espiègle ». Ces deux groupes sont ensuite regroupés au niveau supérieur. Enfin, troisième caractéristique, les structures de performance en français, comme celles de l'anglais, sont constituées d'un certain nombre d'unités prosodiques de longueurs plus ou moins égales,

5. Notons que ceci est vrai pour les structures de chacun des sujets. En effet, nous avons corrélé, pour chaque phrase, les valeurs V + P de chaque sujet avec celles de chacun des autres sujets, puis avons obtenu la moyenne des corrélations. Nous avons ensuite calculé la moyenne pour chaque phrase et avons finalement moyenné ces 9 moyennes. Le résultat final est de 0,84, ce qui signifie que la structure de performance d'une phrase produite par un sujet est fort semblable à celles produites par les autres sujets. Bien qu'il existe quelques petites divergences entre sujets, la configuration générale des structures individuelles est pratiquement identique et révèle les trois mêmes caractéristiques, à savoir, la symétrie, la hiérarchie et la longueur plus ou moins égale des unités de base.

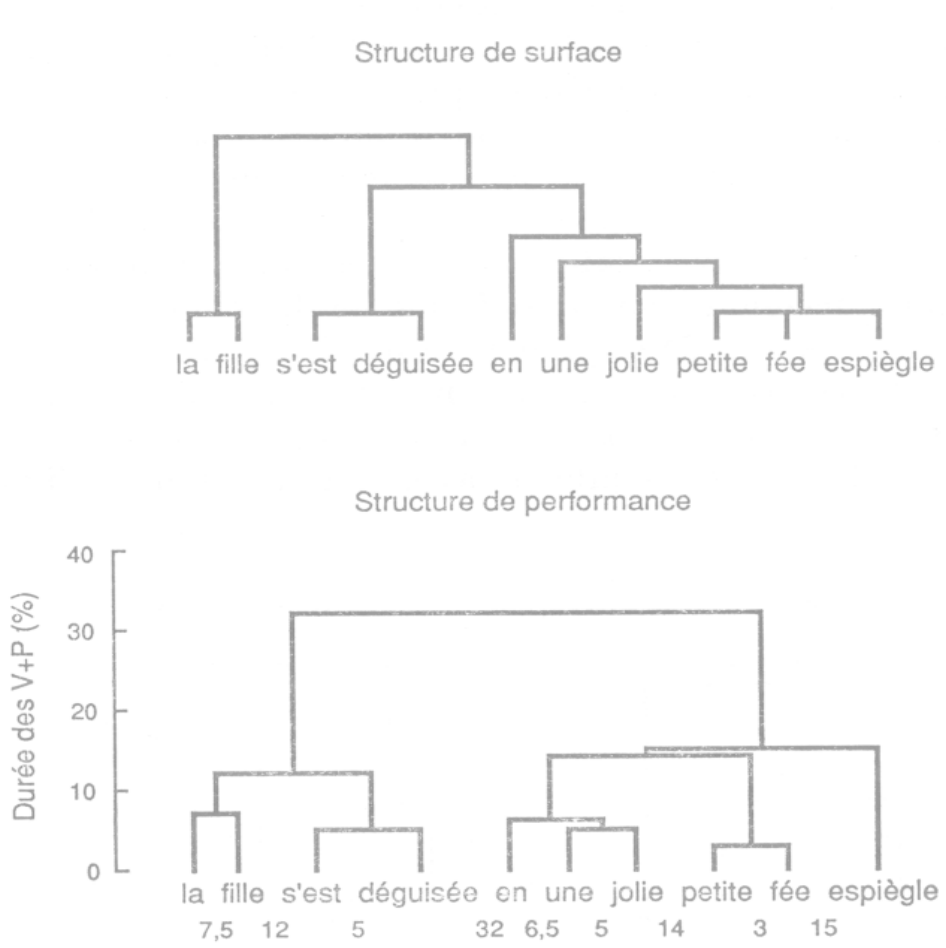


Fig. 1. — Structure de surface et structure de performance de la phrase 1. La structure de performance est élaborée à partir des durées des V + P converties en pourcentages

*The surface structure and performance structure of sentence 1. The latter is based on percentage vowel + pause durations*

unités qui ne correspondent pas toujours à un constituant syntaxique. Par exemple, dans la figure 1, le syntagme prépositionnel « en une jolie petite fille espiègle » est décomposée en trois unités prosodiques distinctes, « en une jolie », « petite fée » et « espiègle ». Ces unités de base ne semblent pas excéder un certain nombre de syllabes qui est sensiblement le même pour chaque unité. Ainsi, toujours dans la même phrase, on distingue 5 groupes prosodiques de base dont le nombre de syllabes varie entre 2 et 4 : « la fille », « s'est déguisée », « en une jolie », « petite fille » et « espiègle ». Sur l'ensemble des phrases, la longueur moyenne des unités prosodiques séparées par une V + P de

psy_0003-8033_1993_num_93_1 /F

12 % ou plus (dans la distribution des effectifs des V + P il existe une rupture à 12 %) est de 3,46 syllabes (écart type de 1,43). Cette troisième caractéristique explique le découpage des syntagmes nominaux sujets que l'on trouve dans les phrases 6 et 9 de la figure 2, syntagmes qui ont exactement la même structure syntaxique. D'un côté « la dame chic » (phrase 6) forme un seul groupe prosodique de 3 syllabes alors que dans la phrase 9, « le propriétaire accommodant » est formé de deux unités prosodiques distinctes de 5 et de 4 syllabes respectivement.

