

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 principalement sur l'anglais et nous nous tournons donc dans cette étude vers les structures de performance du français en caractérisant les structures obtenues à l'aide d'une tâche de lecture orale.

1. 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 & Lane, 1979 ; Dommergues & Grosjean, 1981 ; Gee & Grosjean, 1983). Cette structure, que l'on a nommée "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 & 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 & Fitzpatrick, 1990 ; Quené & Kager, 1992 ; Wang & Hirschberg, 1992).

Obtention des structures de performance

Un certain nombre d'études sur l'anglais ont montré que la fréquence et la durée des pauses entre les constituants d'une phrase sont liées à sa structure syntaxique, mais aussi à des aspects non syntaxiques. Dans une première étude, Grosjean, Grosjean & Lane (1979) ont mis en évidence les structures de performance de l'anglais à l'aide des pauses insérées entre les mots d'une phrase lors de la lecture lente puis ont tenté de prédire ces structures.

Pour construire "l'arbre de performance" à partir des données expérimentales, les auteurs ont employé la procédure itérative suivante :

- a) Prendre la pause la plus courte dans la phrase.
- b) Relier les deux mots (ou groupes de mots) séparés par cette pause sous un nœud commun.
- c) Ne plus tenir compte de cette pause et répéter le processus jusqu'à ce que toutes les pauses soient effacées.

Quant à la hauteur des nœuds, elle est proportionnelle à la durée des pauses considérées : plus la pause est longue plus le nœud est élevé.

Les résultats ont montré que les sujets ne groupaient pas automatiquement le verbe avec le syntagme nominal qui suit comme une structure syntaxique traditionnelle le prédirait mais que dans bien des cas, une coupure principale était observée entre le syntagme verbal et l'objet.

D'autre part, cette étude a montré que les structures de performance obtenues ne sont pas le résultat de contraintes de respiration. En effet, la même expérience conduite avec un autre groupe de 14 sujets mais sans que les sujets reprennent leur souffle au cours de l'enregistrement a fourni des résultats pratiquement identiques. Grosjean, Grosjean & Lane ont également mis en évidence le fait que ces structures sont invariables d'une tâche à l'autre. Un troisième groupe de sujets avaient pour tâche de trouver une coupure principale dans chacune des phrases, de la numéroter du chiffre 1 puis en considérant indépendamment les deux parties obtenues, de les segmenter à leur tour en numérotant la frontière la plus

importante du chiffre 2 et ainsi de suite jusqu'à ce que chaque frontière de mot soit numérotée.

Les structures hiérarchiques obtenues à partir de ces valeurs de segmentation ont permis de constater que les résultats sont très comparables. Il y a une forte corrélation entre les indices de segmentation et les durées de pauses converties en pourcentages.

Propriétés des structures de performance

Trois propriétés principales ressortent de l'examen des structures de performance :

- a) Les structures prosodiques sont composées d'unités de base de longueurs plus ou moins égales. Sur l'ensemble des quatorze phrases utilisées dans l'étude de Grosjean, Grosjean & Lane, on obtient une moyenne de trois syllabes par groupe prosodique, le nombre de syllabes variant de un à huit.
- b) Les structures de performance sont organisées hiérarchiquement. Les unités de base se regroupent en unités de niveau supérieur qui, à leur tour, se regroupent au sein d'unités plus grandes.
- c) Les structures de performance sont plus ou moins symétriques. Cette propriété de symétrie fait que les structures de performance, au contraire des structures syntaxiques, sont rarement dérivées à droite.

Les structures de performance en français

Dans cette étude, nous nous tournons vers les structures de performance du français afin de déterminer dans quelle mesure les principes adoptés pour 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 les structures de performance du français oral. Nos objectifs sont, 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 et de quantifier l'importance des coupures qui existent entre elles.

Dans cette étude nous caractériserons les structures de performance à l'aide d'une tâche de lecture orale. Pour ce faire, nous utiliserons l'approche adoptée dans les études anglaises (voir Grosjean, Grosjean & Lane, 1979) et ceci pour des raisons de simplicité expérimentale. Nous avons donc choisi de calculer l'importance des frontières interlexicales afin de caractériser la structure prosodique de nos énoncés. 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.

2. Méthode

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

Matériaux : 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.

