

L'accès au lexique de code-switchs chez le bilingue : effets de la densité et du contexte [◇]

Markus Leuenberger

Résumé

L'élaboration de modèles de reconnaissance de mots (accès au lexique) chez les bilingues a pris une certaine importance ces dernières années. Grosjean (1988), par exemple, a proposé un modèle d'activation interactive pour rendre compte de l'accès au lexique en mode de communication bilingue. L'objectif de la présente étude est de permettre d'affiner ce modèle par rapport aux effets contextuels. Nous avons fait varier la contrainte sémantique et la densité des code-switchs lors de la reconnaissance de code-switchs dans une tâche de gating appliquée à des sujets bilingues suisse-allemand - français. Les résultats significatifs obtenus pour les deux variables testées semblent montrer que la présence d'un code-switch précédent agit probablement comme information ascendante en augmentant le niveau d'activation de la langue la moins utilisée tandis que la contrainte sémantique active de manière descendante les mots de cette langue.

1. Introduction

Lors de la reconnaissance de mots, l'auditeur est confronté à deux tâches simultanées : d'une part, il doit analyser l'onde acoustique pour en extraire les unités linguistiques (phonèmes, syllabes) et d'autre part, à l'aide de ces unités, il doit identifier les mots énoncés. Reconnaître un mot n'est pas simplement une projection des caractéristiques acoustiques et phonétiques dans le lexique mental. C'est un processus compliqué qui comprend diverses étapes d'activation et d'inhibition, de correction, d'anticipation, etc. La recherche qui touche à la reconnaissance des mots se penche à la fois sur les effets qui rendent compte de l'accès au lexique et sur les modèles qui réunissent les effets en un tout cohérent.

[◇] Cette étude a pu être entreprise et menée à bien grâce à un subside du FNRS (12-33582.92)

Ce n'est que récemment que les chercheurs se sont tournés vers la reconnaissance de mots chez le bilingue. La réalité du bilingue est différente de celle du monolingue dans la mesure où le bilingue, dans sa vie de tous les jours, se trouve dans divers modes de communication. En effet, il peut passer d'un mode complètement monolingue à un mode bilingue. Dans le mode monolingue, le bilingue s'approprie la langue de l'interlocuteur monolingue et désactive, au mieux, son autre langue. Dans le mode bilingue, les interlocuteurs choisissent une des langues pour communiquer ensemble (c'est ce qu'on appelle la langue de base). Ensuite, tout en se servant de cette langue, ils peuvent introduire des éléments de l'autre langue (appelée également langue alternative). Une des possibilités est d'utiliser un code-switch, c'est-à-dire un mot, un syntagme ou une phrase de l'autre langue.

Bien que la plupart des modèles monolingues de reconnaissance pourraient être modifiés pour tenir compte de la réalité du discours bilingue, le type de modèle qui est pour l'instant le plus prometteur est celui de l'activation interactive. Grosjean (1988) propose un modèle de ce genre pour l'accès au lexique au niveau oral, modèle basé en partie sur TRACE de McClelland et Elman (1986). Selon Grosjean, quand un bilingue se trouve dans le mode bilingue, les deux langues sont activées, mais la langue de base l'est davantage. L'activation d'une unité linguistique (phonème, syllabe, mot, etc.) dans une des langues et de l'unité correspondante dans l'autre (si elle existe) dépend de leur degré de similitude. L'activation des unités spécifiques à une langue augmente l'activation globale de cette langue et accélère donc la reconnaissance des mots de celle-ci. Par contre, l'activation d'unités linguistiques similaires dans les deux langues ralentit l'identification des unités de l'autre langue et, par conséquent, les mots qui en font partie. Enfin, la fréquence des homophones inter-langues (avec leurs seuils d'activation différents) et la configuration phonétique des mots (prononcés comme code-switchs ou comme emprunts) influencent le processus de reconnaissance des mots de l'autre langue en accélérant ou en freinant leur accès.

