

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Faculté de psychologie et des sciences de l'éducation
Unité de psychologie cognitive



LA RÉGULATION DES FLUX D'INFORMATION LORS DE L'ENCODAGE GRAMMATICAL DE LA PHRASE. L'ACCORD COMME OUTIL D'INVESTIGATION

Julie Franck

*Thèse présentée en vue de l'obtention du titre de
Docteur en sciences psychologiques*

*Promoteur : Professeur M. Hupet
Jury : Professeurs G. Altmann, J. Costermans,
P. Feyereisen et G. Vigliocco*

Mai 2000

Ce travail n'aurait pas vu le jour sans le soutien du professeur Michel Hupet qui a dirigé ma thèse avec attention tout en me laissant la plus grande autonomie. C'est avant tout lui qui m'a convaincue de revenir travailler à Louvain-la-Neuve, alors que j'entreprenais mes recherches sur l'accord en français à University College London. Il est, avec Brian Butterworth, parmi ceux qui m'ont donné confiance en moi pour me lancer dans cette entreprise intellectuelle.

Je remercie également mon comité d'encadrement, les professeurs Gabriella Vigliocco et Pierre Feyereisen, ainsi que les membres du jury, les professeurs Gerry Altmann et Jean Costermans, qui ont accepté de prendre le temps de lire ce travail et de le commenter précieusement. Je m'adresse tout particulièrement à Gabriella Vigliocco avec qui j'ai mené plusieurs recherches. Notre collaboration a été, et continue d'être mon principal point d'ancrage dans le monde de la psycholinguistique. Les séjours de travail passés chez elle, les centaines d'échanges par courrier électronique, la confiance mutuelle que nous nous témoignons et la proximité de nos points de vue ont fait d'elle bien plus qu'un superviseur ou une collaboratrice.

Merci aussi aux amis et collègues de l'unité EXCO et de la Faculté pour leurs conseils et l'ambiance détendue grâce à laquelle le travail a toujours été un plaisir.

Papa, maman, Jean-Sé, John, Jean, Gérard et ceux que je ne peux citer faute de place mais qui, je l'espère, se reconnaissent dans ce paragraphe ; merci d'avoir éveillé et/ou entretenu ma curiosité intellectuelle et ma sensibilité politique sans lesquels je n'aurais jamais trouvé le fil conducteur d'un tel travail théorique de recherche.

Tout ce temps que ma thèse a absorbé ces derniers mois, j'aurais voulu le passer avec toi, Elise. Mais tu comprendras un jour qu'une maman, pour être vraiment heureuse, a besoin de travailler.

Jean-Sé, ce travail est pour toi. Merci pour tout ce que nous vivons ensemble depuis douze ans.

Table des matières

Table des matières	i
Prélude	1
Chapitre 1 Introduction générale	11
1.1 Le langage : un système modulaire ou interactif ?	11
1.1.1. Le concept de modularité	12
1.1.2. La perspective modulaire de la production	16
1.1.3. La perspective non modulaire de la production	25
1.2 L'accord grammatical : input minimal ou input maximal ?	34
1.2.1. Les recherches en faveur de l'hypothèse d'un input minimal	35
1.2.2. Les recherches en faveur de l'hypothèse d'un input maximal	41
Chapitre 2 When sex and syntax go hand in hand: Gender agreement in language production	49
2.1 Introduction	49
2.2 Agreement in Models of Sentence Production	51
2.3 Gender Systems in Italian and French	55
2.4 The Plan of the Study	57
2.5 Experiment 1: Italian	58
2.5.1. Method	58
2.5.2. Results	61
2.5.3. Discussion	62
2.6 Experiment 2: French	62
2.6.1. Method	63
2.6.2. Results	65
2.6.3. Discussion	66
2.7 Experiment 3: Italian	68
2.7.1. Method	68
2.7.2. Results	70
2.7.3. Discussion	71
2.8 Experiment 4: French	71
2.8.1. Method	71
2.8.2. Results	73
2.8.3. Discussion	73
2.9 General Discussion	74
2.9.1. Comparisons Between Number and Gender Agreement	74
2.9.2. Relevance of the Present Studies for Theories of Sentence Production	77
2.10 Conclusion	80

Chapitre 3 When Sex hits Syntax: Syntactic accuracy in language production	81
3.1 Introduction	81
3.2 Experiments 1(Italian) and 2 (French)	86
3.2.1. Method	86
3.2.2. Results and discussion	89
3.3 Experiments 3 (Italian) and 4 (French)	92
3.3.1. Method	92
3.3.2. Results and discussion	93
3.4 General discussion	94
3.5 Conclusion	98
Chapitre 4 Intermezzo : l'hypothèse lexicale	99
4.1 Introduction	99
4.2 Interprétation de l'effet conceptuel	101
4.3 Expérience	107
4.3.1. Méthode	107
4.3.2. Résultats	109
4.4 Discussion	110
4.4.1. Remise en question de l'hypothèse notionnelle	110
4.4.2. Interprétation des différences entre le français et l'anglais	117
4.5 Conclusion	118
Chapitre 5 Quand morphophonologie et syntaxe collaborent	121
5.1 Introduction	121
5.2 L'interface entre phonologie et syntaxe : les observations	122
5.3 La morphophonologie du genre en français	125
5.4 Expérience	126
5.4.1. Méthode	127
5.4.2. Résultats	129
5.4.3. Discussion	130
5.5 Aperçu de la recherche en italien	131
5.5.1. La morphophonologie du genre en italien	131
5.5.2. Expérience	131
5.6 Discussion générale	132
5.6.1. Interprétation théorique de l'effet morphophonologique	133
5.6.2. Hypothèses explicatives des différences interlinguistiques	140
1.7 Conclusion	144
Chapitre 6 Attraction in sentence production: The role of syntactic structure and processing complexity	147
6.1 Introduction	147
6.2 Interpreting the attraction effect	149
6.3 A preview of the experiments	154
6.4 Experiment 1 (French)	155

6.4.1. Method	157
6.4.2. Results	159
6.4.3. Discussion	160
6.5 Experiment 2 (English)	161
6.5.1. Method	161
6.5.2. Results	161
6.5.3. Discussion	163
6.6 Experiment 3 (French)	166
6.6.1. Method	167
6.6.2. Results	168
6.7 Experiment 4 (English)	170
6.7.1. Method	170
6.7.2. Results	170
6.7.3. Discussion of Experiments 3 and 4	172
6.8 General discussion	173
6.8.1. The role of syntactic structure	173
6.8.2. The role of processing complexity	175
6.8.3. Cross-linguistic considerations	177
6.8.4. Alternative explanation: the role of conceptual salience	180
6.9 Conclusion	181
Chapitre 7 L'hypothèse de la saillance conceptuelle	183
7.1 Introduction	183
7.2 Expérience	189
7.2.1. Méthode	189
7.2.2. Résultats	190
7.3 Discussion	191
7.4 Conclusion	193
Chapitre 8 Synthèse théorique	195
8.1 Les contraintes impliquées dans la réalisation de l'accord	196
8.1.1. Les contraintes relatives au sens	196
8.1.2. Les contraintes relatives à la morphophonologie	204
8.1.3. Les contraintes relatives à la syntaxe	206
8.1.4. Les contraintes extra-linguistiques	213
8.2 La variabilité dans l'application des contraintes	214
8.2.1. La variabilité au sein des langues	214
8.2.2. La variabilité entre les langues	221
Conclusion générale	225
Références bibliographiques	229
Annexes	241

Prélude

Le traitement syntaxique sous-jacent à l'encodage et au décodage de la phrase mobilise-t-il des informations non syntaxiques, à savoir des informations pragmatiques, conceptuelles, lexicales et phonologiques ? Cette question est fondamentale pour le chercheur qui veut expliciter les processus engagés dans la production et la compréhension d'énoncés. Elle a suscité de très vifs débats. Expliquer les flux d'information entre les différents niveaux de traitement permettrait non seulement de dégager une propriété du système langagier, mais aussi de définir les avantages et inconvénients potentiels de cette propriété.

C'est en matière de compréhension de phrases que ces débats ont été les plus développés. Ils opposent d'une part les partisans de modèles autonomes pour qui le traitement syntaxique est envisagé comme une étape obligatoire, préalable à toute interprétation et complètement autonome (ex. Frazier, 1990), d'autre part les partisans de modèles interactifs selon lesquels le traitement syntaxique obligatoire s'effectue parallèlement à d'autres traitements, sur la base d'échanges mutuels d'informations (ex. Altmann, 1989 ; Trueswell, Tanenhaus & Garnsey, 1994), et enfin les partisans de modèles intégratifs pour lesquels il n'y a pas lieu de faire intervenir une représentation syntaxique intermédiaire, la représentation conceptuelle étant construite d'emblée (ex. Marslen-Wilson & Tyler, 1987).

La question de l'autonomie de la syntaxe a moins envahi la littérature concernant la production de la phrase, probablement en raison de la difficulté d'investiguer expérimentalement les processus en jeu. Pourtant, contrairement aux mécanismes de compréhension, qui peuvent souvent aboutir à un décodage satisfaisant sans réelle analyse grammaticale, la mise en application de patterns grammaticaux est indispensable à la création d'un énoncé correct. Autrement dit, s'il y a lieu de mettre à jour une composante proprement syntaxique, c'est justement sur les processus de production qu'il s'agit de se pencher. L'étude des mécanismes de production semble donc offrir un terrain particulièrement propice à l'étude de la composante syntaxique, et plus particulièrement à l'étude des relations entre la syntaxe et les autres aspects de la construction de la phrase.

Sur le plan théorique, l'hypothèse autonomiste ou *modulariste* postule l'existence d'un 'module' syntaxique contenant des procédures responsables d'élaborer des patterns spécifiquement syntaxiques, indépendamment du contenu des éléments de la phrase. On trouve déjà une ébauche de cette perspective, selon laquelle les formes syntaxiques sont indépendantes de leur contenu, dans les travaux d'Edward Sapir qui l'expose de la façon suivante :

All languages evince a curious instinct for the development of one or more particular grammatical processes at the expense of others, tending always to lose sight of any explicit functional value that the process may have had in the first instance. (...) This feeling for form as such, freely expanding along predetermined lines and greatly inhibited in certain directions by the lack of controlling types of patterning, should be more understood than it seems to be. (Sapir, 1921, pp. 60-61).

Tandis que les réflexions de Sapir concernent le développement de patterns formels au cours de l'évolution de la langue, la psycholinguistique aujourd'hui s'interroge quant à elle sur l'utilisation de ces patterns grammaticaux par un individu à un moment précis de l'évolution diachronique de la langue. Les premiers travaux de psycholinguistique visant à mettre en évidence l'existence de procédures syntaxiques autonomes remontent aux années soixante. Menés au départ par George Miller (ex. 1962) sous l'impulsion de la théorie de la grammaire générative et transformationnelle du linguiste Noam Chomsky (1957), ils visaient essentiellement à démontrer la détermination des processus de production par la composante syntaxique de la langue. Cette conception extrême de la production a toutefois rapidement donné lieu à un courant opposé mettant l'accent sur le rôle du sens et de la perspective communicative dans l'élaboration de la forme des énoncés, reléguant la syntaxe à un statut secondaire, voire même inexistant (ex. Osgood, 1971, cité dans Bock & Kroch, 1989).

On peut considérer que ces deux positions théoriques constituent les pôles extrêmes du fonctionnalisme linguistique (Bates & MacWhinney, 1982) : selon la première position, les primitives syntaxiques sont à elles seules suffisantes pour rendre compte de la création d'un énoncé, selon la seconde, le système syntaxique n'est ni suffisant ni même nécessaire pour expliquer la production de la phrase. Au centre du débat sur le fonctionnalisme linguistique se situe la question des relations entre forme et fonction, les uns mettant l'accent sur le rôle de la forme et son indépendance par rapport à la fonction qu'elle remplit, les autres articulant leur théorie sur le rôle joué par la fonction, la forme n'étant considérée que comme un moyen de remplir la fonction et donc comme peu intéressante pour comprendre l'utilisation du langage. Dans notre thèse, une position intermédiaire est adoptée selon laquelle formes et fonctions sont liées par des contraintes de fonctionnement pesant à la fois sur le système cognitif général (ex. limitations attentionnelles et liées à la mémoire de travail) et sur le système spécifique de production du langage. Ce sont ces dernières qui ont attiré notre attention, et plus particulièrement les contraintes relatives à la limitation des flux d'informations entre le niveau syntaxique et les autres niveaux de traitement de la phrase.

Tout le monde s'accorde aujourd'hui pour reconnaître que la construction de la structure syntaxique de la phrase est sensible à des facteurs tels que le contexte dans lequel elle apparaît, ou encore à différentes caractéristiques du message qu'elle porte. De nombreuses études expérimentales ont mis en évidence que les

phrases commencent souvent par des constituants référant à des entités animées, concrètes, prototypiques de leur catégorie sémantique, ou encore par de l'information connue plutôt que nouvelle, etc. Or, ces propriétés ne sont pas de nature syntaxique mais relèvent de la connaissance qu'a le locuteur de ce dont il parle, ou du partage des connaissances entre le locuteur et son interlocuteur (ex. ce que le locuteur pense que son interlocuteur sait ou ne sait pas, ce qui a déjà été dit dans la phrase précédente, etc.).

Ces résultats ont inévitablement impliqué une remise en question de la position défendue par les tenants des théories 'basées sur la syntaxe'. L'enjeu du débat aujourd'hui n'est donc plus de trancher entre le parti de la modularité stricte et celui de l'interdépendance totale de la syntaxe et des autres composantes de la production, mais bien de définir les *limites* de l'autonomie syntaxique. La pertinence de la question relève dès lors d'une analyse fine de la composante syntaxique pour en spécifier les aspects directement déterminés par des éléments conceptuels et ceux relevant de catégories et de représentations intrinsèquement linguistiques. Il va par exemple de soi aujourd'hui que la construction syntaxique de la phrase opère dans un premier temps sous le contrôle partiel d'éléments relatifs au système cognitif global. Par exemple, pour assigner une fonction syntaxique (ex. sujet, objet direct) à un mot, il est nécessaire de se référer au sens du message spécifiant quel est le rôle conceptuel de ce mot (ex. agent, patient). Si l'on veut exprimer l'idée que « l'âne mange la carotte », l'élaboration de la structure syntaxique de la phrase ne peut pas se contenter de se baser sur la sélection des mots *âne*, *mange* et *carotte* correspondant aux concepts ; la structure conceptuelle du message doit encore être disponible pour signaler à la composante syntaxique que *âne* doit recevoir la fonction sujet tandis que *carotte* doit être en position d'objet du verbe. Si cette structure n'est pas disponible au locuteur, il risque tout simplement de produire la phrase « la carotte mange l'âne ».