- 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 : Avant d'être présentées aux sujets, les phrases ont été tapées dans un ordre aléatoire sur des bandes de papier. 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 MacRecorder 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 l'approche utilisée pour l'anglais, où on ne mesure 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 consonnes¹. 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 phrase².

Avant même de construire les structures de performance 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 neuf 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 décrite dans l'introduction. (Nous avons simplement remplacé la pause par la valeur V+P). 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.

¹ 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.

² 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.

3. Résultats et discussion

La série de structures obtenue montre que les trois propriétés fondamentales des structures de performance qui ressortent en anglais (Grosjean & Dommergues, 1983 ; Gee & 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 base³. 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 (voir page suivante), 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 même figure, 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, 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 décomposé 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

³ 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 de base.

PHRASE 1

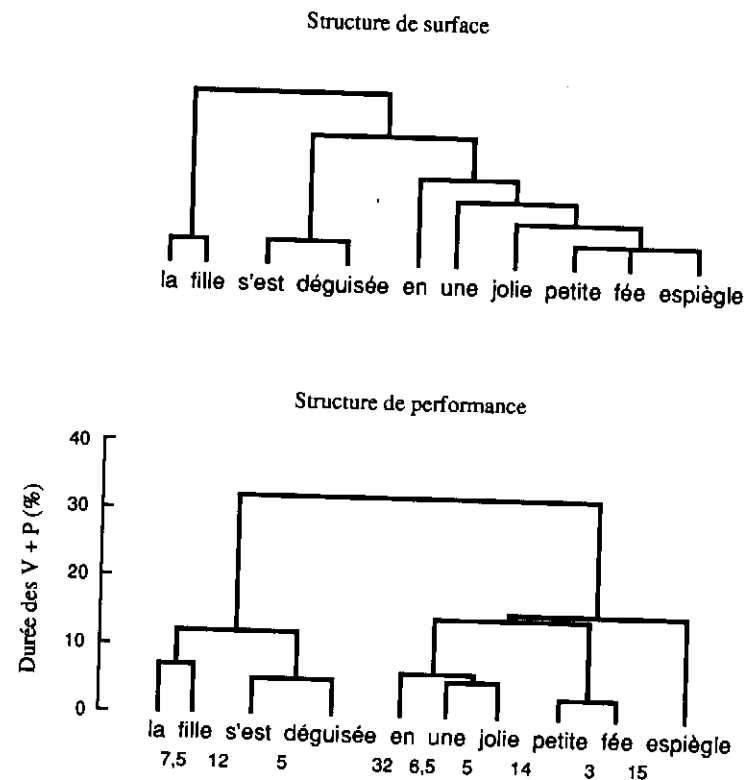


Fig. 1. - Structure de surface et structure de performance de la phrase 1. La structure de performance est élaborée à partir des durées des V + P converties en pourcentages.

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 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 "la dame chic possède un spectaculaire chat indonésien" et 9 "le

propriétaire accommodant possède un beau chat blanc", 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 particulier. En effet, dans la phrase 1 (figure 1, structure du bas), on observe une V+P importante entre "fille" et "espiègle" et dans la phrase 6 (figure 2, structure du haut à gauche ; voir page suivante) on remarque une V+P plus longue (24,5) entre le substantif ("chat") et l'adjectif postposé ("indonésien") qu'entre l'adjectif préposé ("spectaculaire") et le substantif (13,5), 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 devons 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 (V+P de 7,5 entre "chat" et "blanc" dans la phrase 9), 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 (figure 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 (figure 2), "un" est rattaché prosodiquement à "possède" et non, comme on pouvait s'y attendre, à "beau".