Grosjean a proposé que dans le mode bilingue le niveau d'activation de l'autre langue est augmenté ou abaissé selon le nombre de code-switchs ou d'emprunts produits pendant l'interaction. Cette proposition demeure mal définie cependant. La reconnaissance d'un code-switch se fait-elle

plus rapidement si celui-ci est précédé par d'autres code-switchs, ou l'effet de la langue de base est-il tellement fort que le niveau d'activation de la langue alternative reste toujours très bas ? Le modèle ne précise pas non plus le rôle de la contrainte sémantique dans la reconnaissance des code-switchs. Celle-ci influence-t-elle leur accès ou non ? Ce sont ces deux questions qui ont été à la base de notre recherche.

Dans cette étude, nous avons manipulé deux variables contextuelles : la force de la contrainte sémantique et la densité des code-switchs. Nous avons déterminé l'apport de chacune d'elles lors de la reconnaissance de code-switchs et avons tenté de découvrir s'il existe une éventuelle interaction entre les deux. Pour ce faire, nous avons utilisé la tâche du gating (Grosjean, 1980) pour présenter des code-switchs dans divers contextes. Les sujets ont entendu les stimuli dans des segments successivement plus longs (50 msec d'accroissement par présentation) et, après chaque présentation, nous leur avons demandé d'écrire le mot qu'ils entendaient et d'indiquer le degré de confiance qu'ils avaient dans leur réponse. Cette tâche nous a permis d'étudier trois variables dépendantes : le point d'isolement des mots (à savoir, leur point d'identification), le degré de confiance à la fin des mots, et les candidats proposés avant le point d'isolement. Ces derniers nous ont donné un aperçu du processus de reconnaissance et de l'apport du contexte dans l'activation des candidats de l'une, de l'autre ou des deux langues. Les résultats ont ensuite été intégrés dans le modèle d'activation interactive de Grosjean.

2. Méthode

Sujets : Quarante-huit bilingues (suisse-allemand - français), sans problèmes d'audition, ont pris part à l'expérience qui a duré deux heures. Tous les sujets étaient des étudiants de français de l'université de Bâle. Pour obtenir une homogénéité maximale du groupe, nous nous sommes limités à des étudiants qui avaient déjà fait au moins quatre semestres d'étude. Les sujets utilisaient les deux langues régulièrement dans leur vie de tous les jours, parlant le suisse-allemand en famille et dans la communauté bâloise et parlant le français pendant les cours et séminaires de l'université ainsi qu'avec des connaissances francophones.

Matériel : Nous avons choisi 32 substantifs (à l'exclusion des substantifs composés) qui faisaient partie de la vie quotidienne des sujets et qui pouvaient être utilisés comme éléments de la langue de base ou comme éléments de l'autre langue (c'est-à-dire, comme code-switchs). Tous les mots ont été placés dans la deuxième partie d'une phrase à deux propositions lue en suisse-allemand. Ces stimuli ne se trouvaient jamais à la fin de la deuxième proposition, mais étaient toujours suivis de un à quatre mots. La première proposition contenait un substantif qui pouvait être énoncé soit dans la langue de base soit

dans l'autre langue (code-switch). Cette structure bipartite nous a permis de construire quatre types de phrases expérimentales où la deuxième proposition était identique :

Type A) la première proposition (en italique ci-dessous) ne contenait pas de code-switch (en majuscules) et n'exerçait pas de contrainte sur le mot expérimental (en gras) de la deuxième proposition. Exemple : *Me lehrt in dr Sek, dass DELEMONT e schöni Stadt isch* (traduction : *On apprend au collège que DELEMONT est une belle ville*).

Type B) la première proposition ne contenait pas de code-switch mais était contraignante pour le stimulus. Exemple : *Es säge alli Jurassier, dass DELEMONT e schöni Stadt isch* (traduction : *Tous les jurassiens disent que DELEMONT est une belle ville*).

Type C) la première proposition contenait un code-switch mais n'exerçait pas de contrainte sur le mot expérimental. Exemple : *Me lehrt im COLLEGE, dass DELEMONT e schöni Stadt isch* (pour la traduction voir la phrase du type A).

Type D) la première proposition contenait un code-switch et était contraignante pour le stimulus. Exemple : *Es säge alli JURASSIENS, dass DELEMONT e schöni Stadt isch* (pour la traduction voir la phrase du type B).