Un processus comme l'assignation des fonctions syntaxiques doit donc en tout cas opérer sous le contrôle de la structure conceptuelle pour assurer une concordance entre la forme et le sens de l'énoncé. L'hypothèse que certains traitements syntaxiques peuvent fonctionner en termes de catégories et représentations purement linguistiques ne peut par conséquent porter que sur certains de ces traitements syntaxiques, subséquents à l'assignation des fonctions syntaxiques (ex. Bock, 1987 ; Garrett, in press). Les théories modularistes actuelles avancent que ces traitements opéreraient non seulement indépendamment d'aspects cognitifs plus larges, mais qu'ils seraient également insensibles aux caractéristiques phonologiques de l'énoncé. Dans ce cadre théorique, la formulation du langage interagirait avec le système cognitif général jusqu'à un certain point : c'est seulement au-delà de cette étape qu'il apparaît raisonnable de chercher quelque traitement syntaxique autonome, impénétrable, isolé des autres processus cognitifs. La question de la modularité de la syntaxe consiste aujourd'hui d'une part à déterminer où se situe la limite de la perméabilité de la syntaxe aux

informations conceptuelles, d'autre part à déterminer dans quelle mesure la syntaxe est effectivement aveugle aux aspects phonologiques de l'énoncé.

Outre la question de la modularité de la syntaxe, une seconde question est posée par les théoriciens de la production de la phrase. Elle concerne l'organisation des mots lors de la formulation de la structure syntaxique de la phrase. La spécification de cette organisation des segments de la phrase au stade du traitement syntaxique de la production est importante étant donné qu'il est supposé qu'elle joue un rôle déterminant dans la réalisation des opérations syntaxiques ayant lieu à ce stade. La question consiste à déterminer si le locuteur accède progressivement aux mots et à la structure dans laquelle ils s'insèrent au fur et à mesure que les mots deviennent accessibles, ou s'il existe un stade dans la production, précédant la mise en ordre linéaire des mots, au cours duquel les mots sont placés dans une structure hiérarchique établissant les relations syntaxiques qu'ils entretiennent les uns avec les autres. Autrement dit, il s'agit de préciser les liens entre la récupération des mots et l'élaboration de la structure syntaxique dans laquelle ils s'insèrent. Celle-ci est-elle formée au fur et à mesure que les mots sont rendus accessibles au locuteur, la chaîne de mots finalement prononcée n'étant que le reflet de l'ordre dans lequel les mots lui sont venus à l'esprit ? Ou bien doit-on envisager l'existence d'un stade intermédiaire dans la production de la phrase au cours duquel les mots sont organisés selon une structure particulière, non linéaire, reflétant non pas l'ordre de prononciation des mots mais bien les relations syntaxiques entre les différents segments de la phrase ?

L'objectif de notre thèse de doctorat est d'éclairer ces deux débats théoriques sur les processus d'encodage syntaxique, à savoir la question de son autonomie de fonctionnement et celle de l'organisation des segments syntaxiques à ce stade. Il s'agit de mettre en évidence les différentes contraintes, i.e. de nature conceptuelle, lexicale, phonologique et syntaxique, à l'œuvre dans la réalisation d'une opération syntaxique typique : l'accord grammatical. Les deux débats à l'origine de nos travaux peuvent se résumer à la question plus générale de la régulation des flux d'informations, d'une part entre le niveau de traitement syntaxique et les autres niveaux de production, d'autre part entre les éléments à accorder, au sein même de la composante syntaxique.

Les recherches empiriques que nous avons menées se basent sur l'analyse des erreurs d'accord en genre et en nombre, et plus particulièrement sur l'analyse des variables pouvant influencer leur occurrence. Selon nous, quatre raisons font de l'accord un outil de travail précieux pour éclairer le fonctionnement de la composante syntaxique en production. Premièrement, l'accord est une opération syntaxique par laquelle deux éléments reçoivent la même caractéristique de nombre, de genre ou de personne parce qu'ils entretiennent une relation

particulière. Le lien unissant ces unités existe à la fois au niveau du sens du message (par exemple, le sujet et le verbe sont liés par une relation de type <agent/action>), au niveau de la syntaxe utilisée pour exprimer ce message (un verbe sera pluriel si le sujet est pluriel) et au niveau de la forme des mots sélectionnés (le pluriel se retrouve à la fois dans la morphologie du nom sujet et dans celle du verbe). Deuxièmement, l'accord est un exemple typique d'opération syntaxique dont la réalisation repose sur l'application de règles combinatoires assurant une fonction de cohésion de l'énoncé. Troisièmement, ces règles d'accord sont relativement simples, précises et peu soumises aux irrégularités. De plus, l'accord est une opération hautement automatisée, ce qui la rend relativement peu sensible aux variations individuelles, contextuelles (comme l'étude en laboratoire) ou encore socio-culturelles. Enfin, l'accord est requis dans plus de trois-quarts des langues naturelles ; l'étude en parallèle des erreurs d'accord dans différentes langues peut donc s'avérer éclairante quant à la question plus générale de l'universalité des contraintes sous-jacentes aux opérations syntaxiques.

Le phénomène d'accord est donc un outil privilégié pour les psycholinguistes soucieux de déterminer dans quelle mesure certaines variables non syntaxiques (conceptuelles ou phonologiques) peuvent assurer le contrôle d'une opération typiquement syntaxique : la présence simultanée de la relation d'accord aux niveaux conceptuel, syntaxique et phonologique permet d'aborder directement la question du flux d'information entre les niveaux de production de la phrase. Il constitue également un terrain d'investigation particulièrement pertinent pour comprendre l'organisation des mots au stade de la formulation syntaxique puisque cette organisation est déterminante pour la réalisation des opérations syntaxiques effectuées à ce stade. Deux grandes questions concernant l'accord grammatical ont donc guidé nos recherches : (1) Des contraintes de nature conceptuelle et phonologique peuvent-elles contrôler la réalisation de l'accord ? (2) Quelles sont les contraintes de nature syntaxique qui contrôlent l'accord ?

(1) Des contraintes de nature conceptuelle et phonologique peuvent-elles contrôler la réalisation de l'accord ?

Notre travail confronte deux hypothèses théoriques sur la question. Selon la première hypothèse, l'information (de nombre, genre ou personne) que porte la cible de l'accord (ex. le verbe dans l'accord sujet-verbe, l'adjectif prédicatif dans l'accord sujet-prédicat) ne porte pas de signification particulière, il ne s'agit que d'un indice abstrait indiquant l'élément source auquel la cible devra être rapportée et donc accordée. Dans une telle optique, la représentation conceptuelle du message détermine la sélection de l'information syntaxique (nombre, genre) de l'élément source, mais seule cette dernière est utilisée pour effectuer l'accord. Par ailleurs, l'information de type morphophonologique se situant sur la source n'est pas non plus pertinente pour la construction de l'accord. Par exemple, la présence de terminaisons typiques du genre sur le nom (ex. *-eur* au masculin et *-euse* au

féminin) n'exerce aucune influence sur l'accord de l'adjectif avec ce nom. Dans cette perspective théorique, seule l'information syntaxique portée par la source de l'accord assure le contrôle de sa construction. Le système de production du langage est conçu comme une succession de stades où seule la séquence finale d'un stade particulier peut influencer l'opération du stade ultérieur. L'information conceptuelle est située trop en amont de l'encodage grammatical pour pouvoir exercer un contrôle sur ce dernier, tandis que l'information morphophonologique, située en aval, n'est pas encore disponible pour pouvoir l'influencer. Dans ce cadre, l'opération syntaxique d'accord est conçue comme modulaire : elle est contrôlée par un minimum d'information, i.e. l'information syntaxique, et demeure isolée des autres composantes de la production. Cette approche, nous la résumons sous le terme d'*hypothèse minimaliste* : l'encodeur grammatical est supposé travailler sur base d'un input minimal limité à l'information syntaxique présente dans le lexique.

A l'inverse de cette option théorique, la seconde hypothèse postule que les informations conceptuelle et morphophonologique régulent, dans une certaine mesure, la construction de l'accord. L'information conceptuelle de nombre et de genre est présente dans le message que le locuteur désire exprimer ; selon qu'il pense à une ou plusieurs choses, à un être masculin ou féminin. L'information conceptuelle coïncide la plupart du temps avec l'information syntaxique, et cette hypothèse postule que le corrélat conceptuel de l'information syntaxique contribue à une réalisation correcte de l'accord en ce qu'il constitue une information supplémentaire disponible en cas de perte de l'information syntaxique. Les corrélats morphophonologiques sont également supposés aider la construction de l'accord. Par exemple, la présence d'un suffixe pluriel *-s* ou d'un suffixe typique du féminin comme *-euse* sur la source de l'accord vient renforcer l'information syntaxique et ainsi favoriser la réalisation correcte de l'accord. Nous résumons cette conception sous le terme d'*hypothèse maximaliste* : l'encodeur grammatical bénéficie d'un input maximal d'informations pour mener à bien sa tâche de construction de la structure syntaxique.

(2) *Quelles sont les contraintes de nature syntaxique qui contrôlent l'accord ?*

L'opération d'accord est réalisée lors de l'encodage grammatical au moment où les relations syntaxiques entre les mots sont spécifiées. On considère généralement qu'elle se déroule après la sélection des unités lexicales et après leur assignation à une fonction syntaxique particulière (sujet, objet, verbe, etc.). Une telle information est effectivement nécessaire puisque l'accord se fait entre deux éléments jouant un rôle syntaxique spécifique dans la phrase (ex. en français, l'adjectif prédicatif ne s'accorde pas avec le genre de n'importe quel nom le précédant mais bien avec le genre du nom occupant la position sujet). La question se pose de savoir quel type de structure unit les segments syntaxiques lors de la construction de l'accord à ce stade de l'élaboration syntaxique de la phrase. Selon

une première hypothèse, il s'agit d'une *structure linéaire*, similaire à la chaîne de mots finale telle qu'on la trouve lors de l'articulation de l'énoncé. Dans ce cadre, le flux d'information entre la source et la cible de l'accord sera déterminé par l'agencement linéaire des segments. Une autre hypothèse postule au contraire l'existence d'une *organisation de type hiérarchique* des segments au stade de l'encodage grammatical. Si une telle structure est effectivement construite avant la mise en ordre linéaire des mots dans la phrase, elle doit avoir un impact sur le flux d'information entre les éléments à accorder.

Pour éclairer ces deux débats, nous avons largement exploité sur le plan expérimental un phénomène observé en de faibles proportions dans des situations naturelles de langage : la production d'erreurs d'accord en nombre et en genre. En situation d'élocution spontanée, on dénombre en moyenne une erreur tous les mille mots produits (Bock, 1991), le taux de productions incorrectes tombant même à moins d'un demi pour cent des phrases si l'on ne considère que les erreurs liées à la structure syntaxique de la phrase (Deese, 1984). Néanmoins, bon nombre des erreurs d'accord spontanées se caractérisent par le fait qu'elles ont tendance à apparaître lorsqu'un mot de nombre ou de genre opposé à celui de la source de l'accord se situe dans son entourage proche. Appelées erreurs de proximité ou *erreurs d'attraction*, elles consistent à accorder le mot cible non pas avec le mot source qui commande l'accord mais avec un autre mot, dit *nom local*, plus proche de la cible dans la phrase (Hanse, 1987). Quelques erreurs d'attraction sont illustrées dans l'Exemple 1.

Ex. 1 : *La fille des voisins sont parti(e)s en voiture
*Il les emporteront en vacances
*Le pays dans laquelle la recherche se développe le plus (...)

Le paradigme expérimental que nous avons utilisé, dit de complètement de phrases, consiste à présenter aux locuteurs adultes des débuts de phrases qu'ils doivent compléter soit librement, soit en respectant certains critères tels que l'utilisation d'un verbe ou encore d'un adjectif particulier. Ces débuts de phrases contiennent toujours un nom local, de nombre ou de genre soit identique soit opposé à celui du nom sujet. Certaines propriétés conceptuelles, phonologiques et syntaxiques des noms sujets et des noms locaux ont ainsi été manipulées pour étudier leur impact sur l'opération d'accord.

Plusieurs recherches présentées dans cette thèse ont été réalisées en collaboration avec le Dr. Vigliocco de l'Université de Wisconsin-Madison (USA), actuellement chercheur invité à l'Institut Max Planck de psycholinguistique à Nimègues. Cette collaboration nous a permis d'aborder la problématique dans une perspective comparative, puisque plusieurs études parallèles à celles que nous avons nous-même menées en français ont pu être conduites en italien ou en

anglais¹. Nous pensons en effet que seule l'étude comparative de différentes langues divergeant quant à leurs propriétés structurales permettra d'avancer un cadre général des contraintes pesant sur le fonctionnement du langage. Cette perspective interlinguistique constitue un versant important des réflexions théoriques avancées pour rendre compte de nos observations.

L'Introduction générale (Chapitre 1) aborde la question de la modularité de la syntaxe en production de phrases. Le concept de module tel qu'il a été proposé par Fodor (1983) pour rendre compte de la compréhension du langage est examiné et appliqué à la production du langage. Nous présentons dans ce cadre les principaux modèles de production du langage réservant une place au traitement syntaxique. Cet aperçu théorique est suivi d'une revue de littérature dans laquelle sont exposées l'ensemble des observations issues des recherches expérimentales sur l'accord grammatical en production orale. La revue est articulée autour des hypothèses théoriques de l'input minimal et maximal. Certaines parties de la revue sont directement reprises d'un article actuellement en révision pour publication (Franck & Hupet, en révision, *L'Année Psychologique*).