En plus de ces trois propriétés fondamentales, deux autres caractéristiques ressortent de nos résultats. La première concerne les adjectifs postposés qui semblent avoir un statut prosodique

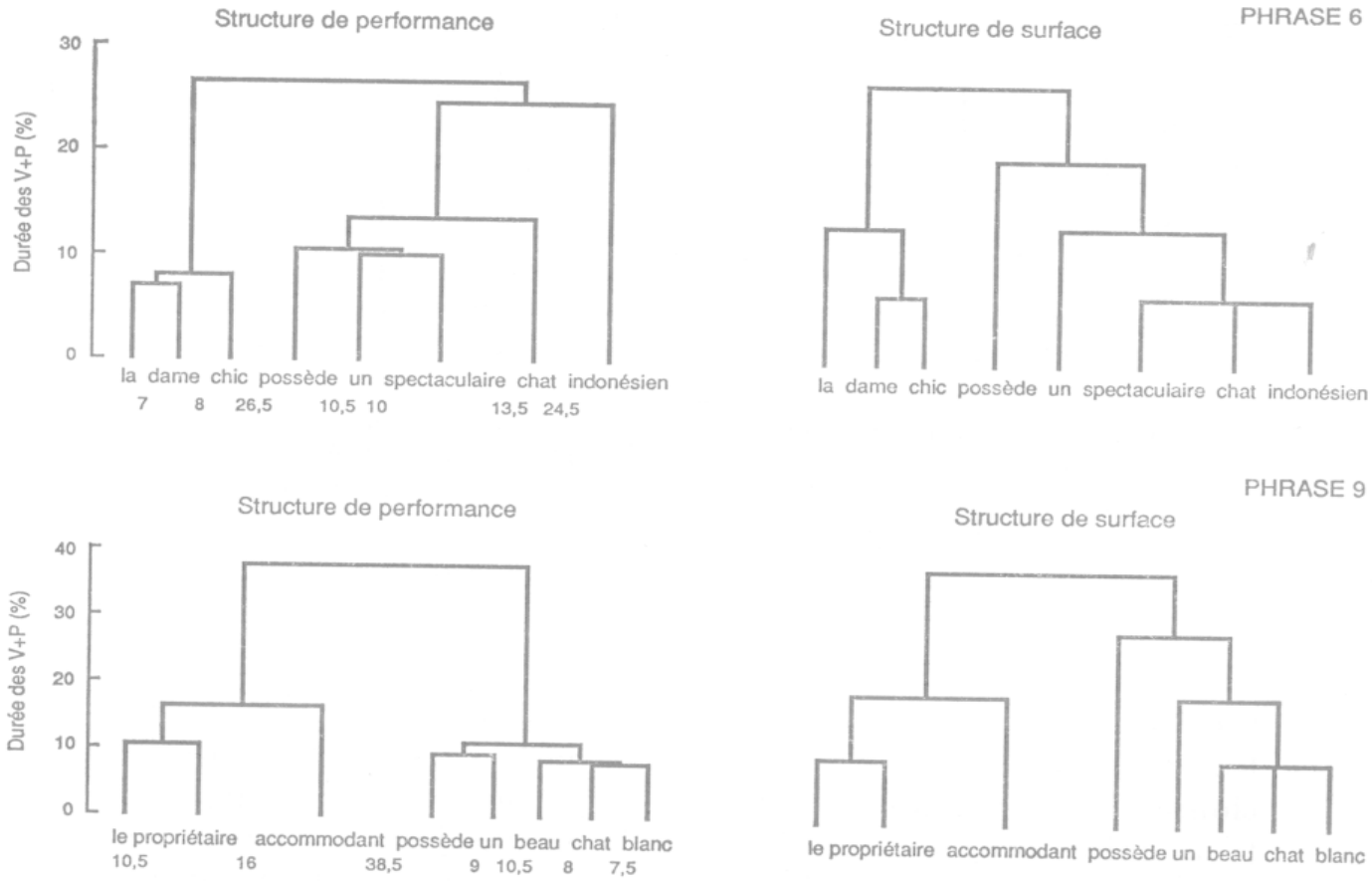


Fig. 2. — Structure de surface et structure de performance des phrases 6 et 9. Les structures de performance sont élaborées à partir des durées des V + P converties en pourcentages

*The surface structure and performance structure of sentences 6 and 9. The performance structures are based on percentage vowel + pause durations*

F.

particulier. En effet, dans la phrase 1 (fig. 1, structure du bas), on observe une V + P importante entre « fille » et « espiègle » et dans la phrase 6 (fig. 2, structure du haut à gauche) on remarque une V + P plus longue entre le substantif (« chat ») et l'adjectif postposé (« indonésien ») qu'entre l'adjectif préposé (« spectaculaire ») et le substantif, bien que les deux adjectifs aient la même longueur. Notons que l'importance de la coupure entre le substantif et l'adjectif postposé, dont nous devrons tenir compte lors de la prédiction des structures de performance, peut être réduite lorsque d'autres variables entrent en jeu, telles que la longueur de l'adjectif (voir « blanc » dans la phrase 9, fig. 2), le lien sémantique qui existe entre l'adjectif et le substantif, et la nature des phonèmes présents de part et d'autre de la coupure (deux voyelles, une consonne et une voyelle, etc.).