En nous basant sur ces différents types de construction, nous avons construit 120 phrases : 60 phrases expérimentales qui contenaient un code-switch à reconnaître et 60 phrases de remplissage qui contenaient un substantif en suisse-allemand, ceci afin d'éviter que les sujets s'attendent systématiquement à entendre un code-switch dans la deuxième proposition. A ces 120 phrases, nous avons ajouté 8 phrases d'entraînement. Nous avons bloqué le type de contexte, c'est-à-dire que le groupe A n'entendait que des phrases du type A), le groupe B celles du type B), le groupe C celles du type C) et finalement le groupe D celles du type D). Chaque groupe a donc entendu 15 phrases expérimentales, 15 phrases de remplissage et 2 phrases d'entraînement.

Pour nous assurer que les premières propositions des phrases du type B) et D) étaient contraignantes et que celles des phrases du type A) et C) ne l'étaient pas, nous avons demandé à 13 personnes bilingues d'évaluer à l'aide d'une échelle de 1 (peu contraignant) à 7 (contraignant) la contrainte des propositions. Un test-t apparié a confirmé que les propositions contraignantes avaient un niveau de contrainte significativement plus élevé que les propositions non contraignantes ($t = -8.803$, $p < 0.001$).

Les 128 phrases ont été lues par une locutrice bilingue sans accent notable à un débit normal et avec une prosodie régulière. Elles ont ensuite été digitalisées sur Macintosh à l'aide du logiciel SoundDesigner II. Afin que les deuxièmes propositions soient acoustiquement identiques pour les quatre types de phrases, nous avons sélectionné un exemplaire de la deuxième proposition parmi les quatre enregistrées et l'avons apposée aux premières propositions. Nous avons ensuite préparé les segments expérimentaux de la façon suivante : lors de la première présentation, la phrase était entendue jusqu'à la fin de "dass" (voir exemples ci-dessus) ; puis, à chaque nouvelle présentation, nous avons ajouté 50 msec de plus. Donc, à la première présentation d'un mot expérimental, le sujet entendait 0 msec de celui-ci, à la deuxième, il entendait 50 msec du mot, à la troisième, 100 msec, etc. Lorsque la durée d'un mot expérimental ne correspondait pas à un multiple de 50, le dernier segment était incrémenté de la durée restante. Il faut préciser que, pour les mots suisse-allemands contenus dans les phrases de remplissage, nous avons ajouté à chaque incrémentation 60 msec et non 50 msec (une différence de 10 msec n'étant pas perceptible pour les sujets) puisque ces mots étaient plus longs que les mots expérimentaux.

Lorsqu'un mot stimulus avait été entièrement préparé, la suite (de un à quatre mots) était segmenté mot par mot. Par exemple, en reprenant les phrases présentées ci-dessus,

les sujets ont entendu les segments suivants : *dass Delémont e*; *dass Delémont e schöni*; *dass Delémont e schöni Stadt* et *dass Delémont e schöni Stadt isch*.

Après avoir préparé les séries de phrases, nous les avons enregistrées sur cassette en mode mono. Un signal acoustique annonçait le début d'une série et était suivi d'un silence de trois secondes. Chaque série était suivie d'une pause de huit secondes.

Procédure : Nous avons réparti les 48 sujets en quatre groupes de 12, selon les quatre types de phrases. Chaque groupe a été subdivisé en deux sous-groupes. Le premier a d'abord entendu 16 phrases et, après une pause d'environ cinq minutes, 16 autres. Le deuxième sous-groupe a entendu les phrases dans l'ordre inverse, c'est-à-dire les 16 dernières phrases en premier et les 16 autres en dernier. La première phrase de chaque partie était une phrase d'entraînement.