Les Chapitres 2 à 4 traitent des *contraintes conceptuelles* sur l'accord. Le second chapitre présente les résultats de quatre expériences que nous avons menées en parallèle en français et en italien. Ces recherches ont mis en évidence le fait que la présence d'un corrélat conceptuel du genre sur le nom sujet réduit le risque d'erreurs d'accord en genre de l'adjectif prédicatif avec le sujet, comparativement à l'absence d'un tel corrélat. Ce chapitre a fait l'objet d'une publication avec le Dr. Vigliocco (Vigliocco & Franck, 1999, *Journal of Memory and Language*). Dans le Chapitre 3, nous montrons par le biais de quatre nouvelles expériences, également menées en français et en italien, que la discordance entre l'information conceptuelle contextuelle et l'information lexico-syntaxique est une source significative d'erreurs d'accord en genre. Ces recherches ont fait l'objet d'une présentation à la conférence de la CUNY (Franck & Vigliocco, 1998) et sont actuellement en révision pour publication (Vigliocco & Franck, under revision, *Journal of Memory and Language*). Dans le quatrième chapitre, une analyse critique de l'hypothèse d'input maximal avancée pour rendre compte des résultats précédents est avancée. Les résultats d'une expérience sur le français suggèrent que le genre conceptuel est différemment représenté du genre grammatical au niveau lexical, et que les effets conceptuels rapportés peuvent être interprétés en termes d'effets lexico-sémantiques (Franck & Vigliocco, in prep).

¹ L'ensemble des recherches rapportées en français ont été conçues et conduites par nous-même, tandis que celles rapportées en italien (dans les Chapitres 2 et 3) et en anglais (dans le Chapitre 6) ont été menées par le Dr. Vigliocco.

Le cinquième chapitre présente deux recherches conduites en français ainsi que deux études menées en parallèle en italien. Elles traitent *des contraintes morphophonologiques* pesant sur l'accord. Ces recherches montrent que des indices relatifs à la morphophonologie du genre du sujet exercent un contrôle sur l'accord dans ces deux langues. En français, il est montré que l'information morphophonologique de genre présente sur le déterminant du nom et sur sa terminaison favorise la réalisation correcte de l'accord en genre. Un effet similaire de l'information sur le suffixe du nom est rapporté en italien. Ces données ont fait l'objet d'une présentation à la conférence de l'ESCOP (Franck, Collina et Vigliocco, 1999) et sont actuellement en cours de rédaction pour publication.

Dans le Chapitre 6, nous exposons quatre études conduites en français et en anglais visant à mettre en évidence le rôle d'une *contrainte syntaxique* sur l'accord. Celle-ci concerne la structure syntaxique unissant les mots dans la phrase au stade de l'encodage grammatical. Nous montrons que dans ces deux langues la distance syntaxique entre les segments est un facteur déterminant de l'occurrence des erreurs d'accord, tandis que la distance linéaire n'a pas d'impact sur la réalisation de l'accord. Ces études nous ont également amenée à nous interroger sur le rôle de la complexité liée au traitement du pluriel. La recherche en français a fait l'objet d'une présentation à la conférence de la CUNY (Franck & Butterworth, 1997) et le chapitre est actuellement soumis pour publication (Franck, Vigliocco & Nicol, submitted, *Journal of Language and Cognitive Processes*). Dans le Chapitre 7, le concept de distance syntaxique comme facteur explicatif des observations rapportées dans le chapitre précédent est discuté plus en détails. Nous présentons une recherche exploratoire que nous avons menée sur le français visant à déterminer dans quelle mesure les résultats obtenus dans les quatre recherches précédentes peuvent être interprétés par une hypothèse alternative selon laquelle la saillance conceptuelle du nom local est le facteur déterminant des erreurs d'attraction.

Une synthèse théorique de l'ensemble des résultats est finalement proposée, dans laquelle nous tentons de fournir une interprétation globale de la construction de l'accord. L'objectif de cette synthèse est de rendre compte de l'apparente hétérogénéité des observations. Un premier facteur d'hétérogénéité relève de l'interprétation théorique des effets d'input maximal et minimal. Les différentes hypothèses explicatives sont mises en regard et discutées quant aux contraintes pesant sur la réalisation de l'accord et plus généralement à la question de la modularité des opérations syntaxiques. Le second facteur d'hétérogénéité est relatif à la variabilité intra- et interlinguistique. Il est examiné et interprété en référence aux notions de validité et de coût de l'utilisation des différentes informations.

Chapitre 1

Introduction générale

Dans cette introduction, nous commençons par exposer le débat sur la perméabilité de la syntaxe en production de phrases. Les propriétés des modules avancées par Fodor (1983) pour rendre compte du système de compréhension du langage sont examinées, ainsi que les implications pour l'étude de la production de la phrase. Les thèses sur les processus de production développées par les modularistes et les interactionnistes sont examinées, l'accent étant mis sur les premières qui sont les plus explicites quant aux processus syntaxiques en jeu en production de phrases.

Nous examinons ensuite les arguments théoriques et expérimentaux issus des travaux sur l'accord grammatical menés dans différentes langues et venant éclairer ce que nous avons appelé les hypothèses de l'input maximal et de l'input minimal. Selon l'hypothèse de l'input minimal, seule l'information syntaxique contrôle la réalisation de l'accord. Le système assurant le processus syntaxique d'accord est conçu comme cloisonné : ni la structure conceptuelle du message, ni les caractéristiques morphophonologiques ne peuvent exercer de contrôle sur son traitement. Cette isolation du système lui assure une certaine rapidité de fonctionnement ainsi qu'elle limite les risques d'interférences lorsque les différentes sources d'information ne coïncident pas. L'hypothèse de l'input maximal suppose que, lors de la construction de la structure syntaxique de la phrase, l'encodeur grammatical tire avantage de toutes les sources d'information pertinentes pour réaliser l'accord. Dans ce cadre, l'information syntaxique, mais également les informations issues des niveaux de traitement en amont et en aval du système, i.e. les niveaux conceptuel et morphophonologique, assurent le contrôle de la procédure d'accord. Il est avancé que la redondance d'information garantit une réalisation correcte de l'accord puisque la perte d'une information aura de moindres conséquences sur sa construction.

1.1 Le langage : un système modulaire ou interactif ?

En s'inspirant des développements de la théorie linguistique, les premiers pas de la psycholinguistique ont consisté à distinguer les différentes composantes pouvant intervenir dans le traitement du langage. Il a ainsi été suggéré que le système langagier comporte une composante *lexicale*, assurant la gestion des mots de la langue correspondant aux concepts, une composante *syntactique*, chargée du

traitement des combinaisons entre les mots en fonction des règles de la grammaire et de l'intention communicative du locuteur, et enfin une composante *phonologique*, dont la tâche est de spécifier la structure sonore des mots. Le message que le locuteur désire exprimer forme la composante dite *conceptuelle*. Bien que cette dernière ne soit pas de nature linguistique, elle constitue le matériau d'entrée sur lequel se base le système de production.

Tandis que l'identification de ces différentes composantes est adoptée à l'unanimité par les psycholinguistes, la question des relations que ces composantes, ou niveaux de traitement, entretiennent entre elles est encore aujourd'hui au cœur d'un vif débat. Sur le plan théorique, les controverses portent sur la définition des *limites fonctionnelles* de chaque niveau, sur la façon dont on envisage le *décours temporel* de chacun des traitements, et sur la spécification de la *direction du flux d'information* entre les niveaux de traitement. Ces différentes problématiques s'articulent autour d'une même question : celle de la modularité des processus cognitifs impliqués dans le traitement du langage. Il s'agit d'une question centrale à l'élaboration de tout modèle de traitement du langage, que cela soit en compréhension ou en production. La section suivante vise à éclairer le concept de modularité tel qu'il est envisagé par Fodor (1983). Nous examinons ensuite comment la psycholinguistique actuelle se situe par rapport à ce concept et les modèles proposés pour rendre compte de la production du langage.

1.1.1. Le concept de modularité

C'est le philosophe Jerry Fodor qui a sans doute théorisé de la façon la plus concise la question de la modularité des représentations et des processus cognitifs responsables du traitement de l'information (Fodor, 1983). Son ouvrage, visant à réactualiser la psychologie des facultés, propose de définir le programme de recherches dans cette perspective et les domaines auxquels l'appliquer. Il est à l'origine de nombreux débats et travaux en sciences cognitives dans des domaines tels que le langage, la vision, l'audition, le contrôle moteur ou encore le traitement des visages.

La première originalité de l'hypothèse de Fodor est de considérer que le langage possède des propriétés similaires à ce qu'il appelle les systèmes périphériques, et que donc tous deux doivent être étudiés selon le même programme. Les systèmes périphériques et le langage ont pour fonction de fournir de l'information aux systèmes centraux, i.e. la pensée. Plus précisément, ils constituent une interface entre ce que Fodor appelle les transducteurs (les systèmes sensoriels) et les mécanismes cognitifs centraux (la pensée). Les systèmes périphériques et le langage produisent des représentations qui spécifient la disposition des choses dans le monde ; il ne s'agit pas d'une traduction, comme c'est le cas pour les transducteurs, mais bien d'un premier traitement de l'information par lequel les systèmes périphériques effectuent des inférences sur les représentations que leur communiquent les transducteurs, pour les rendre

compréhensibles aux systèmes centraux. Ainsi, le langage, comme les systèmes périphériques, fournit à la pensée des symboles mentaux qui sont, par leur forme, des représentations adéquates pour les processus cognitifs. Pour assurer une telle fonction, le langage et les systèmes périphériques sont caractérisés par une série de propriétés. Nous en présentons ici les principales puisqu'elles constituent les critères de définition du fonctionnement modulaire.

Le langage est propre à un domaine. Les systèmes d'analyse perceptuelle du langage opèrent exclusivement sur des signaux qui sont considérés comme étant linguistiques. Autrement dit, seule une classe suffisamment restreinte de représentations peut déclencher leur opération. Par ailleurs, si un domaine de stimuli est excentrique, i.e. si son analyse exige un ensemble d'informations très particulières, alors les processus définis sur ce domaine correspondent à des systèmes computationnels spécialisés. Il a par exemple été montré qu'un signal qui est perçu comme le début d'une consonne dans un environnement linguistique sera perçu comme un sifflement lorsqu'il est présenté en isolation. Ce résultat suggère que les systèmes perceptuels responsables de l'analyse phonétique de la parole sont propres à un domaine, en ce sens qu'ils sont distincts des mécanismes responsables de l'analyse perceptuelle de signaux auditifs non linguistiques.

L'opération du langage est obligatoire. Elle est déclenchée automatiquement par les stimuli qui constituent son domaine d'application. Quand bien même il est demandé aux sujets de ne pas prêter attention à des stimuli langagiers, en situation d'écoute dichotique par exemple, il a été montré que l'information provenant du canal que les sujets doivent ignorer est quand même analysée, et cela sans que les sujets n'en aient vraiment conscience (ex. Lackner & Garrett, 1972). Il semble donc que l'information est automatiquement traitée par le système du langage, mais qu'elle n'ait plus accès à certains processus centraux. Sur le plan fonctionnel, Fodor souligne qu'il est indispensable que l'opération du langage, et plus généralement celle des systèmes périphériques, soit obligatoire puisqu'elle est l'unique voie par laquelle le produit des transducteurs parvient aux processus centraux.

Les systèmes centraux n'ont qu'un accès limité aux représentations calculées par le système du langage. Autrement dit, seul le produit final des traitements effectués par le langage est accessible aux processus cognitifs qui sous-tendent le comportement observable ; les niveaux intermédiaires sont inaccessibles à la conscience. La compréhension d'une phrase implique non seulement son codage acoustique, mais aussi le traitement de son contenu phonétique et lexical ainsi que de sa forme syntaxique. Le résultat du traitement à chacun de ces niveaux n'est pas connu ; l'information finale à laquelle l'individu accède, i.e. le sens du message, est la seule à laquelle il a accès. L'utilité fonctionnelle de cette propriété serait de limiter le coût en termes de mémoire et d'attention que nécessiterait la conservation des représentations intermédiaires. Plusieurs recherches semblent

effectivement indiquer que l'inaccessibilité relative des premiers niveaux de traitement du langage dépend avant tout de la façon dont le système est relié à la mémoire : certains niveaux sont en effet souvent accessibles moyennant un effort de mémoire et d'attention. Par ailleurs, des traitements de niveaux intermédiaires peuvent ne pas accéder à la conscience (ex. deux sons n'ayant pas de valeur contrastive dans la langue sont jugés comme identiques) tout en influençant le comportement final (ralentissement du temps de réponse). De façon générale, Fodor avance l'hypothèse que plus on est loin des transducteurs, plus les représentations sont accessibles aux systèmes cognitifs centraux.

Le langage est un système rapide. Par exemple, des études sur le pistage (qui consiste à répéter un flot de parole à mesure qu'il se déroule) montrent que les sujets sont capables à la fois de répéter et de comprendre ce qu'ils répètent en un intervalle de 250 ms entre le stimulus et la réponse (Marslen-Wilson, 1973). Cette propriété du langage est sans doute liée au fait qu'il s'agit d'une opération obligatoire ; parce que ces processus sont automatiques et ne tiennent compte que d'un sous-ensemble de choix computationnels, ils font l'économie du temps.