Un deuxième aspect qui ressort des structures de performance du français concerne le rattachement de certains mots grammaticaux. Normalement ceux-ci sont rattachés prosodiquement aux mots lexicaux de droite comme dans le cas de « une » qui est rattaché à « jolie » dans la phrase 1 (fig. 1). Cependant, il semblerait que dans certains cas les mots grammaticaux peuvent se rattacher aussi aux mots lexicaux de gauche. Ainsi, dans la phrase 9 (fig. 2, structure du bas à gauche), « un » est rattaché prosodiquement à « possède » et non, comme on pouvait s'y attendre, à « beau ». Nous devrons tenir compte de cette particularité dans notre algorithme de prédiction.

## 2. PRÉDICTION DES STRUCTURES DE PERFORMANCE

Dans un premier temps, nous avons tenté d'utiliser les méthodes de prédiction employées dans les travaux sur l'anglais moyennant quelques légères modifications des algorithmes. Nous avons d'abord corrélé les valeurs des V + P obtenues aux indices de complexité syntaxique (IC) des phrases (voir Grosjean *et al.*, 1979, pour la procédure d'indexation) et avons ensuite appliqué chacun des trois algorithmes utilisés jusqu'à maintenant (GGL, CPC et PHI ; voir Gee et Grosjean, 1983) pour obtenir des indices de prédiction que nous avons ensuite corrélés avec les valeurs des V + P (voir l'annexe 1 pour une description succincte des trois algorithmes),

Comme nous pouvons le constater dans la figure 1 et dans la figure 2 (phrase 6, en haut) la seule structure syntaxique de la phrase n'est pas normalement un très bon prédicteur de la structure de performance. En effet, la moyenne des corrélations entre les IC et les V + P pour l'ensemble de nos phrases n'est que de 0,68 (elle était de 0,76 en anglais ; Grosjean *et al.*, 1979). Lorsqu'on applique les algorithmes de prédiction développés pour l'anglais, les résultats sont meilleurs mais pas encore réellement satisfaisants. L'algorithme GGL prédit assez bien la structure hiérarchique des constituants prosodiques majeurs mais moins bien l'organisation interne des constituants mineurs (la moyenne des corrélations n'est donc que de 0,78). L'algorithme CPC arrive à prédire assez bien la coupure prosodique principale et cela pour l'ensemble des 9 phrases, mais la composante de bissection, qui ne s'applique qu'une seule fois pour l'ensemble de nos phrases, engendre de mauvais indices au sein de syntagmes longs (la moyenne des corrélations est donc de 0,57 ce qui est significativement moins élevé que la moyenne obtenue avec les IC ; $t = 3{,}29$, $p < 0{,}01$). Quant à l'algorithme PHI adapté au français[6], il prédit relativement bien les coupures prosodiques majeures mais prévoit parfois un indice trop élevé après le verbe lorsque celui-ci est suivi d'un mot grammatical qui se rattache prosodiquement à lui (la moyenne des corrélations est de 0,83). Notons que les moyennes obtenues avec GGL et PHI sont significativement différentes de la moyenne des corrélations calculées avec l'IC ($t = 3{,}04$, $p < 0{,}01$ et $t = 2{,}95$, $p < 0{,}01$, respectivement) mais quelles ne sont pas significativement différentes l'une de l'autre ($t = 1{,}36$, NS).

Face aux manquements des algorithmes existants, nous avons décidé de développer un nouvel algorithme (l'algorithme MG) en nous basant sur les données obtenues, certains principes de l'algorithme PHI et les travaux de Bailly (1989), Pasdeloup (1990) et Bachenko et Fitzpatrick (1990). Bailly (1989) met l'accent sur l'importance du rythme dans la production d'énoncés en français. Il montre que le découpage prosodique des énoncés est fonction de l'importance de la frontière syntaxique et du

6. Pour le français, il ne nous a pas été possible d'appliquer toutes les règles de l'algorithme PHI. La tête d'un syntagme occupe parfois une position plus à gauche (ex. « la petite fille espiègle ») et nous avons décidé, entre autres, de considérer comme syntagme phonologique individuel les adjectifs postposés à un nom.

nombre de syllabes des différents groupes prosodiques. Par conséquent, le mot prosodique, qui est une unité qui ne comporte qu'un seul accent, est constitué de quatre syllabes au maximum. Pasdeloup (1990), dans le même sens que Bailly, souligne l'alternance rythmique des groupes accentuels et met en évidence le fait que la durée des syllabes d'un mot varie selon leur position dans celui-ci. En outre, elle montre que les durées syllabiques augmentent en allant de gauche à droite et que la durée de la syllabe porteuse de l'accent est la plus longue. Bachenko et Fitzpatrick (1990), quant à eux, insistent sur le fait que la syntaxe joue un rôle déterminant mais pas suffisant dans la structure prosodique de l'anglais. Ils montrent également l'importance de la longueur des mots lexicaux dans la formation des mots phonologiques. L'algorithme que nous proposons ne traite pour l'instant que de phrases simples et devra donc être complété dans l'avenir.

## A) Les étapes de l'algorithme MG

Afin d'illustrer le déroulement de l'algorithme, nous utiliserons la phrase suivante : « La fille s'est déguisée en une jolie petite fée espiègle. » Pour des cas particuliers, nous ferons appel à d'autres phrases.

1) *Identification des noyaux de constituants prosodiques.* — En procédant de gauche à droite, identifier les noms, les verbes principaux et les adjectifs postposés. Ceux-ci reçoivent le statut de noyau de constituant prosodique et sont dominés par ∅.