Nous avons conduit les séances de passation en suisse-allemand, la langue habituelle de conversation entre l'expérimentateur (lui-même bilingue) et les sujets. Les instructions étaient en allemand, la langue écrite de la Suisse alémanique. Les sujets savaient qu'ils allaient devoir reconnaître des substantifs suisse-allemands et des substantifs français dans la deuxième proposition d'une phrase suisse-allemande. Nous leur avons demandé d'écouter ces phrases et, après chaque présentation, de faire trois opérations : 1) écrire le mot qu'ils pensaient avoir entendu ; 2) indiquer un degré de confiance pour le mot proposé en entourant un chiffre sur une échelle de 1 (peu certain) à 10 (certain) ; et 3) indiquer si le substantif appartenait à la langue suisse-allemande ou à la langue française en entourant la lettre "D" pour le suisse-allemand et la lettre "F" pour le français à droite de l'échelle de confiance. Les feuilles de réponse étaient organisées de manière à ce que les sujets écrivent d'abord un mot, puis donnent un degré de confiance et finalement indiquent la langue d'origine du mot. Les sujets avaient 8 secondes pour accomplir ces trois tâches. Nous avons demandé à nos sujets de donner une réponse après chaque présentation, même s'ils étaient très peu sûrs de leur réponse.

Analyse des données : Nous avons analysé trois variables dépendantes : le point d'isolement du mot stimulus, le degré de confiance à la fin du stimulus et les candidats proposés avant le point d'isolement.

Le point d'isolement, à savoir le point où le sujet propose le stimulus et ne change pas d'avis par la suite, nous a permis de déterminer à quel moment le mot a été identifié correctement. Nous avons exprimé ce point en pourcentage de la durée totale du mot. L'orthographe et l'indication "D" ou "F" des sujets nous ont aidé à définir de manière exacte le point d'isolement. Nous avons également étudié le degré de confiance à la fin du stimulus. Finalement, nous avons analysé les candidats erronés avant le point d'isolement en fonction de la langue à laquelle ils appartenaient.

Lorsqu'une mesure manquait (point d'isolement, degré de confiance), ce qui arrivait lorsque le sujet n'identifiait pas du tout le stimulus ou seulement après la fin de celui-ci, ou encore s'il refusait de donner un degré de confiance, la mesure était remplacée par la moyenne de tous les sujets pour le stimulus concerné. Cela a représenté 5.97% des points d'isolement et 6.67% des degrés de confiance.

3. Résultats et discussion

Le point d'isolement

La Figure 1 (page suivante) présente le point d'isolement, exprimé en pourcentage, en fonction de la contrainte sémantique (faible et forte) et de l'occurrence d'un code-switch précédent (sans et avec).

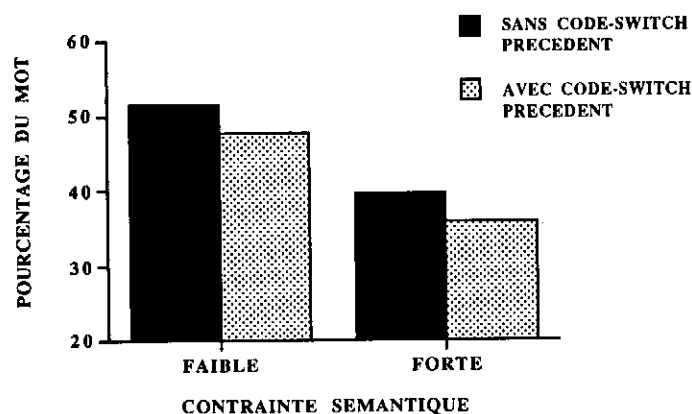


Figure 1. Points d'isolement du mot (%) en fonction de la contrainte sémantique (faible et forte) et de l'occurrence d'un code-switch précédent (sans et avec).

Comme nous pouvons l'observer, il existe un effet de contrainte (hauteur différente des deux groupes de barres) et un effet de densité (hauteur différente des deux barres dans chaque groupe). En effet, dans la condition de contrainte sémantique faible (les phrases des types A et C), les sujets reconnaissent un code-switch après 51.3% lorsqu'il n'y a pas de code-switch précédent et après 47.7% s'il est précédé par un autre code-switch. Dans la condition de contrainte sémantique forte (les phrases des types B et D), le point d'isolement pour les stimuli sans code-switch précédent se trouve à 39.5% et à 35.7% lorsqu'ils sont précédés d'un autre code-switch. Nous avons fait deux analyses de variance et avons trouvé un effet significatif de la contrainte sémantique lors de l'analyse par mots : $F(1,14) = 12.22$, $p < 0.01$, ainsi que lors de l'analyse par sujets : $F(1,44) = 55.04$, $p < 0.001$. Ces analyses ont également montré qu'il existait un effet de densité des code-switchs entre les phrases contenant un code-switch précédent et les phrases n'en contenant pas, aussi bien dans l'analyse par mots, $F(1,14) = 5.09$, $p < 0.05$, que dans l'analyse par sujets, $F(1,44) = 5.27$, $p < 0.05$. Aucune interaction n'a été trouvée entre les deux variables.