Le langage est informationnellement cloisonné. Fodor entend par là que les processus qui sous-tendent la compréhension du langage sont sous-déterminés par les données sensorielles et insensibles aux croyances et désirs de l'organisme, autrement dit au système conceptuel de la pensée. L'exemple donné par Fodor est que s'il dit « j'ai une girafe dans ma poche », nous n'aurons aucun mal à comprendre cet énoncé, bien que rien dans le contexte n'aurait pu prédire ni sa forme ni son contenu. Cette propriété d'insensibilité au contexte et plus généralement aux attentes du sujet, autrement dit, d'impénétrabilité cognitive, est nécessaire pour permettre au système de compréhension d'analyser des stimuli nouveaux, non redondants, comme c'est le cas des stimuli langagiers. Pour effectuer ses traitements, le système perceptuel du langage ne prend en compte qu'une partie de l'information qui pourrait en principe être pertinente. Le cloisonnement informationnel des modules place une contrainte supplémentaire sur la régulation du flux d'information, à savoir que l'information ne circule que dans un sens : du bas vers le haut (en ce qui concerne la compréhension du langage ; nous verrons plus loin ce qu'il en est de la production), sans possibilité de rétroaction. Cette affirmation est néanmoins fort délicate parce qu'elle repose entièrement sur la délimitation entre le traitement effectué par le module (le système de compréhension) et celui opéré par le système cognitif central, i.e. elle repose sur la détermination du niveau où se situe l'interface entre les processus de compréhension du langage et les processus cognitifs. Par exemple, si la représentation finale d'une phrase ne spécifie que son type (et non pas l'acte de langage accompli), alors l'observation d'un effet de contexte sur l'estimation des intentions communicatives ne démontrerait en rien la pénétrabilité cognitive du système de compréhension. Pour démontrer la pénétrabilité d'un système, il est nécessaire de spécifier la nature des représentations finales de ce système. Seule

l'observation d'une interaction entre les informations cognitives d'arrière-plan et les niveaux intermédiaires de traitement du module constitue un argument contre l'hypothèse de cloisonnement informationnel. En effet, il va de soi que l'information fournie par les systèmes périphériques doit, d'une façon ou d'une autre, être rendue compatible avec les connaissances d'arrière-plan du sujet. Tout le monde fait l'expérience de rectifier l'analyse proposée par l'input en fonction de ses croyances².

On ne peut donc pas séparer la question de l'évaluation du degré de cloisonnement informationnel des systèmes périphériques de la question de l'individuation de ces systèmes et de la caractérisation des représentations finales qu'ils fournissent. (Fodor, 1983, p. 100)

Fodor avance pour cela une autre propriété du système perceptuel du langage, à savoir que *le produit final de son traitement est superficiel*. Par superficiel, il signifie que les représentations fournies par ce système sont limitées. Cette propriété est nécessaire pour avoir 'une chance' de reconnaître la modularité du système ; si les représentations qu'il construit sont trop profondes, alors il devient difficile de soutenir le cloisonnement informationnel par rapport au système conceptuel. Mais qu'entend par superficiel ? Pour répondre à cette question, Fodor suggère de la formuler en termes de *quel est le maximum qu'un système cloisonné puisse calculer ?* autrement dit, quels aspects d'un stimulus peuvent être analysés sans faire appel aux données d'arrière-plan ? L'analyse de la forme syntaxique devrait à l'évidence faire partie des calculs effectués par le système de compréhension du langage. L'argument principal pour affirmer que la reconnaissance de la forme linguistique ne peut pas être orientée par le contexte est que ce dernier ne détermine pas la forme, autrement dit, il n'y a pas (beaucoup) de redondance entre les variables contextuelles et la forme syntaxique de l'énoncé. En ce qui concerne le lexique, la reconnaissance superficielle des mots serait assurée par le système modulaire (déterminer si oui ou non il s'agit d'un mot de la langue) mais ce système n'effectuerait pas l'analyse sémantique interne des unités lexicales. En résumé, la sortie du système d'analyse du langage définirait la structure syntaxique ainsi que les mots qu'elle comporte sans spécifier encore la représentation sémantique de l'énoncé.

Dans la perspective développée par Fodor, l'affirmation qu'un système comme la compréhension du langage est informationnellement cloisonné doit être soigneusement distinguée de l'hypothèse que l'information circule de haut en bas à *l'intérieur* de ce système. Considérons le résultat selon lequel la restitution de

²Le raisonnement peut facilement devenir circulaire : on veut démontrer qu'un système est modulaire et pour cela qu'il possède la propriété d'être cloisonné ; dès lors, pour rendre compte des données indiquant l'influence des informations d'arrière-plan sur ce système, on restreint son champ d'application de sorte que l'effet d'interaction observé porte sur une représentation qui ne ressortit plus du module. Pour éviter une telle circularité méthodologique, il est indispensable de spécifier *a priori* les limites du système. Nous verrons que dans le domaine de la production, la circularité méthodologique constitue parfois un obstacle à la compréhension du fonctionnement du système (voir Levelt, Roelofs & Meyer 1999).

phonèmes est facilitée par le fait que le phonème fait partie d'un mot de la langue. Cet effet semble indiquer que le mécanisme de reconnaissance des phonèmes est sensible aux connaissances lexicales des sujets, et donc que l'information circule de haut en bas au cours de la perception de la parole. Ce résultat n'est cependant pas en défaveur de l'hypothèse du cloisonnement informationnel puisque le niveau de représentation lexicale fait lui aussi partie du système de compréhension du langage ; l'information peut donc circuler de haut en bas à l'intérieur même du module du langage. Il est donc indispensable de distinguer la question de la pénétrabilité cognitive (du cloisonnement) et celle des effets internes au module, ces deux questions relevant toutes deux de la définition des limites du système.

1.1.2. La perspective modulaire de la production

Jusqu'à présent, le seul aspect du langage analysé par Fodor a été le système de compréhension. Les opérations effectuées par ce système, que Fodor postule être modulaire, sont propres à un domaine, obligatoires, elles ne sont que partiellement accessibles aux processus centraux, elles sont rapides, informationnellement cloisonnées et leur résultat final est superficiel. Dans son ouvrage, l'auteur ne réserve que très peu d'attention aux processus de production du langage, qui selon lui relèvent des processus cognitifs centraux et n'entrent donc pas dans la problématique de la modularité. Comme il l'écrit :

Nous utilisons le langage (entre autres moyens) pour donner notre avis sur le monde. Mais cela n'est possible que si les mécanismes de production de la parole ont accès à ce qui nous permet de voir (d'entendre, de nous rappeler, de penser) le monde tel qu'il est. Puisque par hypothèse, ces mécanismes constituent une « interface » entre des facultés verticales (i.e. les systèmes périphériques et la compréhension du langage, ndla), ils ne peuvent être eux-mêmes propres à un domaine. (Fodor, 1983, p. 134).

Pourtant, une série d'arguments en faveur du cloisonnement informationnel en production de phrases peuvent être avancés, tant sur le plan théorique qu'empirique. En se basant sur l'analyse des erreurs spontanées présentes dans le discours de sujets adultes sains, le psycholinguiste Merrill Garrett a ainsi suggéré d'adapter le concept de modularité au système de production du langage. Suivi par d'autres auteurs (ex. Bock, 1987 ; Bock & Levelt, 1994 ; Levelt, 1989 ; Levelt, Roelofs, & Meyer, 1999), et également soutenu par une série de données expérimentales, Garrett propose un modèle en niveaux de traitements, où les niveaux conceptuel et syntaxique sont séparés et où l'impact de facteurs conceptuels sur les processus syntaxiques est strictement limité. Par ailleurs, il pousse encore plus loin la décomposition en modules proposée par Fodor en suggérant un découpage plus fin des systèmes de traitement lexical et syntaxique : ceux-ci sont eux-mêmes décomposés en sous-systèmes, dont les propriétés s'apparentent à celles des modules. Ce modèle modulaire de la production de phrases est illustré par la Figure 1.

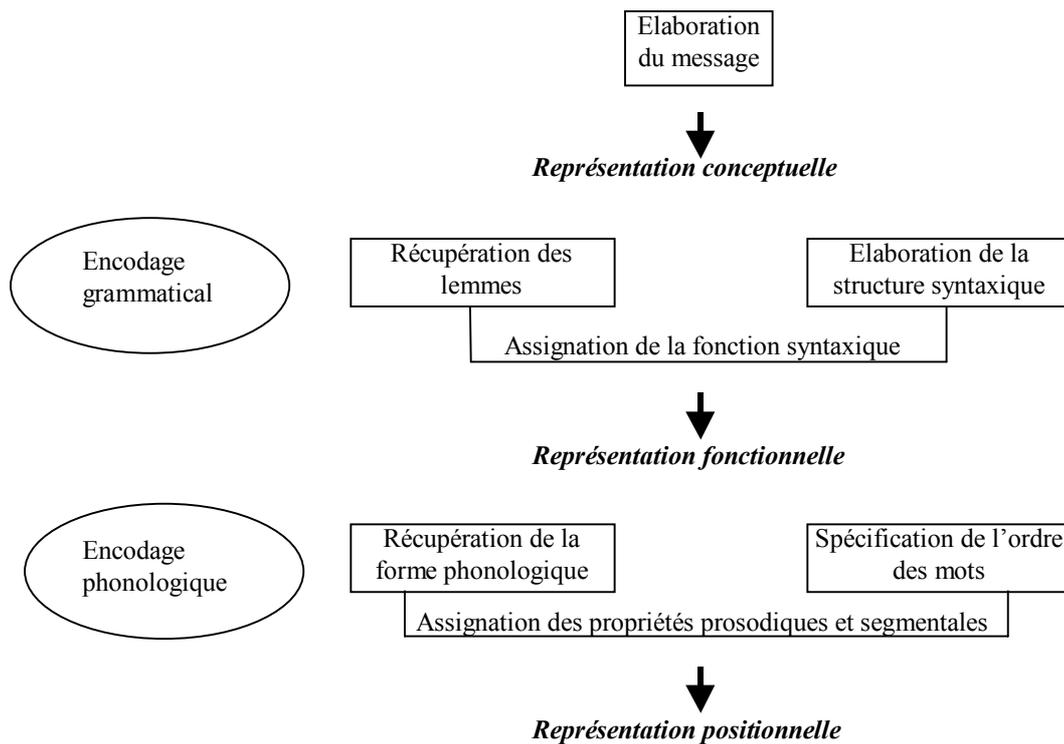


Figure 1. Un modèle modulaire de production, adapté de Garrett (1988).

Le premier stade de la production consiste à élaborer une structure du sens et de l'intention communicative du locuteur. Cette structure, qui n'est pas encore de nature linguistique, constitue le matériel d'entrée sur lequel opère le système de production du langage. Sur base de la structure conceptuelle, des représentations abstraites des mots correspondant aux concepts sont sélectionnées dans le lexique. Selon un certain nombre d'auteurs (ex. Garrett, 1982, Levelt, 1989), ces représentations, appelées lemmes, spécifient l'information syntaxique du mot, à savoir sa catégorie grammaticale (nom, verbe, adjectif, etc.), sa fonction grammaticale (sujet, objet, etc.), le type de structure syntaxique dont il peut faire partie (syntagme nominal, syntagme verbal, etc.), ainsi que certaines caractéristiques spécifiques à la langue (ex. le genre). Une série d'auteurs considèrent également que les lemmes sont organisés au sein de champs sémantiques et spécifient ainsi le sens du mot (Garrett, 1982, 1992, Butterworth, 1983). A l'issue de ce stade de la production dit d'encodage grammatical, la structure fonctionnelle de la phrase est élaborée. C'est seulement au stade suivant qu'a lieu l'encodage phonologique au cours duquel la forme phonologique du mot, aussi appelée lexème, est récupérée du lexique, ainsi que les caractéristiques prosodiques et segmentales de chaque mot. L'ordre final linéaire dans lequel ils seront arrangés est déterminé et les morphèmes grammaticaux (ex. déterminants, flexions de nombre, prépositions, etc.) sont insérés dans ce cadre. La structure

issue de ce stade est dite positionnelle puisque chaque élément, spécifié pour son articulation, est positionné dans la phrase³.

Nous l'avons dit, ce modèle avancé par Garrett visait tout d'abord à rendre compte des lapsus produits en situation d'élocution spontanée. Un premier type d'erreurs analysé concerne les substitutions de mots. Garrett (1993) constate que ces erreurs impliquent soit deux mots entretenant une relation de sens (ex. en français), soit deux mots entretenant une relation au niveau de leurs sons (ex. en français). De cette observation, il déduit qu'il doit exister deux composantes différentes impliquées dans la récupération d'un mot : la première est responsable de récupérer les caractéristiques sémantiques du mot (ce qu'il est convenu d'appeler le lemme), la seconde est responsable de récupérer ses caractéristiques phonologiques (la forme du mot). Un dysfonctionnement dans le travail de la première composante se marquera par une similarité sémantique entre le mot récupéré et le mot cible, tandis qu'une erreur au niveau de la seconde composante se traduira par une similarité phonologique entre ces deux unités.

Une série d'observations viennent confirmer l'hypothèse d'une séparation du travail de récupération des lemmes et de la forme phonologique, comme des études sur les 'mots sur le bout de la langue' (TOT, Tip Of the Tongue). Ces études ont montré que des locuteurs peuvent révéler certaines propriétés sémantiques et syntaxiques de mots (comme le genre grammatical) qu'ils sont pourtant incapables de dénommer, autrement dit dont ils ne parviennent pas à récupérer la forme phonologique (ex. Vigliocco, Antonini & Garrett, 1997 ;). Par ailleurs, certains patients aphasiques anomiques présentent, eux aussi, des difficultés à trouver leurs mots tout en maîtrisant les caractéristiques sémantiques, syntaxiques ou même relatives à la structure syllabique. Ces résultats indiquent à nouveau que le traitement phonologique peut ne pas être assuré sans que cela ne compromette ni le traitement des caractéristiques sémantiques, ni même le traitement des caractéristiques syntaxiques ou syllabiques des mots.

Outre la distinction entre une composante de récupération des lemmes et une composante de récupération de la forme phonologique, le modèle spécifie également que ces deux composantes opèrent de manière strictement séquentielle. Il semble aller de soi que la récupération du sens doit précéder celle de la forme, puisque le locuteur, en produisant sa phrase, commence par se former une idée claire du sens de ce qu'il va dire. Des données expérimentales viennent conforter cette hypothèse. Il a ainsi été montré qu'un mot amorceur lié sémantiquement à un mot cible à dénommer sur la base d'une image peut, lorsqu'il est présenté *avant* la présentation de l'image, inhiber la récupération du mot cible, tandis qu'un mot relié phonologiquement à la cible facilite sa récupération s'il est présenté *après* la présentation de l'image (ex. Schriefers, Meyer & Levelt, 1990). Des

³ Les termes de structure fonctionnelle et positionnelle proviennent de Garrett (ex. 1988) tandis que les termes d'encodage grammatical et phonologique sont de Levelt (ex. 1989).