Ex. « La (fille) s'est (déguisée)
en une jolie petite (fée) (espiègle). »

2) *Formation des constituants prosodiques de base par regroupement des mots autour des noyaux.* — La formation des constituants prosodiques se fait selon les sous-étapes suivantes.

*a) Mots lexicaux.* — Lorsque le noyau est précédé de mots lexicaux (substantifs, verbes, adjectifs, adverbes) qui font partie du même constituant syntaxique de premier niveau que le noyau, regrouper les mots lexicaux et le noyau sous un même nœud ∅ au moyen d'un branchement à droite. Ils forment ainsi un

constituant prosodique plus large. Nous entendons par « premier niveau » le niveau des nœuds terminaux.

Ex. « La (fille) s'est (déguisée)
en une (jolie (petite (fée))) (espiègle). »

*b) Rattachement des mots grammaticaux.* — L'auxiliaire est toujours rattaché à l'aide d'un # au verbe situé à droite. Quant aux autres mots grammaticaux, ils sont rattachés à l'aide d'un # au noyau ou constituant prosodique de droite dans tous les cas sauf dans celui où un mot grammatical seul, rattaché au noyau ou constituant prosodique de gauche, formerait un nouveau constituant d'une longueur *inférieure* à quatre syllabes.

Ex. de la situation normale :

« (La#(fille)) (s'est#(déguisée))
(en#(une#(jolie (petite (fée))))) (espiègle). »

Ex. de l'exception :

« Le professeur sympathique
((possède)#un) (joli (chat)) (tigré). »

3) *Indexation des constituants prosodiques de base.* — Attribuer à chaque frontière marquée par un # un indice de hiérarchie égal à un. Au sein des constituants prosodiques, attribuer à chaque frontière séparant deux mots lexicaux un indice de hiérarchie correspondant à l'indice syntaxique de cette frontière[7].

Ex. « (La#(fille)) (s'est#(déguisée))
1 1
(en#(une#(jolie (petite (fée))))) (espiègle). »
1 1 2 1

4) *Formation des constituants prosodiques supérieurs.*

*a) Regroupement des constituants prosodiques à gauche et à droite du ∅ verbe.* — En procédant par niveau de hiérarchie, de bas vers le haut, rassembler les ∅ qui appartiennent au même constituant syntaxique sous un même nœud ∅ au moyen d'un

7. Les indices syntaxiques correspondent au nombre de nœuds non terminaux de chaque côté de la frontière en question en remontant et en incluant le nœud dominant.

branchement à droite. Ce rassemblement nous donne des constituants prosodiques supérieurs (CPS) à gauche et à droite du verbe.

Ex. « (La#(fille)) (s'est#(déguisée))
1 1
((en#(une#(jolie (petite (fée))))) (espiègle)). »
1 1 2 1

*b) Rattachement du ∅ verbe.* — Regrouper le ∅ verbe au CPS de droite sauf si le nombre d'embranchements de la structure formée par le ∅ verbe et le CPS de gauche est inférieur ou égal au nombre d'embranchements du *seul* CPS à droite (sans compter le ∅ verbe donc),

Ex. d'un regroupement du ∅ verbe à droite :

« ((Le#professeur) (sympathique))
(((possède)#(un)) ((joli (chat)) (tigré))). »

Ex. d'un regroupement du ∅ verbe à gauche :

« ((La#(fille)) (s'est#(déguisée)))
1 1
((en#(une#(jolie (petite (fée))))) (espiègle)). »
1 1 2 1

*c) Regroupement final.* — Terminer l'arbre prosodique à l'aide d'un ∅ dominant.

Ex. « (((La#(fille)) (s'est#(déguisée)))
1 1
((en#(une#(jolie (petite (fée))))) (espiègle))). »
1 1 2 1

5) *Indexation des constituants prosodiques supérieurs.* — Pour les frontières n'ayant pas encore reçu un indice de hiérarchie, procéder de la manière suivante : attribuer à chaque frontière séparant deux ∅ un indice de hiérarchie correspondant au nombre de nœuds immédiatement à gauche et à droite de cette frontière en remontant et en incluant le nœud dominant.

Ex. « (((La#(fille)) (s'est#(déguisée)))
1 3 1 5
((en#(une#(jolie (petite (fée))))) (espiègle))). »
1 1 2 1 3

*Cas spécial.* — Si immédiatement à gauche d'une frontière séparant deux ∅ se trouve un mot grammatical, cette frontière reçoit automatiquement un indice de hiérarchie égal à un.

Ex. : « ((Le#professeur) (sympathique))
(((possède)#(un)) ((joli (chat)) (tigré))). »
1

6) *Ajustements :*

*a) Longueur des constituants et des mots.*

1) *Longueur sujet-verbe :* si le nombre de syllabes des mots lexicaux du SN sujet et du verbe (en incluant l'auxiliaire) est inférieur ou égal au nombre de syllabes des mots lexicaux du syntagme syntaxique (et non prosodique) qui suit le verbe, alors l'indice de hiérarchie de la frontière séparant le SN sujet du verbe est diminué de 1.

Ex. « ((La#(fille)) (s'est#(déguisée)))
en une jolie petite fée espiègle. »
2

2) *Longueur des mots lexicaux au sein d'un même ∅ de base :* si le nombre de syllabes de deux mots lexicaux connexes est supérieur à quatre syllabes, alors l'indice de hiérarchie entre ces deux mots est augmenté de un.