De ces analyses, nous pouvons retenir les points suivants. Premièrement, l'effet de la contrainte sémantique est présent et influence le traitement des code-switchs dans le mode bilingue en accélérant leur reconnaissance. Ceci est valable pour les phrases sans code-switch précédent ainsi que pour celles avec code-switch précédent. Ce résultat confirme l'effet "contexte" que l'on trouve à maintes reprises dans la littérature mais ajoute un fait nouveau : le contexte influence les mots des deux langues chez l'auditeur bilingue, ceux de la langue de base et ceux de la langue alternative (au moins en mode de communication bilingue). Deuxièmement, l'auditeur identifie plus rapidement un code-switch dans une phrase si celui-ci est précédé d'un autre code-switch. La présence d'éléments de l'autre langue active donc davantage le lexique de cette langue et l'accès se fait plus rapidement puisque les mots qu'elle contient ont besoin de moins d'excitation pour atteindre leur seuil d'activation. Troisièmement, il n'existe pas d'interaction entre l'effet contrainte et l'effet densité. Cela signifie que les deux effets accélèrent la reconnaissance d'un code-switch mais qu'ils ne s'influencent pas mutuellement.

A partir de ces résultats, nous pouvons émettre l'hypothèse que les deux variables influencent probablement de manière différente le processus de reconnaissance. Le contexte interviendrait depuis un niveau supérieur et serait donc une information descendante. Par contre, la densité des code-switchs aurait un effet ascendant, car ce sont les informations phonétiques des code-switchs précédents qui augmentent le niveau d'activation de la langue d'où ils proviennent. Lorsqu'un code-switch apparaît, la langue alternative est encore plus activée et la reconnaissance du mot se fait donc plus rapidement.

Le degré de confiance

La Figure 2 (page suivante) montre le degré de confiance à la fin du mot en fonction de la contrainte sémantique (faible et forte) et de l'occurrence d'un code-switch précédent (sans et avec).

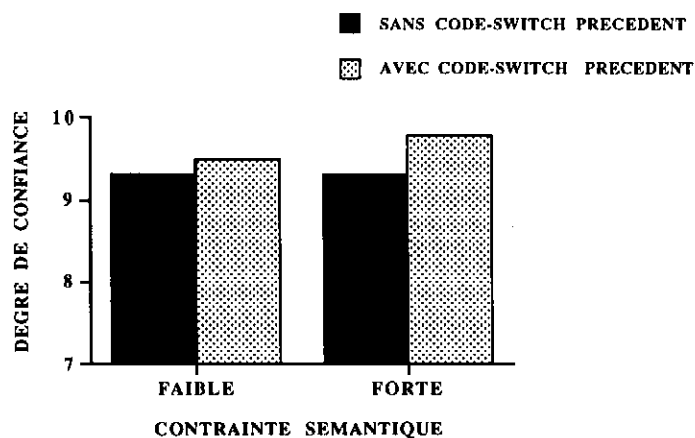


Figure 2. Degrés de confiance à la fin du mot en fonction de la contrainte sémantique (faible et forte) et de l'occurrence d'un code-switch précédent (sans et avec).

Nous constatons peu de différence au niveau de la contrainte sémantique mais une différence plus grande au niveau de la densité. En effet, il existe un degré de confiance moyen de 9.3 dans la condition de contrainte faible ainsi que dans la condition de contrainte forte lorsque le stimulus n'est pas précédé d'un autre code-switch (phrases des types A et B). Par contre, si le stimulus est précédé d'un autre code-switch, la moyenne globale pour la condition de contrainte faible est de 9.5 et pour la contrainte forte de 9.9 (phrases des types C et D). Deux analyses de variance confirment qu'il existe une différence significative entre les stimuli précédés d'un autre code-switch et les stimuli qui ne le sont pas. Nous avons trouvé cet effet dans l'analyse par mots ($F(1,14) = 9.51, p < 0.01$) et dans l'analyse par sujets ($F(1,44) = 5.93, p < 0.05$). Par contre, l'effet de contrainte n'est pas significatif et il n'existe pas d'interaction.