études récentes en imagerie cérébrale relatives aux potentiels de préparation latéralisés (Lateralized Readiness Potentials) menées par Van Turennout et ses collaborateurs montrent, quant à elles, des réponses cérébrales à des activations sémantiques et phonologiques distinctes et séquentielles (Van Turennout, Hagoort & Brown, 1997).

Les propriétés syntaxiques des lemmes sont, elles aussi, récupérées *avant* la spécification de la forme phonologique dans ce modèle. Des mesures électrophysiologiques (LRPs) soutiennent cette hypothèse, indiquant que le genre du nom est accessible avant la récupération de son phonème initial (Van Turennout, Hagoort & Brown, 1999). Toutefois, la sélection des propriétés syntaxiques serait conditionnée par la présence d'un environnement syntaxique nécessitant leur utilisation. Une série de recherches ont ainsi été menées sur l'accès au genre des noms en utilisant le paradigme d'interférence image-mot. Les participants doivent dénommer des images précédées d'un mot (distracteur) ayant soit le même genre que le mot cible à dénommer, soit un genre différent. Il a été montré que la présentation d'un mot ayant le même genre que le mot cible accélèrait la récupération de ce dernier mais seulement lorsque les participants devaient produire un syntagme nominal complet, i.e. contenant le nom plus un article et/ou un adjectif (ex. Schriefers, 1993, van Berkum, 1997). La production d'un nom isolé n'est, quant à elle, pas facilitée par la présentation d'un distracteur du même genre (LaHeij, Mak, Sander & Willeboordse, 1998). Cependant, Schriefers et Teruel (cité dans Schriefers & Jescheniak, 1999) ont montré que l'effet de facilitation lors de la production de syntagmes nominaux apparaît même si les autres mots du syntagme à produire ne portent pas de marque de genre sur le plan phonologique (ex. *l'*-masculin/féminin *assiette jaune*-masculin/féminin). L'interprétation de ces observations par les tenants du modèle séquentiel est, d'une part, que la propriété de genre est sélectionnée uniquement si l'environnement syntaxique local nécessite son utilisation (l'accord de l'article ou de l'adjectif) et d'autre part, que dans de telles conditions cette sélection est automatique, même si le genre n'apparaît pas dans la structure phonologique de surface des mots et n'est donc pas requis pour l'encodage phonologique (article et adjectif non marqués pour le genre).

En ce qui concerne le système de récupération des lemmes, il serait lui-même constitué de deux sous-composantes distinctes et séquentielles : la première, responsable du traitement des propriétés lexicales sémantiques (lexico-sémantiques), la seconde, responsable de traiter les propriétés lexicales syntaxiques (lexico-syntaxiques). En utilisant le paradigme d'interférence image-mot, Schriefers (1993) a ainsi mis en évidence le fait qu'une interférence sémantique avec un distracteur précède l'interférence syntaxique (a lieu pour des temps de présentation antérieurs). Ces résultats montrent que l'effet d'amorçage des traits lexico-sémantiques a lieu avant l'effet d'amorçage des traits lexico-syntaxiques, ce qui semble indiquer que le traitement de ces deux caractéristiques se fait en deux

étapes distinctes et successives. Notons que cette décomposition de la récupération lexicale est peu discutée dans la littérature à notre connaissance ; elle n'est par conséquent pas représentée dans la Figure 1

Les observations rapportées jusqu'ici semblent plaider en faveur de l'existence de deux systèmes distincts ; l'un responsable de la récupération des propriétés sémantiques et syntaxiques abstraites du mot (lemme), l'autre assurant la récupération de sa forme phonologique (lexème). Ces deux processus sont séparés et considérés comme informationnellement cloisonnés. Le cloisonnement est implémenté dans le modèle par la contrainte de *séquentialité unidirectionnelle* pesant sur le décours temporel de la production : chaque niveau de traitement détermine uniquement le niveau qui le suit directement et est lui-même uniquement déterminé par celui qui le précède. Le flux d'information se fait obligatoirement de haut en bas et seulement entre un niveau et celui qui lui succède immédiatement. La représentation du message détermine la sélection des lemmes, mais une fois cette sélection opérée, cette représentation n'a plus aucun contrôle sur les opérations suivantes, en l'occurrence sur la sélection de la forme phonologique. Par ailleurs, le flux d'information se faisant de haut en bas, la structure phonologique des mots n'influence en rien la sélection des lemmes. Ainsi, le module responsable de la récupération des propriétés sémantiques du mot reçoit comme input la représentation conceptuelle du message mais est totalement imperméable aux informations relatives à la forme du mot ; il n'y a pas de rétroaction possible. Le système élaborant la récupération de la forme phonologique opère sur base de l'output de la structure fonctionnelle (nous reviendrons sur ce point), il est par contre imperméable aux représentations conceptuelles puisque celles-ci ne le précèdent pas directement (pas d'action des représentations de deux étages en amont). La modularité de ces deux systèmes est illustrée dans la Figure 2.

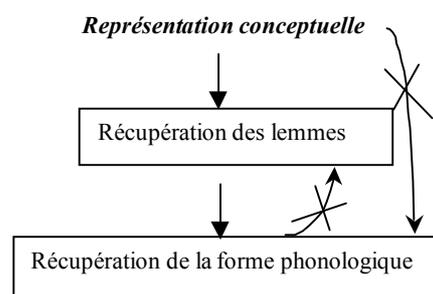


Figure 2. Illustration de la restriction de l'input dans les deux systèmes de récupération lexicale.

La question de la modularité des composantes de la production n'a jusqu'à présent été abordée qu'en ce qui concerne les processus de récupération lexicale, localisés dans les boîtes de la moitié gauche du modèle de la Figure 1. Les deux boîtes de la moitié droite du modèle contiennent quant à elles des processus de

traitement responsables de la construction syntaxique de la phrase. Voyons tout d'abord ce qui justifie la distinction entre deux niveaux de traitement syntaxique. A nouveau, la première source d'observation sur laquelle se base Garrett est l'analyse d'erreurs spontanées, et en particulier ici des erreurs dites d'échange. Ces erreurs consistent à permuter deux unités linguistiques au sein de la phrase, comme illustré en Ex. 2.

Ex. 2 : **Most cities are true of that (pour : That is true of most cities) (Garrett, 1993)**
La plupart des villes sont vraies de ça (pour: C'est vrai de la plupart des villes)

Les erreurs d'échange peuvent, comme les erreurs de substitution, se diviser en deux classes. Les erreurs dans la première classe sont caractérisées par le fait qu'elles impliquent des unités de la classe dite ouverte (Nom, Verbe, Adjectif, Adverbe), que l'échange porte sur des unités de la même catégorie grammaticale (un nom est échangé avec un nom, un verbe avec un verbe, etc.), qu'il n'y a pas de similarité phonologique entre les unités échangées, et que les unités impliquées ont tendance à faire partie de la même proposition mais pas spécialement du même syntagme. Un exemple est présenté en (Ex. 3).

Ex. 3 : **You ordered up ending some fish dish (pour : ended up ordering) (Garrett, 1993)**
Tu commandas par finir un plat de poisson (pour: finiras par commander)

Les erreurs de la seconde classe quant à elles portent plutôt sur des unités de la classe fermée (ex. déterminants, prépositions, flexions, dérivations), elles ne respectent pas la catégorie grammaticale des unités échangées, les deux unités échangées sont souvent proches sur le plan phonologique, et elles se situent au sein du même syntagme ou du même mot. Ce type d'erreur est illustré en (Ex. 4).

Ex. 4 : **That's why they sell the cheaps drink (pour : the drinks cheap) (Garrett, 1993)**
C'est pourquoi ils vendent les bons marché-Pl boisson-Sg (pour: les boissons bon marché)

L'observation de ces deux classes d'erreurs est à l'origine de la distinction entre les niveaux de traitement syntaxique fonctionnel et positionnel. Au cours du premier stade, fonctionnel, les unités lexicales de la classe ouverte, spécifiées pour une série de caractéristiques syntaxiques (ex. le nombre, le genre, la personne, etc.), reçoivent le rôle syntaxique qu'elles devront jouer dans la phrase (ex. Sujet, Objet). Ces unités sont insérées dans une structure de type hiérarchique, que l'on représente sous la forme d'un arbre syntaxique. La structure est dite fonctionnelle parce que les fonctions syntaxiques des mots sont déterminées et parce que cette structure, de nature hiérarchique, spécifie les relations syntaxiques entre les mots. Comme nous le verrons dans le Chapitre 6, l'organisation hiérarchique des segments syntaxiques joue un rôle déterminant dans la réalisation de procédures syntaxiques telles que l'accord ayant lieu à ce stade. L'observation que la grande majorité (80% d'après Garrett, 1980) des échanges de mots se fait entre deux mots

situés au sein de la même proposition suggère que l'unité d'encodage grammatical est la proposition⁴. L'unité d'encodage peut être définie comme :

Such units may be taken to constitute the minimal domain toward which syntactic encoding efforts can be individually directed. (Bock & Cutting, 1992, p. 101)

D'autres données viennent renforcer l'hypothèse que la proposition constitue l'unité d'encodage grammatical, comme par exemple le fait que la détection d'un signal externe apparaissant au sein de la proposition soit reportée à la fin de celle-ci (Ford & Holmes, 1978), la présence de pauses aux frontières propositionnelles (Cooper, Paccia, & Lapointe, 1978) ou encore la faible probabilité de détection d'une erreur au sein de ses propres productions lorsqu'elle se situe en début de proposition (Levelt, 1983). Une étude menée par Bock & Cutting (1992) montre par ailleurs que les erreurs d'accord du verbe avec un nom autre que le sujet de la phrase surviennent principalement lorsque ce mot se situe dans la même proposition que le sujet (Ex. 5), tandis que la présence d'une frontière propositionnelle isolant cet autre nom du sujet réduit fortement le risque d'erreur (Ex. 6). Les éléments présents au sein d'une même proposition semblent facilement entrer en interférence, étant donné qu'ils sont traités en même temps, au sein de la même unité d'encodage.

Ex. 5 : *The claim about the stolen babies were rejected
*L'affirmation à propos des bébés volés ont été rejetées

Ex. 6 : *The claim {that wolves were stealing babies} were rejected
*L'affirmation {que les loups volaient des bébés} ont été rejetées

Au cours du second stade, dit positionnel, les unités syntaxiques, jusque-là organisées fonctionnellement, reçoivent leur position dans la structure linéaire finale de la phrase. C'est également à ce stade que sont récupérés les morphèmes grammaticaux dans la phrase. Les morphèmes grammaticaux sont des unités strictement linguistiques. Ils se divisent en morphèmes grammaticaux libres, i.e. les mots de la classe fermée et en morphèmes grammaticaux liés. Ces derniers sont dits liés parce qu'ils n'apparaissent qu'en présence d'une racine lexicale et n'ont donc pas d'existence indépendamment d'un mot de la classe ouverte. Parmi les morphèmes liés, on distingue les morphèmes flexionnels (ex. la flexion *-s* pluriel ; *fille-s*) et les morphèmes dérivationnels (ex. la dérivation *-eur* ; *serv-eur*). C'est aussi à ce niveau que les caractéristiques sonores des spécifications sémantiques et syntaxiques sont encodées (ex. la liaison du *-s* pluriel avec le mot suivant, l'intonation interrogative). Puisque les traitements opérés au stade positionnel sont à la fois de nature phonologique (ils ne concernent pas des unités de sens) et morphologique (ils portent sur la structure du mot), ils sont souvent qualifiés de processus *morphophonologiques*.

⁴ La proposition est une unité syntaxique (réduite parfois à un seul mot) qui constitue à elle seule une phrase simple ou qui entre comme élément dans la formation d'une phrase complexe (Le Nouveau Petit Robert, 1993).

La question de la modularité se pose à nouveau en ce qui concerne ces composantes. En particulier, il s'agit de déterminer quelles relations le système responsable de l'élaboration de la structure syntaxique de la phrase (au niveau fonctionnel) entretient avec les autres composantes de la production : avec le système responsable de la sélection des lemmes, avec la représentation conceptuelle de l'énoncé (situé en amont dans le modèle), et avec le système responsable du traitement morphophonologique au niveau positionnel de la phrase (situé en aval dans le modèle). Dans le modèle de Garrett, la récupération des lemmes et la construction de la structure syntaxique sont situées au même niveau, le niveau dit d'encodage grammatical, responsable d'élaborer la structure fonctionnelle de la phrase. Ces deux processus sont pourtant fonctionnellement distincts, puisque l'on trouve dans les erreurs spontanées deux types d'erreurs que nous avons décrites plus haut ; les erreurs de substitution, de nature lexicale, et les erreurs d'échange, de nature syntaxique. Certaines données expérimentales confortent également cette distinction. Bock (1986) a par exemple montré qu'il était possible d'obtenir un effet d'amorçage de la structure syntaxique indépendamment du contenu lexical des mots impliqués.

Mais quel est le lien entre ces deux composantes ? La plupart des auteurs s'accordent actuellement à reconnaître le principe de *guidance lexicale* de la construction syntaxique : le lexique contient des propriétés que le système syntaxique doit nécessairement prendre en compte pour élaborer un cadre pour la phrase (ex. Bock, 1987 ; Kempen & Hoenkamp, 1987). Les formulations récentes de l'influente théorie chomskyenne du gouvernement et du liage (*Government and Binding Theory*, ex. Chomsky, 1981) sont, elles aussi, basées sur l'hypothèse que les unités lexicales déterminent directement la construction de la représentation syntaxique (théorie des X-bar). Dans ce cadre, l'information sur laquelle se base le système syntaxique est stockée dans les unités lexicales, et la frontière entre lexique et syntaxe devient difficile à déterminer. Ces contraintes lexicales concernent à la fois des propriétés syntaxiques (ex. le genre grammatical du nom), des propriétés sémantiques déterminant certaines structures syntaxiques (ex. le nombre de compléments obligatoires pour chaque verbe) et des propriétés combinatoires, relatives à la combinaison des mots (contraintes de sous-catégorisation comme la transitivité du verbe, le choix de la préposition suivant l'adjectif, etc. La composante lexicale et la composante syntaxique sont néanmoins distinctes et strictement séquentielles dans le modèle de production de Garrett. La sélection des lemmes précède la construction phrastique et n'est pas influencée par elle. Le principe de guidance lexicale tel que le conçoit Garrett est que la construction phrastique est guidée par une portion limitée de la composante lexicale, à savoir uniquement par ses représentations finales. Seule l'information lexico-syntaxique pénètre donc la composante de construction de phrases ; cette dernière est imperméable aux propriétés sémantiques des lemmes.