4. 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

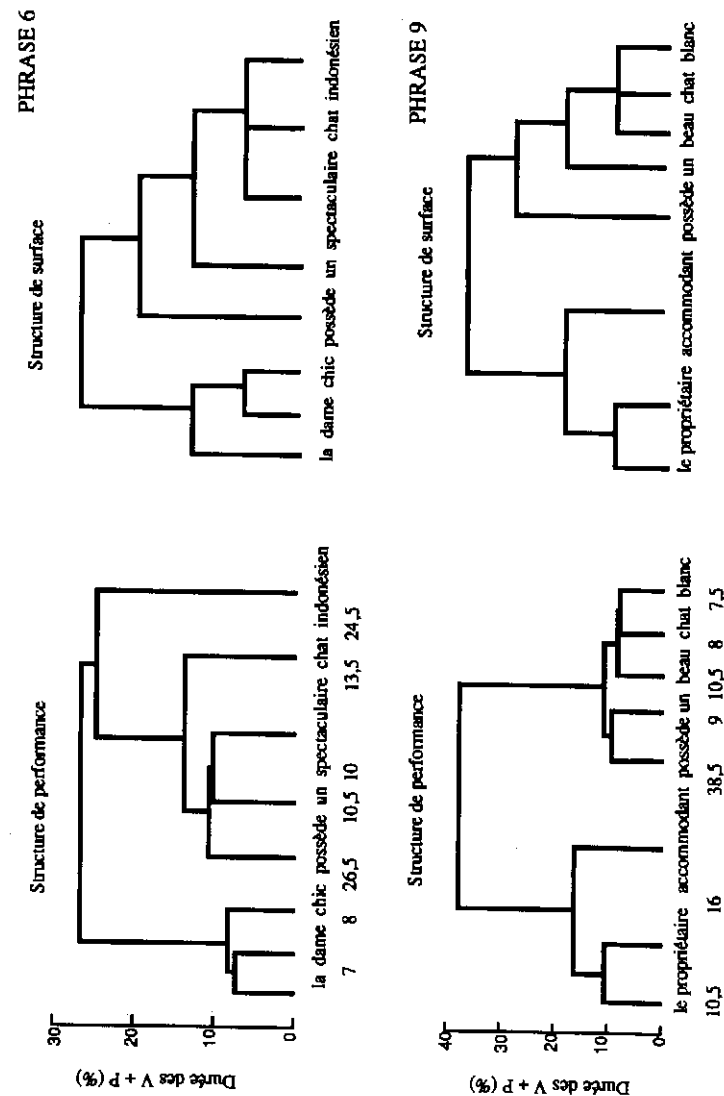


Fig. 2. - Structure de surface et structure de performance des phrases 6 et 9. Les structures de performance sont élaborées à partir des durées de V + P converties en pourcentages.

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.

Dans la continuation de cette étude, nous avons élaboré un algorithme de prédiction⁴ (algorithme MG) permettant de prédire nos structures de performance. Il s'avère être un bon prédicteur des structures de performance des phrases expérimentales et il rend compte des différentes propriétés générales des structures (hiérarchie, symétrie, unités de base de longueur plus ou moins égale). De plus, il traite d'autres 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.

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 & Fitzpatrick, 1990 ; Quené & Kager, 1992 ; Wang & 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.

5. Bibliographie

AUBERGE, V. (1991): *La synthèse de la parole : "des règles aux lexiques"*, Thèse de doctorat, Universités Stendhal et P. Mendès-France.

⁴ Voir Monnin, P. & F. Grosjean (1993): "Les structures de performance en français : caractérisation et prédiction", *L'année psychologique*, 93, 9-30.

BACHENKO, J. & E. FITZPATRICK (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, Université de Provence.

COOPER, W. & J. PACCIA-COOPER (1980): *Syntax and Speech*, Cambridge, MA, Harvard University Press.

DOMMERGUES, J.Y. & F. GROSJEAN (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, Université de Provence.

GEE, J.P. & F. GROSJEAN (1983): "Performance structures: A psycholinguistic and linguistic appraisal", *Cognitive psychology*, 15, 411-458.

GROSJEAN, F. & J.Y. DOMMERGUES (1983): "Les structures de performance en psycholinguistique", *L'année psychologique*, 83, 513-536.

GROSJEAN, F., L. GROSJEAN & H. LANE (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, Université de Provence.

- LEVELT, W. (1989): *Speaking: From Intention to Articulation*, Cambridge, MA, MIT Press.
- LLORCA, R. (1984): *Éléments d'analyse du rythme de la parole en français*, Thèse de doctorat, Université de Franche-Comté.
- PASDELOUP, V. (1990): *Modèle de règles rythmiques du français appliqué à la synthèse de la parole*, Thèse de troisième cycle, Université de Provence.
- QUENE, H. & R. KAGER (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.
- VAISSIERE, J. (1980): La structuration acoustique de la phrase française, *Annali della Scuola Normale Superiore di Pisa*, 10, 529-560.
- WANG, M. & J. HIRSCHBERG (1992): "Automatic classification of intonational phrase boundaries", *Computer Speech and Language*, 6, 175-196.