Ex. « La dame chic
possède (un#(spectaculaire (chat))) indonésien. »
2

3) *Longueur des mots lexicaux entre deux ∅ de base au sein d'un même syntagme syntaxique :* si le nombre de syllabes des mots lexicaux à l'intérieur d'un ∅ de part et d'autre d'une frontière séparant deux ∅ :

— est inférieur à huit syllabes mais supérieur ou égal à quatre syllabes, alors l'indice de hiérarchie à cette frontière est diminué de un.

Ex. « La fille s'est déguisée
((en#(une#(jolie (petite (fée))))) (espiègle)). »
2

— est inférieur à quatre syllabes, alors la valeur de l'indice à la frontière est diminuée de deux.

Ex. « ((La #(dame)) (chic))
1
possède un spectaculaire chat indonésien ».

*b) Expression figée.* — Si deux mots lexicaux forment une locution reconnue (« salle de bain », « petite fille », etc.), alors la valeur de l'indice de hiérarchie séparant ces deux mots est considérée comme nulle.

Ex. : « (La #(jolie (petite (fille)))) s'est déguisée
0
en une fée espiègle. »

*c) Fin de phrase.* — Lorsque les mots lexicaux des deux derniers ø de la phrase ont ensemble un nombre de syllabes supérieur à quatre, alors l'indice à la frontière des deux ø est augmenté de un.

Ex. « La fille s'est déguisée
((en #(une #(jolie (petite (fée))))) (espiègle)). »
3

La figure 3 montre la structure obtenue pour la phrase « La fille s'est déguisée en une jolie petite fée espiègle » au moyen de l'algorithme MG.

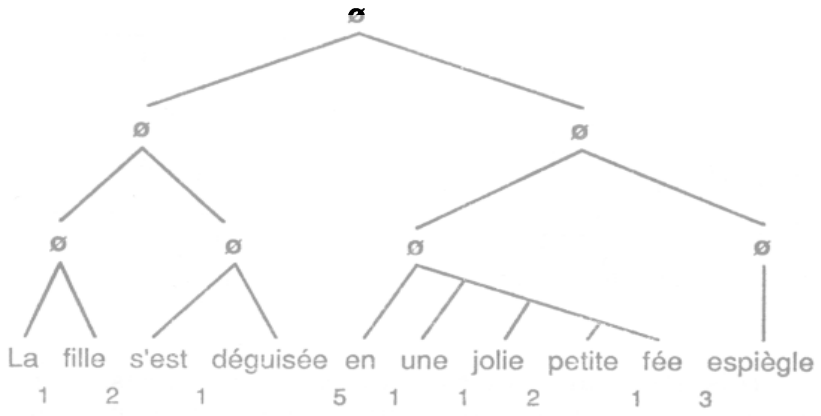


Fig. 3. — Structure de performance de la phrase 1 prédite par l'algorithme MG et indices de complexité

*The performance structure of sentence 1 predicted by the MG algorithm and its complexity indices*

psy_0003-5033_1993_num_93_1 /F.

## B) Prédiction de l'algorithme MG

L'algorithme MG est un bon prédicteur des structures de performance de nos phrases expérimentales. En effet, sur l'ensemble des phrases, nous avons obtenu une corrélation moyenne de 0,94. Celle-ci est significativement différente des corrélations moyennes obtenues avec CPC (0,57 ; $t = 7{,}18$, $p < 0{,}001$), GGL (0,78 ; $t = 3{,}53$, $p < 0{,}01$) et PHI (0,83, $t = 3{,}36$, $p < 0{,}01$). L'algorithme MG rend compte des différentes propriétés générales des structures de performance (hiérarchie, symétrie, unités de base de longueur plus ou moins égale) et traite également d'autres