(...) a clear demonstration of a processing stage/representation that is sensitive to phrasal configuration but not to the semantic force of the lexical content. (Garrett, in press)

Qu'en est-il du contrôle conceptuel sur la construction syntaxique ? L'hypothèse défendue par Garrett ainsi que d'autres auteurs comme Bock (ex. 1987) est que la structure conceptuelle détermine le stade initial de la construction phrastique, à savoir l'assignation des fonctions syntaxiques, mais qu'une fois cette assignation effectuée, la structure conceptuelle n'a plus d'influence sur les mécanismes syntaxiques. Les recherches les plus pertinentes visant à déterminer le contrôle conceptuel sur les mécanismes syntaxiques subséquents à l'assignation des fonctions syntaxiques portent sur le mécanisme d'accord grammatical. Etant donné qu'elles constituent le point de départ de plusieurs des travaux effectués dans le cadre de notre thèse, une revue de littérature leur est consacrée dans la seconde partie de cette introduction.

Nous avons vu que le modèle sépare le traitement syntaxique en deux niveaux, i.e. le niveau fonctionnel et le niveau positionnel, pour rendre compte des deux classes d'erreurs spontanées relevées dans les corpus. Sur le plan théorique, Garrett, suivi de Bock et d'autres auteurs, postule à nouveau le cloisonnement informationnel typique à l'organisation modulaire : le niveau positionnel opère sur la base de l'output de l'encodage grammatical, à savoir la représentation fonctionnelle de la phrase, mais cette dernière n'est en rien influencée par le traitement morphophonologique assuré au stade positionnel. L'analyse des erreurs semble conforter une telle hypothèse, mais elle a été fort peu abordée sur le plan expérimental. Toujours dans la même étude visant à amorcer la structure syntaxique, Bock (1986) a montré qu'il était impossible d'obtenir un tel effet d'amorçage par la forme phonologique du mot. A nouveau, ce sont les quelques études récentes sur l'accord qui procurent les informations les plus pertinentes sur la question, elles sont présentées plus loin dans cette introduction.

Nous avons jusqu'ici tracé les grandes lignes théoriques du modèle de production postulant l'existence de différents systèmes de traitement se caractérisant par un mode de fonctionnement modulaire, i.e. séquentiel et cloisonné. L'information guidant l'opération de chaque système est limitée au minimum, à savoir l'output du système le précédant immédiatement, sans possibilité de rétroaction de l'information des systèmes subséquents. Un premier avantage d'un tel cloisonnement est la rapidité du traitement effectué par ces modules, puisque ces traitements ne reposent que sur une portion limitée d'information. L'organisation modulaire du système procure un second avantage qui est de limiter les risques d'interférences entre les différentes sources d'information, puisqu'elles n'entrent pas toutes en ligne de compte pour le calcul des représentations à chaque niveau. Ce modèle ne semble toutefois pas compatible avec un certain nombre de données, ce qui a amené une série d'auteurs à envisager un autre type de fonctionnement du système de production.

1.1.3. La perspective non modulaire de la production

La conception modulaire de la production, à l'origine développée par Garrett et reprise ensuite par de nombreux auteurs, est la première à avoir été avancée au sein d'un modèle pour rendre compte des processus de production, et est encore aujourd'hui adoptée par de nombreux psycholinguistes. Ce modèle propose des prédictions relativement claires quant aux fonctions opérées par chaque module ainsi qu'aux relations qu'ils entretiennent entre eux. La majorité des travaux en production de phrases se sont donc inscrits dans ce cadre théorique. Certains chercheurs, surtout dans le domaine de la perception du langage, se sont rapidement opposés à la conception strictement séquentielle défendue par ce modèle qui exclut tout retour en arrière une fois le traitement d'un niveau effectué. (ex. Morton, 1969 ; McClelland & Rumelhart, 1981). En ce qui concerne les processus de production, l'hypothèse que différents traitements peuvent être effectués en parallèle aux différents niveaux, avec effets de rétroaction d'un niveau sur l'autre s'est ainsi progressivement imposée. Elle reste toutefois essentiellement centrée sur l'explication des processus lexicaux (ex. Dell, 1986), bien que certains auteurs l'aient également étendue aux processus syntaxiques (ex. Stemberger, 1985a, 1985b).

Ces modèles alternatifs reprennent l'idée d'une décomposition du système en niveaux, tout en intégrant, comme principe de fonctionnement, l'hypothèse que les unités, à chaque niveau et entre les niveaux, forment un réseau interconnecté, similaire à l'organisation cérébrale (i.e. les neurones forment les unités, les synapses, les relations entre ces unités). Le flux d'information se fait par *activation* positive ou négative (*inhibition*) des unités. L'activation se propage d'une unité à l'autre au sein du réseau. Le degré d'activation d'une unité est fonction du nombre de connexions venant l'activer et de son niveau d'activation au repos (*resting level*) de chaque unité, c'est-à-dire de son niveau d'activation lorsqu'elle n'est pas utilisée. Pour qu'une unité soit finalement sélectionnée, il faut que son niveau d'activation soit suffisamment élevé. Un autre principe important est que les unités concurrentes au sein d'un même niveau (ex. les différents synonymes pouvant exprimer un même concept) tendent à s'inhiber mutuellement au sein du réseau qui les relie, tandis que les unités situées à des niveaux différents tendent à s'activer l'une l'autre dans le réseau. Ce principe est nécessaire pour rendre compte de la sélection finale d'un élément à chaque niveau (sinon le système tournerait sur lui-même) et de la progression du flux d'information d'un niveau de traitement à l'autre.

Un exemple de modèle interactif de la production de la phrase est illustré dans la Figure 3. Le processus de production commence lorsque le locuteur formule une intention activant une série d'unités sémantiques et pragmatiques. Ces unités activent à leur tour toutes les unités lexicales reliées à ces unités sémantiques et pragmatiques. Le niveau d'activation de chaque mot dépend de son activation au

repos mais aussi de la quantité de traits sémantiques qu'il possède en commun avec le concept à exprimer. C'est seulement via la sommation de l'ensemble des activations provenant de la structure sémantique qu'un mot particulier sera sélectionné. Ces unités lexicales activent ensuite les différentes unités phonémiques qui les composent et qui en retour, viennent activer, par rétroaction, les unités lexicales auxquelles elles sont associées, c'est-à-dire celles qui contiennent un ou plusieurs de ces phonèmes.

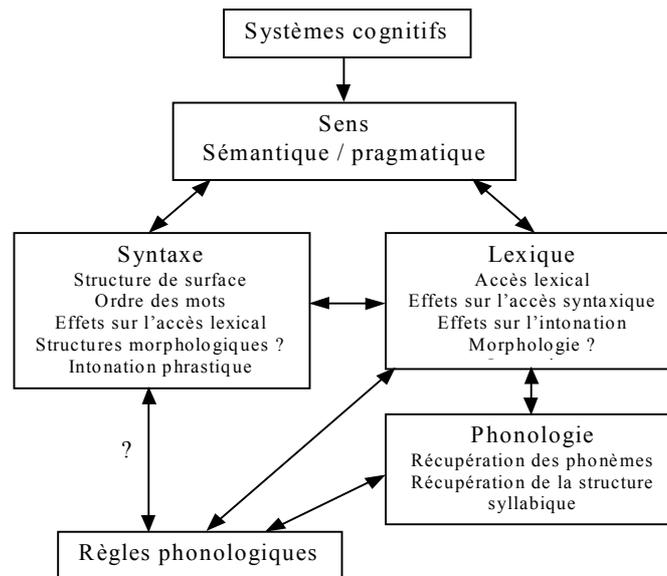


Figure 3. Un modèle de production interactif (d'après Stemberger, 1985a, p. 151).

L'hypothèse d'une rétroaction des constituants phonémiques des mots sur leur sélection lexicale a originellement été proposée pour rendre compte du fait que les mots impliqués dans les erreurs de substitution sémantique présentaient souvent des similarités phonologiques, dans des proportions dépassant le niveau de la chance (Dell & Reich, 1981), une observation répliquée sur le plan expérimental (ex. Levelt & Maasen, 1981). En outre, les mêmes auteurs notent que le résultat des erreurs spontanées de substitution phonémique est dans la majorité des cas un mot de la langue (Dell & Reich, 1981). Un tel biais lexical est facilement expliqué par l'existence de rétroactions du niveau de la forme phonologique sur le niveau de la sélection lexicale. Par ailleurs, une étude de corpus effectuée par Berg (1992) indique que les erreurs de substitution phonémique non contextuelles respectent dans 81,4% des cas le genre du mot, en ce sens que le mot cible et le mot produit sont de même genre. Cette contrainte syntaxique qui semble opérer sur la survenue des erreurs phonologiques peut aussi être expliquée par l'existence de rétroactions : l'activation du genre du lemme se propage à l'ensemble des unités phonologiques liées à ce genre, qui à leur tour se propagent vers le haut en activant les lemmes de même genre, augmentant ainsi la probabilité de sélectionner erronément un mot de

même genre, comparativement à la probabilité d'une sélection d'un mot de genre différent.

Les unités du niveau syntaxique du modèle sont les règles de réécriture (ex. $P = SN + SV$, etc.). La sélection d'une structure syntaxique procède de la même façon que pour les unités lexicales. Les règles de réécriture sont activées par les unités sémantiques et pragmatiques auxquelles elles sont associées. Ce sont d'autres aspects qui les activent que ceux impliqués dans l'activation du réseau lexical. Il s'agit par exemple des rôles syntaxiques ou du caractère défini/indéfini. Chaque règle de réécriture active les autres unités, comme par exemple P (Phrase) active SN (Syntagme Nominal) et SV (Syntagme Verbal). La structure recevant la plus forte activation est finalement sélectionnée, et inhibe alors les structures concurrentes. La détermination de l'ordre des mots se fait directement sur la base des règles de réécriture ; il n'y a pas de distinction entre une structure profonde et une structure de surface. Ce niveau procure donc un cadre syntaxique composé d'une série de *slots* (trous) dans lesquels sont insérés les mots.

Quelles sont les relations entre la composante lexicale et la composante syntaxique dans un tel modèle ? Les deux niveaux opèrent *simultanément* et *en interaction*. Nous l'avons vu dans la présentation du modèle modulaire ; les unités lexicales possèdent certaines caractéristiques qui déterminent la structure de la phrase. Le principe de guidance lexicale qui caractérise le modèle modulaire est également adopté ; la sélection d'une unité lexicale contraint la sélection d'une structure syntaxique particulière. Par ailleurs, contrairement au modèle modulaire, la sélection d'une structure syntaxique joue un rôle important dans la sélection lexicale en ce qu'elle contraint l'ordre dans lequel les mots doivent être sélectionnés. Cette contrainte est implémentée dans le système par le fait que les relations entre les unités formées des règles de réécriture, hiérarchiques par excellence, ont des poids différents : par exemple, l'activation du SN par P est plus rapide que celle du SV, ou encore l'activation du Déterminant par SN précède celle du Nom. Le poids des connexions garantit donc que la sélection lexicale se fait en respectant les règles pesant sur l'organisation linéaire de la phrase (i.e. la branche de gauche de l'arbre syntaxique précède celle de droite dans la linéarisation).

Pour étayer ces hypothèses, Stemberger a lui aussi recours à l'analyse des lapsus. Nous ne reviendrons pas ici sur les données montrant un contrôle lexical sur le traitement syntaxique, puisqu'il s'agit d'une hypothèse partagée par les modèles modulaires (même si certains auteurs comme Garrett considèrent toutefois un flux minimal d'informations lexicales vers la composante syntaxique). L'influence syntaxique sur la sélection lexicale est quant à elle étayée par le fait que les erreurs de substitution lexicale décrites respectent certaines contraintes syntaxiques. Par exemple, elles préservent la plupart du temps la catégorie grammaticale des mots. Le fait que la catégorie grammaticale des mots soit préservée lors des substitutions indique, selon Stemberger, que la syntaxe exerce

une contrainte sur la sélection lexicale. Notons que si l'on considère l'information relative à la catégorie grammaticale comme une propriété lexico-syntaxique de l'unité lexicale, présente dans le lexique, cette observation ne constitue pas un argument en faveur d'une contrainte de la composante proprement syntaxique sur le fonctionnement lexical.

Dans le modèle interactif illustré par la Figure 3, le système syntaxique opère sous le contrôle de la structure sémantique et pragmatique du message. Cette relation n'est cependant pas discutée par Stemberger. Une série d'études sur l'accord procurent des données expérimentales compatibles avec l'hypothèse d'un contrôle conceptuel sur le traitement syntaxique ; elles sont décrites plus loin dans cette introduction et font l'objet des Chapitres 2 et 4.