Ces résultats confirment en partie ceux obtenus pour le point d'isolement. Un code-switch est reconnu avec plus de confiance s'il est précédé d'un autre code-switch que ce soit dans une situation de contrainte faible ou de contrainte forte. Par contre, l'absence d'un effet de contrainte, ici, est probablement due au fait que les valeurs des degrés de confiance des quatre types de contexte se rapprochent les unes des

autres vers la fin des stimuli. Il semble que les informations contextuelles dans les conditions de contrainte forte et faible soient suffisantes pour neutraliser les valeurs vers la fin des stimuli, ce qui produit un effet plafond des courbes lors de la présentation des derniers segments. L'absence d'un effet plafond pour la variable densité s'explique par le fait qu'un premier code-switch indique à l'auditeur bilingue la présence possible d'autres éléments de la langue alternative ce qui implique, lors de la présence effective d'un deuxième code-switch, que la reconnaissance de celui-ci se fait avec plus de confiance.

Le processus d'isolement des code-switchs

La Figure 3 ci-dessous présente le nombre moyen de candidats français en fonction de la durée des segments et du type de phrase contextuelle (A, B, C, D). Les résultats obtenus pour les sept premiers segments ont été retenus.

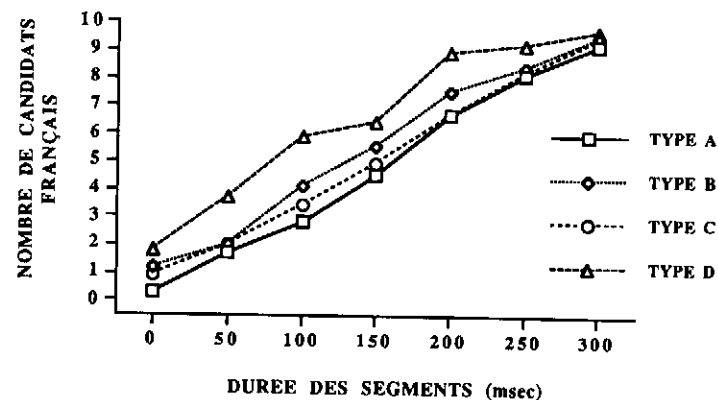


Figure 3. Nombre moyen de candidats français en fonction de la durée des segments et du type de phrase contextuelle (A, B, C, D).

Nous remarquons tout d'abord un effet tout à fait clair de la langue de base. En début de mot, les sujets sont influencés par la langue "porteuse" (le suisse-allemand) et ne proposent donc que peu de candidats français. Ce n'est que lorsque l'information acoustico-phonétique du code-switch arrive (dans les segments successifs) que les sujets

commencent à proposer des candidats de l'autre langue. Nous constatons également que le nombre de candidats français est plus important quand les stimuli sont précédés d'un code-switch ou quand la contrainte sémantique est forte. En effet, nous trouvons 5.8 candidats français en moyenne pour les phrases qui contiennent un code-switch précédent (types C et D) contre 5.1 candidats pour celles qui n'en ont pas (types A et B). De plus, nous obtenons plus de candidats français lorsque la contrainte sémantique est forte (types B et D; moyenne de 5.9) que lorsqu'elle est faible (types A et C; moyenne de 4.9). Une analyse de variance confirme que ces effets sont significatifs (densité : $F(1,98) = 29.68, p < 0.001$; contrainte sémantique : $F(1,98) = 21.68, p < 0.001$).

4. Discussion générale

Les résultats de cette étude ont montré l'importance des deux variables testées. La première, la densité des code-switchs, a été mise en évidence par le point d'isolement des mots (les stimuli précédés d'un autre code-switch sont identifiés plus rapidement que les autres), par le degré de confiance (les stimuli sont reconnus avec plus de confiance) et par les candidats proposés lors des premiers segments (les phrases qui sont précédées d'un code-switch obtiennent plus de candidats français que celles qui ne le sont pas). Pour expliquer ces résultats, nous avons émis l'hypothèse qu'un code-switch précédent active davantage de candidats de l'autre langue, ce qui fait que le mot stimulus, s'il s'agit d'un code-switch, sera reconnu plus rapidement.