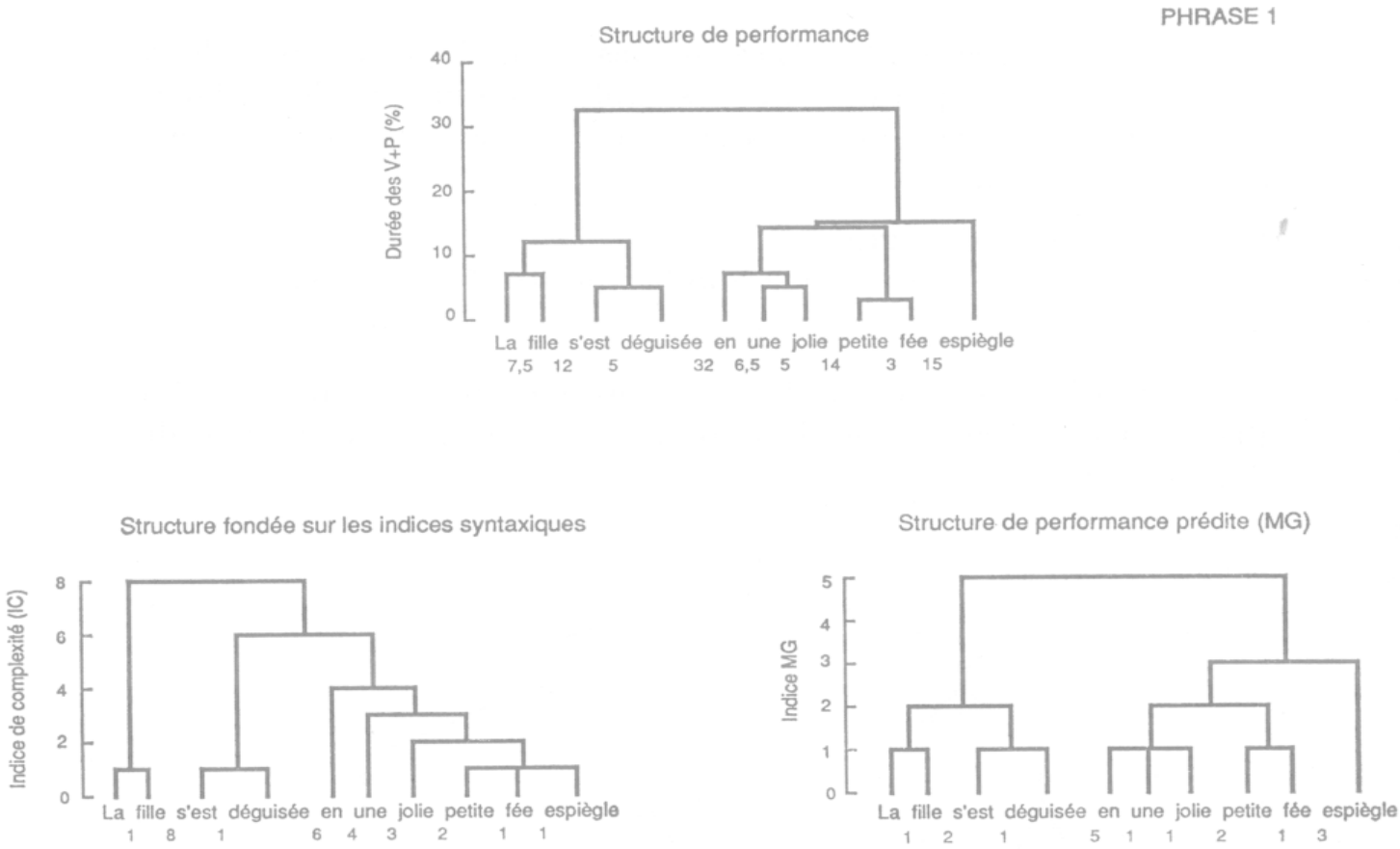


Fig. 4. — Trois structures hiérarchiques de la phrase 1 : la structure de performance élaborée à partir des durées des V + P converties en pourcentages, la structure fondée sur les indices de complexité syntaxique et la structure de performance prédite à l'aide de l'algorithme MG.

*Three hierarchical structures for sentence 1 : the performance structure based on percentage vowel + pause durations, the syntactic structure based on complexity indices and the performance structure predicted by the MG algorithm.*

F

aspects que nous avons trouvés spécifiquement dans notre étude : le cas de l'adjectif postposé et celui du rattachement à gauche de certains mots grammaticaux, entre autres. Comme nous le voyons sur la figure 4, l'algorithme MG, contrairement à la seule structure syntaxique (structure en bas à gauche), prédit correctement non seulement les frontières prosodiques majeures mais également les frontières mineures. En effet, on remarque que l'algorithme MG engendre un découpage prosodique comparable à celui observé dans les structures de performance, c'est-à-dire une segmentation de la phrase en un certain nombre d'unités prosodiques (« la fille », « s'est déguisée », « en une jolie », « petite fée » et « espiègle ») qui ne sont pas forcément identiques à des unités syntaxiques telles qu'elles sont engendrées par la grammaire.

La prédiction satisfaisante de l'algorithme est expliquée en grande partie par le fait que nous avons pu bénéficier de l'apport de PHI : la procédure à utiliser pour la constitution et l'indexation des constituants prosodiques de base et des constituants prosodiques supérieurs, la méthode de rattachement des mots grammaticaux aux noyaux et du constituant du verbe aux constituants à droite ou à gauche de celui-ci, l'ajustement pour les mots lexicaux longs et les fins de phrase, etc. Notre apport a été de construire les arbres de prédiction de manière différente (voir les premières étapes), de mettre plus d'emphase sur l'équilibre en longueur des unités prosodiques et de mettre en valeur le statut spécial des adjectifs postposés et de certains mots grammaticaux qui se rattachent à gauche plutôt qu'à droite. Au sujet des adjectifs postposés, et d'un point de vue formel, il n'est pas possible de les considérer comme des noyaux, c'est-à-dire des têtes de syntagme. Dans la théorie syntaxique générative (convention X barre) les adjectifs postposés ont comme tête un nom. Cependant, dans notre algorithme, nous avons délibérément choisi de considérer les adjectifs postposés comme des têtes de syntagme. Ce choix est basé sur des considérations avant tout empiriques car, dans la majorité des cas, on observe une pause plus ou moins marquée entre le constituant prosodique et l'adjectif postposé à ce constituant. De ce fait, il semble que prosodiquement, ces adjectifs se comportent comme des syntagmes prosodiques (s'ils sont assez longs, bien entendu).

Le rattachement à gauche de certains mots grammaticaux prête également à discussion. D'un point de vue syntaxique, les

mots grammaticaux sont rattachés aux mots lexicaux de droite. En effet, sur le plan théorique, le découpage d'un syntagme verbal du type « possède un magnifique chat » en deux unités prosodiques « possède un » et « magnifique chat » n'est pas possible. Cependant nous avons remarqué que du point de vue prosodique les sujets avaient tendance à rattacher le mot grammatical « un » au verbe « possède » lorsque les mots lexicaux suivants étaient nombreux et/ou longs. Nous avons donc tenu compte de ce comportement prosodique dans notre algorithme. On notera également dans l'algorithme une série d'ajustements (étape 6) qui tiennent compte de la longueur des mots et des constituants, de l'allongement produit en fin de phrase et du caractère figé de certaines expressions. Ces ajustements, qui ont un caractère encore légèrement *ad hoc*, devront être étudiés de plus près avant de trouver une place centrale dans l'algorithme. Leur inclusion ne change en fait pas grand-chose aux résultats que nous avons obtenus. En effet, en excluant l'étape des ajustements (l'étape 6), notre algorithme prédit encore très bien les données obtenues : $r = 0{,}89$. Cette moyenne est encore significativement différente des résultats obtenus à l'aide des trois autres algorithmes (GGL : $t = 2{,}65$, $p < 0{,}05$ ; PHI : $t = 3{,}05$, $p < 0{,}01$ ; CPC : $t = 5{,}55$ $p < 0{,}001$) mais pas de la moyenne calculée lorsque la sixième étape est inclue (0,94 ; $t = 1{,}62$, NS).