Qu'en est-il enfin des relations entre la composante syntaxique et la composante phonologique ? Aucune flèche ne relie directement ces deux composantes dans le modèle de Stemberger, il ne discute par ailleurs pas cette relation. Contrairement aux observations de Garrett, l'analyse des erreurs flexionnelles effectuée par Stemberger indique que seule une très faible proportion des erreurs flexionnelles sont phonologiques. Par conséquent, plutôt que de placer le traitement des morphèmes liés au stade de l'encodage phonologique, comme le propose Garrett, Stemberger suggère qu'il est assuré conjointement par les composantes lexicale et syntaxique. Le point de départ de son analyse du traitement morphologique vient des erreurs qui consistent à permuter deux mots tout en maintenant en place leurs flexions (*stranding errors*, voir Ex. 3). Stemberger remarque que ces erreurs ne respectent pas la seule contrainte liée à la catégorie grammaticale des mots mais aussi d'autres contraintes de nature syntaxique. Par exemple, en anglais, une racine lexicale portant la flexion *-ing* est couramment échangée avec une racine portant la flexion *-ed* lorsqu'il s'agit du *-ed* au temps parfait (*perfect*), tandis qu'elle n'est jamais échangée avec une racine portant la même flexion *-ed* indiquant le temps passé (*past*). Il se trouve justement que les contraintes d'utilisation de la forme *-ing* sont similaires à celles du parfait (toutes deux se forment avec un auxiliaire ainsi qu'avec la forme *be*), mais qu'elles diffèrent par contre fortement de celles du passé. Par ailleurs, les analyses montrent que le *-s* final du pluriel constitue une exception à la règle du *stranding* : la racine se déplace souvent avec le *-s* au lieu de le laisser dans son slot. Ce type d'erreur est interprété par Garrett comme témoin qu'elle a eu lieu au stade fonctionnel parce que le pluriel n'a pas été spécifié sur la bonne unité. Stemberger suggère quant à lui de l'expliquer par le fait que l'utilisation de la flexion *-s* du pluriel est très peu contrainte sur le plan syntaxique (le pluriel de la plupart des noms se forme en effet par l'ajout de cette flexion en anglais), et que par conséquent le mouvement du *-s* ne fait pas l'objet d'un contrôle syntaxique important.

De ces différentes observations, Stemberger suggère deux points. Premièrement, les flexions et les racines lexicales ont des représentations lexicales

distinctes, puisque les erreurs spontanées indiquent qu'elles peuvent circuler séparément dans la phrase. Stemberger postule également sur la base de son corpus que les formes fléchies des unités lexicales sont représentées séparément des formes non fléchies dans le lexique ; ce dernier contient donc des unités pour les formes non fléchies, des unités pour les formes fléchies et des unités pour les seules flexions (ex. *-ing*, *-s*, *-ed*). Ces dernières sont représentées une seule fois dans le lexique, il s'agit de représentations partagées entre les différentes unités lexicales qui peuvent les porter. Deuxièmement, l'activation de ces unités flexionnelles est contrainte par les règles de la syntaxe. En résumé, la morphologie est traitée en parallèle par la composante lexicale et par la composante syntaxique qui s'assure que la sélection des morphèmes flexionnels respecte les règles de la syntaxe. La composante phonologique n'intervient donc pas directement dans le traitement syntaxique ; elle peut toutefois théoriquement exercer un effet indirect sur cette dernière via les rétroactions sur la sélection lexicale et l'interaction entre les composantes lexicale et syntaxique. Dans le Chapitre 5, nous discutons l'implication théorique, sur l'interprétation des relations entre morphologie et syntaxe, du statut qui est réservé au traitement de la morphologie flexionnelle dans les modèles de Garrett et Stemberger.

Assez peu de recherches expérimentales ont été conduites pour tester la perspective interactive du traitement syntaxique en production de phrases. Par ailleurs, le modèle porte essentiellement sur les composantes lexicale et syntaxique de la production et leurs relations. Les aspects de la composante syntaxique examinés se limitent en plus à la 'syntaxe lexicale' ; le modèle n'aborde pas la question des opérations syntaxiques non lexicales, i.e. combinatoires, comme l'accord, ni par conséquent celle de leurs relations avec les différents niveaux de la production. Les prédictions sur le plan empirique sont par conséquent fort vagues concernant ces opérations. Il nous semble toutefois que c'est justement au niveau de ces opérations-là que la question de la modularité de la syntaxe semble la plus pertinente, puisqu'elles requièrent l'intervention de procédures indépendantes des unités lexicales, et donc proprement syntaxiques. Dans les paragraphes qui suivent, nous décrivons brièvement le modèle de compétition développé par Bates et McWhinney (1989). Il adopte les mêmes principes de fonctionnement que celui de Stemberger, à savoir l'activation en parallèle et en interaction des différents niveaux de traitement. Bien que ce soit surtout dans le domaine de la compréhension qu'il s'est le plus illustré, le modèle de compétition avance un cadre théorique que nous serons amenée à considérer dans plusieurs chapitres.

Le modèle de compétition vise à rendre compte de la compréhension mais aussi de la production du langage dans une perspective issue du fonctionnalisme linguistique, selon laquelle les formes linguistiques sont créées, gouvernées, contraintes, acquises et utilisées au service de la fonction communicative (Bates & MacWhinney, 1989). Le système linguistique n'est pas conçu comme isolé des aspects généraux de la cognition humaine ; non seulement ses propriétés de

fonctionnement dérivent de celles de l'esprit humain, mais aussi il opère en interrelation étroite avec les différentes sources d'information.

Contrairement aux deux modèles présentés jusqu'ici, le modèle de compétition ne spécifie que deux niveaux structuraux : un niveau formel et un niveau fonctionnel. Au niveau formel sont représentées les formes de surface, également appelées indices linguistiques. Elles se composent des unités lexicales, des marques morphologiques, de la spécification de la position dans la phrase ou des contours prosodiques, etc. Le niveau fonctionnel contient les représentations des aspects sémantiques et pragmatiques qui sous-tendent les formes. Formes et fonctions ne sont pas en relation biunivoque ; une même forme peut exprimer plusieurs fonctions tandis qu'une même fonction peut recevoir plusieurs formes linguistiques. Par exemple, les unités telles que le rôle d'agent ou de thème dans la structure fonctionnelle de la phrase sont associées à différentes formes linguistiques (connexions verticales). Ces deux fonctions sont également liées entre elles, et les différentes formes entretiennent aussi des relations (connexions horizontales).

L'approche quantitative statistique de la langue est au centre du modèle de compétition. L'hypothèse est que les connexions peuvent être plus ou moins fortes en fonction de la corrélation ou cooccurrence entre les deux unités dans l'usage de la langue. Il est possible de calculer ces corrélations, et cette estimation quantitative reflète le poids des différentes représentations du système tout en permettant de rendre compte de son fonctionnement. Ces valeurs sont progressivement calculées par le système au cours de l'acquisition de la langue et varient d'une langue à l'autre. Une des caractéristiques du modèle, et qui le distingue de l'ensemble des autres modèles, est par conséquent sa capacité à rendre compte des différences entre les langues. Les langues se différencient sur le plan qualitatif quant aux indices qu'elles ont mis au point pour guider le langage, elles se différencient également au niveau quantitatif quant à la pertinence relative de chaque indice pour guider le traitement. Prenons par exemple l'association entre la position dans la phrase et le rôle syntaxique de sujet. En anglais, cette association est extrêmement forte : l'ordre des mots est globalement limité à l'ordre dit canonique, i.e. Sujet-Verbe-Objet ; une fois le premier mot sélectionné, le rôle de sujet lui est automatiquement attribué et conditionne aussi le choix du verbe et de l'objet. L'italien, par contre, permet une grande variabilité dans l'ordre de ces trois constituants. Un locuteur peut donc entamer une phrase sans encore avoir spécifié le rôle syntaxique du premier syntagme nominal, et donc sans s'engager quant à la structure qui suivra. Il semble en effet que les locuteurs italiens mis en situation de conversation accélérée produisent beaucoup plus de structures de type OVS ou OSV que sous des conditions d'élocution standard (Antinucci, Duranti, & Gebert, 1979). Autrement dit, la sélection d'une unité lexicale pour une certaine position dans la phrase est un indice extrêmement contraignant pour déterminer le rôle

syntactique de cet élément en anglais, tandis qu'il n'est qu'un piètre indice en italien.

Les indices guident les processus de compréhension et de production ; le modèle suppose que des représentations identiques sous-tendent ces deux processus. En compréhension, les indices sont utilisés pour accéder au sens du message, tandis qu'en production ils servent à en créer la forme. Les indices sont caractérisés par deux propriétés : leur validité et leur coût. La *validité* d'un indice par rapport à une fonction (ou une forme) dépend de la fréquence de son association avec cette fonction. Celle-ci détermine sa disponibilité (le nombre de fois que cet indice est présent lorsque la fonction est activée) et sa fiabilité (le nombre de fois que cet indice est associé à cette fonction par rapport au nombre de fois où il est associé à une autre). Etant donné les interconnexions du réseau, plusieurs indices sont souvent activés, et peuvent soit concorder soit être en discordance. Dans ce cas, une compétition s'installe entre les indices, elle sera remportée par celui ayant la plus grande validité. Un indice peut toutefois posséder une grande validité mais perdre la compétition au profit d'un indice ayant un moindre coût. Le *coût* d'un indice dépend de sa perceptibilité⁵ (ex. la marque du pluriel n'est pas audible sur la troisième personne du pluriel en français) et de sa disponibilité, i.e. la possibilité de l'utiliser dès son apparition, sans devoir le stocker en mémoire de travail pour pouvoir bénéficier de l'information dont il est porteur.

C'est surtout pour expliquer les processus de compréhension que la puissance du modèle de compétition apparaît. Malheureusement, les recherches visant à tester le modèle quant aux processus syntaxiques en production sont encore peu nombreuses. Bates et Devescovi ont toutefois mis en évidence l'effet, sur la production de structures syntaxiques complexes (formes), telles que les relatives et les passives, d'une série de variables pragmatiques et sémantiques (fonctions) contextuelles comme le contraste entre information nouvelle ou ancienne, le thème du discours ou encore le maintien de la perspective. Les conclusions théoriques proposées par Bates et Devescovi dans le cadre du modèle de compétition restent néanmoins relativement superficielles ; les principes d'engagement et d'accessibilité sont mis en avant sans être réellement développés. Par ailleurs, les relations entre les formes syntaxiques et phonologiques ne sont pas abordées à notre connaissance.

L'approche quantitative des relations entre les différents niveaux de représentation linguistique et non linguistique est également au centre des théories visant à expliquer le traitement des ambiguïtés syntaxiques en compréhension de phrases (ex. Trueswell, Tanenhaus & Garnsey, 1994 ; MacDonald, Pearlmutter &

⁵ La notion de perceptibilité est spécifique des processus de compréhension ; elle repose sur les caractéristiques phonologiques de l'indice. On peut toutefois faire l'hypothèse que la perceptibilité morphophonologique d'un indice influence aussi son utilisation en production, comme il est discuté dans le chapitre 5.

Seidenberg, 1994). Ces théories, aussi appelées modèles de satisfaction des contraintes ou modèles probabilistes, suggèrent un fonctionnement hautement interactif du système de décodage syntaxique déterminé par la fréquence des unités lexicales et des structures syntaxiques dans la langue. Nous ne développerons pas ici ces modèles qui ont été conçus pour rendre compte des processus de compréhension, et particulièrement du traitement de variables lexicales ; les contraintes de fonctionnement qu'ils proposent, ainsi que leurs implications pour l'interprétation de certaines de nos données, sont discutées plus loin.

Deux conceptions fondamentalement opposées de la composante syntaxique en production de phrases ont été exposées. Selon la première, la syntaxe est une composante informationnellement cloisonnée, fonctionnant sur la base d'un input strictement limité aux représentations finales que lui fournit le système de sélection lexicale situé en amont dans la chaîne séquentielle de production. Un certain flou subsiste quant à la nature exacte des informations lexicales pouvant pénétrer le système syntaxique. Les tenants d'une perspective modulaire radicale soutiennent ainsi que le système n'est perméable qu'à certaines informations lexicales, à savoir les informations syntaxiques véhiculées par les lemmes. Une telle perspective radicale nécessite une décomposition du système de sélection lexicale en deux étapes séquentielles, la première responsable de la récupération des propriétés lexico-sémantiques, la seconde de la récupération des propriétés lexico-syntaxiques. Ces dernières constitueraient les représentations finales du module de sélection lexicale et elles seules serviraient d'input aux traitements syntaxiques.

Selon la seconde conception, le traitement syntaxique est ouvert à l'ensemble des informations potentiellement utiles à son fonctionnement. Chacune de ces informations influence le processus syntaxique en fonction de critères tels que son accessibilité, sa validité, la fréquence de son occurrence dans la langue (son seuil d'activation au repos), la force de la corrélation qu'elle entretient avec la forme syntaxique en question ou encore la présence d'informations concurrentes. Toutes ces informations pénètrent la composante syntaxique en parallèle et entrent en compétition pour assurer son contrôle.

Il est important de noter que les deux conceptions convergent pour reconnaître une certaine *perméabilité représentationnelle*, en ce sens que toutes deux partagent l'idée que la production implique des représentations de natures différentes (lexicale, syntaxique, phonologique, prosodique) ayant chacune leur spécificité. Ces représentations sont construites progressivement au cours de l'apprentissage et régulièrement remises à jour avec l'expérience ; leur élaboration doit par conséquent être sensible au système cognitif général. Les divergences entre ces deux conceptions concernent ce que l'on pourrait appeler la *perméabilité*

opérationnelle des processus, à savoir la façon dont les systèmes de traitement comportant ces représentations opèrent on-line, en situation de production.

Nous l'avons vu, une des principales motivations de ces modèles réside dans leur capacité à expliquer la survenue de certains patterns d'erreurs spontanées. Toutefois, leur objectif final se doit avant tout d'être capable de rendre compte des caractéristiques de la production normale. L'une des propriétés principales de la production est la nécessité de réguler le flux d'information à chacun des niveaux de traitement tout en assurant un *flux continu* de parole. A tout moment, lorsque nous parlons, nous devons à la fois garder à l'esprit l'idée que nous voulons exprimer, trouver les mots justes, la syntaxe adéquate, prononcer l'ensemble, tout en continuant à planifier le reste de la phrase. Sans cette continuité dans le processus de production, les énoncés seraient parcourus de silences ; le système articulatoire attendant de recevoir les nouvelles instructions pendant qu'un nouveau paquet d'idées est traité. Chaque locuteur est donc capable de procéder à un *traitement simultané* des différents aspects de la phrase. La simultanéité des traitements est un présumé des modèles interactifs, mais comment une organisation modulaire de la production peut-elle assurer la fluidité malgré le cloisonnement et la séquentialité caractérisant le fonctionnement des différents systèmes de traitement ?