Quant à la deuxième variable, la contrainte sémantique, nous avons trouvé un effet significatif avec le point d'isolement (qui est précoce en contexte contraignant) et les candidats proposés lors des premiers segments (qui appartiennent à l'autre langue lorsque le contexte active celle-ci). Nous en avons conclu que la contrainte sémantique fonctionne de manière descendante et active les mots de l'autre langue lorsque le contexte fait appel à elle.

A partir de ces résultats, certaines modifications peuvent être apportées au modèle de Grosjean (1988). Celles-ci sont en italique afin de les différencier des propositions d'origine.

1. En mode de communication bilingue, les deux langues sont activées mais la langue de base l'est plus fortement. Le niveau d'activation de la

langue alternative peut être modifié tout au cours de l'interaction. *Deux variables peuvent augmenter ce niveau : la présence de code-switchs précédents et la contrainte sémantique et pragmatique. La première variable (la densité des code-switchs) fonctionne comme information ascendante et augmente l'activation de la langue alternative, facilitant ainsi la reconnaissance des code-switchs. La contrainte sémantique, quant à elle, est une information descendante qui active plus fortement les mots de la langue alternative lorsque celle-ci est appropriée à la situation (interlocuteurs, sujets, etc).*

2. L'activation d'une unité dans une langue (phonème, syllabe, mot, etc.) et de l'unité correspondante dans l'autre (si elle existe) dépend de leur degré de similitude.
3. L'activation des unités spécifiques à une langue (traits, phonèmes, syllabes) augmente l'activation globale de cette langue et accélère donc la reconnaissance des mots de celle-ci.
4. L'activation d'un mot spécifique à une langue augmente également l'activation globale de celle-ci et accélère donc la reconnaissance des mots qui lui appartiennent.
5. L'activation de mots qui se ressemblent dans les deux langues (homophones inter-langues) ralentit la reconnaissance des code-switchs et des emprunts. La fréquence des homophones (avec leurs taux d'activation différents) et la configuration phonétique du mot influencent le processus de reconnaissance en l'accéléralant ou en le freinant.
6. *Lorsque le contexte sémantique et pragmatique exerce une contrainte sur la langue alternative, celle-ci devient plus active et les mots qui en font partie sont reconnus plus rapidement.*

La présente étude a réussi à affiner le modèle de Grosjean (1988) en lui permettant de rendre compte de deux nouveaux effets : la densité des code-switchs et la contrainte sémantique. Une autre étude sera nécessaire pour confirmer les résultats présentés ici et pour montrer que ce n'est pas une contrainte sémantique quelconque qui active les mots de la langue alternative mais seulement celle qui correspond aux domaines d'utilisation de cette langue.

5. Bibliographie

GROSJEAN, F. (1980): "Spoken word recognition and the gating paradigm", *Perception and Psychophysics*, 28, 267-283.

GROSJEAN, F. (1988): "Exploring the recognition of guest words in bilingual speech", *Language and Cognitive Processes*, 3 (3), 233-274.

McCLELLAND, J. & J. ELMAN, (1986): "The TRACE model of speech perception", *Cognitive Psychology*, 18, 1-86.

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 ; Padeloup, 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