## CONCLUSION

La caractérisation des structures de performance en français fait appel à des caractéristiques prosodiques plus complexes que pour l'anglais. Déjà dans l'obtention de la mesure dépendante (durée V + P), nous avons dû prendre en considération non seulement la durée de la pause silencieuse interlexicale, comme cela a toujours été le cas dans les études sur l'anglais, mais également la durée de la voyelle de la syllabe en fin de mot. Bien que les structures de performance du français partagent les trois mêmes propriétés fondamentales que celles observées en anglais (symétrie, hiérarchie et unités prosodiques de longueur plus ou moins égale), nous avons pu mettre en valeur d'autres caractéristiques tout à fait intéressantes : le statut de noyau prosodique de l'adjectif postposé et le rattachement à gauche de certains mots grammaticaux. L'algorithme MG que nous

proposons tient compte des variables impliquées et prédit donc bien les résultats que nous avons obtenus dans notre étude. Il sera important à l'avenir de tester cet algorithme sur de nouvelles phrases où l'on varierait la longueur et la complexité syntaxique et sémantique. Il est évident qu'il devra subir des modifications parfois importantes avant de pouvoir prédire la structure de performance d'un large ensemble de phrases françaises.

Deux domaines de recherche sont susceptibles de bénéficier des travaux portant sur les structures de performances. D'une part, en traitement automatique de la parole, une meilleure description des structures prosodiques du français oral et le développement d'algorithmes de prédiction peuvent contribuer à améliorer la qualité de la synthèse à partir du texte, comme cela a déjà été le cas en anglais (voir Bachenko et Fitzpatrick, 1990 ; Quené et Kager, 1992 ; Wang et Hirschberg, 1992). D'autre part, en psycholinguistique, une compréhension approfondie de ces structures aura des conséquences sur l'élaboration des modèles de production et de perception de la parole. Du point de vue de la production, nous pourrons spécifier davantage, dans les étapes de production d'un énoncé, celle qui a trait à la phonétique. Dans le modèle de Garrett (Garrett, 1980), par exemple, une révision du niveau phonétique préciserait non seulement la forme phonologique des mots en fonction de leur environnement syllabique (ex. : « bel abricot » au lieu de « beau abricot »), mais spécifierait également la structure prosodique des énoncés. C'est à ce niveau que le regroupement en syntagmes phonologiques se ferait, ce qui produirait, pour la phrase « Jean a mangé un bel abricot rouge », la structure prosodique suivante : ((Jean) (a mangé) ((un bel abricot) (rouge))). Quant à la perception, il faudra expliquer le rôle que joue la structure prosodique de l'énoncé dans le traitement de la parole, ou tout au moins, la fonction des éléments qui la composent (syllabes plus ou moins accentuées, syntagmes phonologiques, etc.). Il se pourrait, par exemple, que l'analyse se fasse par syntagmes phonologiques (voir Grosjean et Gee (1987) pour une suggestion similaire pour l'anglais). L'auditeur isolerait chaque syntagme en localisant tout d'abord la tête de celui-ci (il s'agit toujours d'un mot lexical à forte charge sémantique et marqué par la prosodie) et le décomposerait ensuite en ses éléments lexicaux afin de pouvoir effectuer l'analyse syntaxique, sémantique et pragmatique. Sans vouloir trop nous avancer sur ce terrain, faute de modélisation adéquate

et d'études expérimentales suffisantes, nous pouvons au moins remarquer qu'il serait surprenant que la structure prosodique de la phrase soit uniquement un phénomène de production et qu'elle n'ait aucun rôle à jouer en perception.

ANNEXE *1. — Les algorithmes GGL, CPC et PHI*

*Algorithme GGL.* — Cet algorithme, proposé par Grosjean, Grosjean et Lane (1979), repose sur l'idée que l'importance d'une coupure interlexicale particulière dépend de l'importance relative de cette coupure en termes, d'une part, de structure de surface de la phrase et, d'autre part, de la position de cette coupure dans le constituant analysé — plus elle est proche du milieu du constituant, plus cette coupure est importante. Ainsi cet algorithme tient compte de deux exigences quelquefois contradictoires : la tendance du locuteur à respecter la structure linguistique de la phrase ainsi que sa tendance à équilibrer la longueur des constituants lors de leur production. Pour en tenir compte, l'algorithme fonctionne de façon cyclique en combinant pour chaque frontière de mot dans la phrase, un indice de complexité linguistique (IC), fondé sur le nombre de nœuds dominés par le nœud correspondant à la frontière de mot en question, et une mesure de la distance qui sépare cette frontière de mot du milieu du segment considéré.