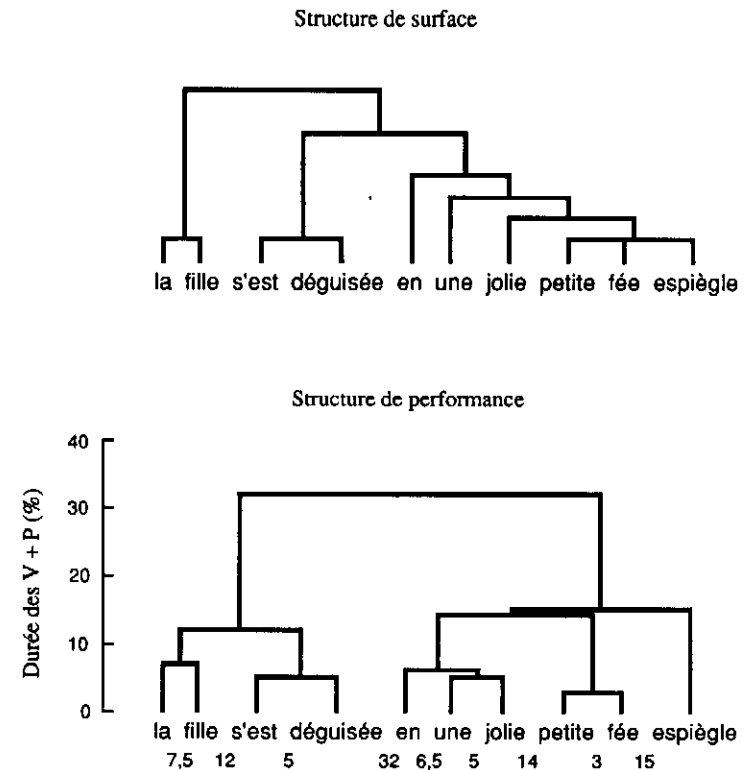


Fig. 1. - Structure de surface et structure de performance de la phrase 1. La structure de performance est élaborée à partir des durées des V + P converties en pourcentages.

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

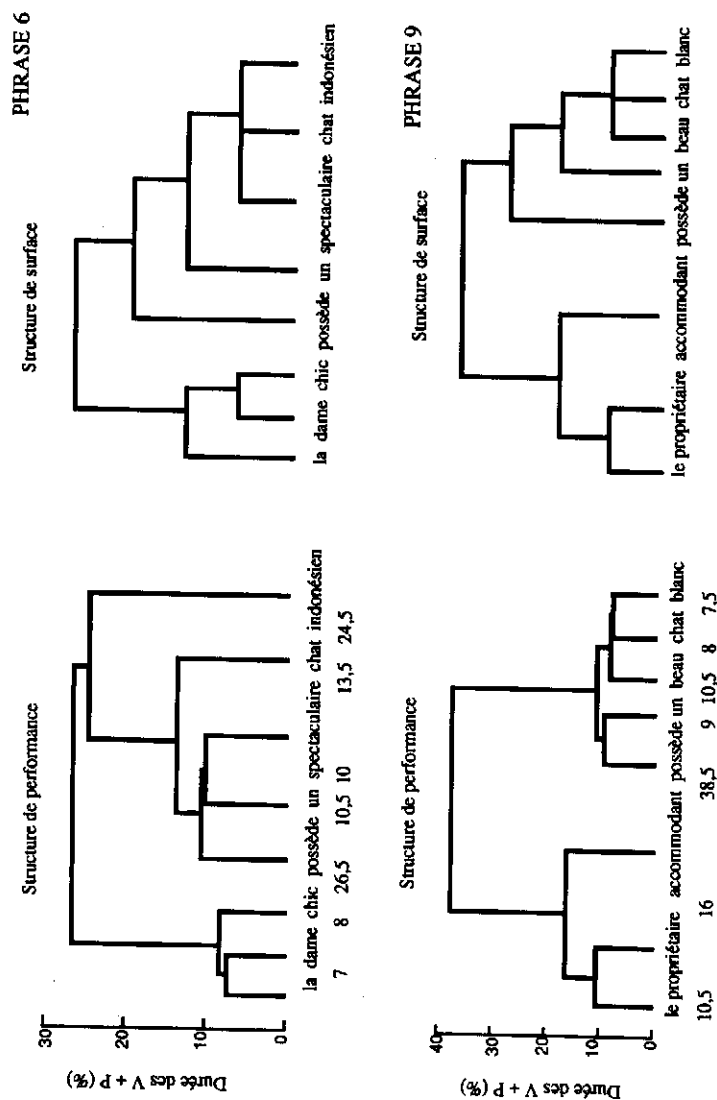


Fig. 2. - Structure de surface et structure de performance des phrases 6 et 9. Les structures de performance sont élaborées à partir des durées des V + P converties en pourcentages.

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.