*Algorithme CPC.* — L'algorithme élaboré par Cooper et Paccia-Cooper (1980) partage de nombreuses caractéristiques avec l'algorithme GGL. Il en diverge néanmoins sur deux points importants. D'abord, il donne plus de poids à l'indice de complexité syntaxique que n'en donne l'arbre de surface lui-même. Au lieu de refléter simplement l'indicateur syntagmatique, comme le ferait le IC (voir ci-dessus), son indice est un véritable algorithme à lui seul, et l'arbre ainsi obtenu diffère largement de l'indicateur syntagmatique de départ. La seconde différence est que la composante de bissection ne s'applique qu'aux constituants dont la taille est égale ou supérieure à sept mots lexicaux. Pour la plupart des phrases de taille moyenne, donc, le processus de bissection ne s'applique qu'une seule fois, alors qu'il s'applique de façon itérative dans l'algorithme GGL.

*Algorithme PHI.* — L'algorithme PHI (Gee et Grosjean, 1983) tient compte de la syntaxe mais intègre également des informations phonologiques et prosodiques. Il fonctionne de gauche à droite dans la phrase, en procédant par syntagmes phonologiques. En se servant de la tête d'un syntagme syntaxique, il isole les éléments qui font partie du syntagme phonologique, délexicalise les mots grammaticaux et dérive à droite les éléments de chaque syntagme. Puis, par une série de règles (règle du constituant, règle du verbe, règle générale), il effectue un regroupement des syntagmes phonologiques en syntagmes intonationnels.

## *RÉSUMÉ*

***Il est possible, grâce à diverses tâches expérimentales, d'obtenir la structure de performance d'une phrase. Celle-ci est caractérisée par un certain nombre de propriétés fondamentales et peut être prédite à l'aide d'algorithmes divers. Jusqu'à présent les recherches ont porté principale-***

*ment sur l'anglais et nous nous tournons donc dans cette étude vers les structures de performance du français. Dans la première partie nous caractérisons les structures obtenues à l'aide d'une tâche de lecture orale, et dans la seconde partie nous prédisons ces structures au moyen d'algorithmes déjà existants (mais légèrement modifiés) et d'un nouvel algorithme que nous avons développé spécifiquement pour le français.*

*Mots clés : psycholinguistique, production, parole, structures de performance, algorithmes de prédiction.*

BIBLIOGRAPHIE

Aubergé V. — (1991) *La synthèse de la parole : « des règles aux lexiques »*, Thèse de doctorat, Grenoble, Universités Stendhal et P. Mendès-France.

Bachenko J., Fitzpatrick E. — (1990) A computational grammar of discourse-neutral prosodic phrasing in English, *Computational Linguistics, 16*, 155-170.

Bailly G. — (1989) Integration of rhythmic and syntactic constraints in a model of generation of French prosody, *Speech communication, 8*, 137-146.

Caelen-Haumont G. — (1991) *Stratégies des locuteurs et consignes de lecture d'un texte : analyse des interactions entre modèles syntaxique, sémantique, pragmatique et paramètres prosodiques*, Thèse de doctorat d'Etat, Aix-en-Provence, Université de Provence.

Cooper W., Paccia-Cooper J. — (1980) *Syntax and speech*, Cambridge (MA), Harvard University Press.

Dommergues J.-Y., Grosjean F. — (1981) Performance structures in the recall of sentences, *Memory and Cognition, 9*, 478-486.

Duez D. — (1987) *Contribution à l'étude de la structuration temporelle de la parole en français*, Thèse de doctorat d'Etat, Aix-en-Provence, Université de Provence.

Garrett M. — (1980) Levels of processing in sentence production, *in* B. Butterworth (Edit.), *Language production*, Londres, Academic Press, 177-220.

Gee J. P., Grosjean F. — (1983) Performance structures : A psycholinguistic and linguistic appraisal, *Cognitive Psychology, 15*, 411-458.

Grosjean F., Dommergues J.Y. — (1983) Les structures de performance en psycholinguistique, *L'Année Psychologique, 83* (2), 513-536.

Grosjean F., Gee J. — (1987) Prosodic structure and spoken word recognition *Cognition, 25*, 135-155.

Grosjean F., Grosjean L., Lane H. — (1979) The patterns of silence : Performance structures in sentence production, *Cognitive Psychology, 11*, 58-81.

Hirst D. — (1987) La *description linguistique des systèmes prosodiques : une approche cognitive*, Thèse de doctorat d'Etat, Aix-en-Provence, Université de Provence.

Levelt W. — (1989) *Speaking : From intention to articulation*, Cambridge (MA), MIT Press.

Llorca R. — (1984) *Eléments d'analyse du rythme de la parole en français*, Thèse de doctorat, Besançon, Université de Franche-Comté.

Pasdeloup V. — (1990) *Modèle de règles rythmiques du français appliqué à la synthèse de la parole*, Thèse de troisième cycle, Aix-en-Provence, Université de Provence.

Quené H., Kager R. — (1992) The derivation of prosody for text-to-speech

from prosodic sentence structure, *Computer Speech and Language*, *6*, 77-98.
Rossi M. — (1985) Intonation et organisation de l'énoncé, *Phonetica*, *42*, 2-3, 135-153.
Vaissière J. — (1980) La structuration acoustique de la phrase française, *Annali della Scuola Normale Superiore di Pisa*, *10*, 529-560.
Wang M., Hirschberg J. — (1992) Automatic classification of intonational phrase boundaries, *Computer Speech and Language*, *6*, 175-196.

(Accepté le 9 décembre 1992.)