Une solution potentielle à ce problème de la fluidité consiste à diviser l'énoncé en fenêtres ou unités de production de plus en plus étroites à chaque niveau, suggérant ainsi une hiérarchie de traitement (Bock, 1991). Les unités d'encodage grammatical sont conçues comme plus grandes que les unités d'encodage phonologique, lesquelles à leur tour dépassent en taille les unités de l'articulateur. L'observation des erreurs produites en langage spontané ou sous condition expérimentale montre, en effet, que les erreurs d'échanges de mots entiers se situent en général au sein d'une même proposition, tandis que les échanges de phonèmes surviennent au sein de l'unité inférieure qu'est le syntagme. Au niveau de l'encodage du programme moteur, la fenêtre de traitement semble encore plus petite puisque les unités seraient de l'ordre du groupe d'accentuation (Bock, 1991). Une telle optique incorpore la notion de traitement parallèle dans le modèle modulaire ; elle permet de concilier un fonctionnement strictement séquentiel et cloisonné avec un traitement en parallèle des différentes unités à chaque niveau de production.

A cette propriété qu'est le flux continu, s'ajoute celle de la *rapidité* de la production. En moyenne, nous produisons deux à trois mots par seconde que nous allons récupérer dans un gigantesque réservoir d'environ trente mille mots utilisés couramment (Levelt, 1989). Cette avalanche de mots, nous les sélectionnons avec une étonnante précision et les organisons dans des phrases suivant des règles extrêmement strictes. Pourtant, on ne dénombre en moyenne qu'une seule erreur tous les mille mots produits (Bock, 1991). Si l'on ne considère que les erreurs liées

à la structure syntaxique unissant les mots dans la phrase, le taux de productions incorrectes tombe même à moins d'un demi pour cent des phrases (Deese, 1984). La production est donc foncièrement *correcte*. Les deux conceptions, modulaire et interactive, rendent compte différemment de ces propriétés du système. Le système modulaire assure une certaine rapidité de fonctionnement et limite les risques d'interférences entre les différentes sources d'informations, tandis que le système interactif assure une production correcte en cas de perte d'information grâce à la redondance introduite par les différentes sources. Cette analyse 'fonctionnelle' des deux modèles sera considérée à plusieurs reprises dans les discussions des travaux que nous avons menés.

1.2 L'accord grammatical : input minimal ou input maximal ?

Les travaux expérimentaux sur l'accord grammatical constituent une source de données capitale quant à la question de la modularité des processus syntaxiques opérés au stade de l'encodage grammatical. Certains de ces travaux semblent étayer la perspective modulaire du traitement syntaxique, tandis que d'autres semblent ne pouvoir s'expliquer que par une vision interactive des relations entre la syntaxe et les composantes conceptuelle et phonologique. La principale raison de ces difficultés provient du fait qu'aucun des modèles de production existants n'avance d'hypothèse suffisamment claire concernant le statut de l'accord. Dans cette section, nous passons en revue les données empiriques présentées dans la littérature⁶. Les implications théoriques de ces résultats ne sont toutefois pas abordées dans les détails ; elles font l'objet de discussions au sein des chapitres 2 à 7 et un cadre d'interprétation des différentes contraintes pesant sur l'accord est proposé dans la Synthèse (Chapitre 8).

Par ailleurs, plutôt que de reprendre les termes de fonctionnement modulaire vs. interactif du système, nous avons choisi d'adopter une terminologie différente. Nous avons proposé de parler d'input minimal versus input maximal (Vigliocco & Franck, 1999) pour se référer aux deux perspectives théoriques que les études sur l'accord confrontent. La raison principale est que nous pensons que le débat entre modularité et interaction est trop large pour encadrer les données actuelles : la question principale posée par ces recherches n'est pas « Peut-on envisager que les mécanismes syntaxiques constituent un module de traitement, avec toutes les propriétés que cela implique au sens de Fodor ? », mais plus modestement « Quelle est la nature des informations qui contrôlent un mécanisme syntaxique typique comme le processus d'accord ? ». Bien sur, des hypothèses sur la modularité peuvent être formulées sur la base des résultats de ces recherches, mais elles sont insuffisantes pour tirer dès à présent des conclusions univoques. Nous reviendrons

⁶ Cette revue de littérature est partiellement reprise de Franck & Hupet (en révision).

dans la Synthèse théorique sur l'implication théorique de ces recherches quant à la question de la modularité du traitement syntaxique.

1.2.1. Les recherches en faveur de l'hypothèse d'un input minimal

Cette première hypothèse, ancrée dans la perspective modulaire, suppose qu'une quantité minimale d'information pénètre le niveau de traitement syntaxique, à savoir l'information strictement nécessaire. Dans cette optique, seule l'information lexicale de sous-catégorisation, i.e. l'information strictement nécessaire, peut contrôler la réalisation de l'accord. Une fois la sélection des lemmes opérée, la composante conceptuelle de la phrase n'a plus aucune prise sur les mécanismes syntaxiques ultérieurs (mise à part l'assignation des rôles syntaxiques). La signification du nombre ou du genre est encodée lors de la sélection lexicale sous forme d'un indicateur abstrait qui, à lui seul, assure le contrôle de l'accord de l'élément cible. La marque de la cible d'un accord (ex. le verbe ou l'adjectif) est conçue comme purement formelle, elle ne porte pas de réelle signification contrairement à la marque de la source (ex. le nom avec lequel le verbe ou l'adjectif s'accorde). Autrement dit, l'opposition entre les formes nominales *chanson* et *chansons* n'est pas de même nature que l'opposition entre les formes verbales *chante* et *chantent* : alors que la première est à la fois de nature conceptuelle et syntaxique, la seconde est de nature strictement syntaxique. Etant de nature purement syntaxique, la marque portée par la cible de l'accord est sélectionnée sur la base de la marque syntaxique de la source.

L'hypothèse de l'input minimal place également une restriction quant à l'impact de l'information morphophonologique sur l'accord. Les corrélats d'ordre morphophonologique de nombre (ex. le *-s* pluriel sur le nom) ou de genre (ex. la terminaison féminine *-euse* sur le nom), spécifiés au stade de l'encodage phonologique, sont eux aussi ignorés puisqu'ils ne constituent pas une information strictement nécessaire. Le statut de l'accord tel qu'il est envisagé par l'hypothèse minimaliste est illustré par la Figure 4.

La plupart des travaux expérimentaux menés sur l'accord reposent sur l'utilisation du paradigme expérimental proposé à l'origine par Bock & Miller (1991). Ce paradigme, dit de complètement de phrase, consiste à présenter aux participants un début de phrase formé d'un nom sujet et d'un nom dit 'local' (ex. *La porte des bureaux*) et de leur demander de compléter ce préambule en construisant une phrase complète, comportant un verbe et éventuellement un complément (ex. *La porte des bureaux est ouverte*). Les productions des participants sont enregistrées et les erreurs d'accord constituent la variable dépendante. Le principe de proximité, selon lequel la présence d'un nom local, proche du nom sujet et de nombre différent de celui-ci, est une source d'erreurs d'accord (voir Prélude), constitue la pierre angulaire de l'ensemble de ces travaux.

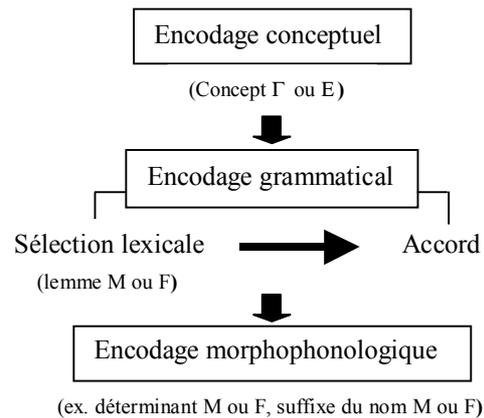


Figure 4. L'accord selon l'hypothèse minimaliste.

L'introduction d'un nom local dans la phrase est utilisée soit pour augmenter le taux d'erreurs d'accord, soit pour manipuler certaines propriétés relatives à la distributivité du sujet (voir paragraphe suivant), ou encore pour étudier l'impact de certaines propriétés conceptuelles ou morphophonologiques du nom local lui-même sur la réalisation de l'accord. L'hypothèse est que si le traitement syntaxique est perméable aux informations conceptuelles et/ou morphophonologiques, un nom local possédant de telles propriétés constituera une plus grande source d'erreurs d'accord qu'un nom local ne possédant pas ces propriétés. Les postulats sous-jacents à cette hypothèse sont discutés dans le Chapitre 4.

Etude de l'input conceptuel

Une première variable étudiée par Bock & Miller (1991) concerne le nombre référentiel du syntagme sujet, et correspond à ce qu'ils ont appelé la distributivité du sujet. Considérons les syntagmes prépositionnels (Ex. 7) et (Ex. 8), repris à Bock & Miller.

Ex. 7 : **The bridge to the islands**
Le pont vers les îles

Ex. 8 : **The label on the bottles**
L'étiquette sur les bouteilles

Le référent dans le cas du syntagme (Ex. 7) est le pont, unique, reliant les différentes îles, tandis que dans le cas de (Ex. 8), l'étiquette se retrouve sur (*est possédée par*) l'ensemble des bouteilles. Le syntagme (Ex. 7) est dit non distribué parce qu'il réfère à une seule entité, le syntagme (Ex. 8) est dit distribué puisqu'il réfère à une entité multiple. Dans les deux cas, le verbe qui suivra devra être au singulier, puisque le syntagme sujet est grammaticalement singulier. Néanmoins, si la référence notionnelle du nombre du sujet peut influencer le processus d'accord, un plus grand nombre d'erreurs devrait être produit lors de l'accord du verbe avec un sujet distribué (ex. **L'étiquette sur les bouteilles sont...*), puisque la référence est plurielle, qu'avec un sujet non distribué (ex. **Le pont vers les îles sont...*),

référant à un seul objet. Bock et Miller n'ont observé aucune différence en termes d'erreurs d'accord entre ces deux types de syntagmes sujets. L'absence d'effet de la distributivité pourrait néanmoins traduire le fait que les locuteurs, en réalité, n'interprètent pas les sujets distribués comme référant à un ensemble d'objets. Or, une épreuve de validation de leur matériel menée sur un groupe de locuteurs anglais a montré que dans 79% des cas, les items distribués sont jugés comme faisant effectivement référence à plusieurs objets, tandis que 18% des items non distribués sont jugés comme y faisant référence. Autrement dit, l'absence d'une différence dans les erreurs d'accord entre les phrases distribuées et celles non distribuées semble bien indiquer que les locuteurs ne prennent pas en considération la référence conceptuelle non distribuée de la quantité, mais bien le nombre grammatical qui la reflète pour construire l'accord du verbe.

Vigliocco, Butterworth et Garrett (1996) ont répliqué ce résultat auprès de locuteurs anglais. Le même matériel que celui présenté dans cette étude a ensuite été soumis à une variante du paradigme de complètement de phrases : à partir des préambules, les participants devaient former des phrases interrogatives. Un adjectif était présenté avant le préambule, afin de contraindre les productions des participants : ils voyaient par exemple *Threatening* (menaçant/e), suivi de *The letter from the lawyers* (La lettre des juristes) et devaient produire la question *Is the letter from the lawyers threatening?* (La lettre des juristes est-elle menaçante?). L'objectif de cette manipulation était de tester dans quelle mesure le fait de devoir produire le verbe, et sa marque de nombre, avant le sujet inciterait les locuteurs anglais à recourir à la représentation conceptuelle du syntagme sujet, celui-ci n'étant pas encore spécifié lors de l'encodage grammatical du verbe. Les résultats se sont avérés négatifs : malgré l'inversion de l'ordre des mots, aucun effet de distributivité n'a été observé dans cette tâche.

Bock et Miller (1991) ont également testé dans quelle mesure des caractéristiques conceptuelles typiques des sujets peuvent influencer l'accord sujet-verbe. Deux propriétés des noms sujets ont été manipulées, reconnues comme prototypiques des sujets (Bates & MacWhinney, 1982) : le caractère animé/inanimé et le caractère concret/abstrait du nom. Des préambules similaires à ceux présentés dans l'expérience précédente ont été créés, comportant un nom sujet et un nom local, chacun pouvant varier sur les critères d'animation et de concrétude. Dans la recherche décrite ici, on manipulait des caractéristiques sémantiques du sujet qui ne concernent pas directement le nombre du syntagme mais la propension du nom à remplir la fonction de sujet du verbe. Contrairement à la distributivité, ces caractéristiques ne relèvent pas d'une interprétation liée à la combinaison du nom sujet et du nom local mais bien des caractéristiques sémantiques des noms. L'hypothèse minimaliste prédit que ces caractéristiques, liées à la sémantique du mot, n'influenceront pas la relation d'accord qui a lieu après l'assignation des fonctions, bien qu'elles puissent jouer un rôle dans la détermination des fonctions syntaxiques. Autrement dit, le verbe s'accorde avec le

sujet de la phrase parce qu'il occupe la position sujet, et non parce qu'il possède telle ou telle caractéristique sémantique généralement associée aux sujets. Bock et Miller ne rapportent d'effet significatif ni de l'animation ni de la concrétude du nom local. Ces auteurs reconnaissent néanmoins que leur manipulation ne porte pas directement sur l'impact de facteurs conceptuels sur les processus d'accord, mais plutôt sur l'influence de ces facteurs sur la détermination des fonctions syntaxiques. Cette influence a une conséquence inévitable sur la relation d'accord puisque celle-ci repose sur l'identification du segment remplissant la fonction sujet dans la phrase. Il est d'ailleurs étonnant qu'aucun effet de ces variables n'ait été décelé, puisque les auteurs reconnaissent eux-mêmes l'irréfutable lien entre ces variables sémantiques, i.e. l'animation et la concrétude, et l'assignation de la fonction sujet. Pour sortir de cette impasse, Bock et Miller suggèrent que l'absence d'un effet de ces variables vient du fait que l'assignation des fonctions syntaxiques dans les phrases présentées était trop simple pour susciter un effet (effet plancher). Effectivement, les auteurs ont montré qu'en utilisant des préambules au sein desquels l'assignation des fonctions syntaxiques est complexe (ex. *The politician that the flag*, Le politicien que le drapeau), il apparaît un effet important de l'animation du nom local : un plus grand nombre d'erreurs sont produites quand il est animé que lorsqu'il est inanimé. Autrement dit, cette recherche de Bock et Miller, bien que souvent avancée (ex. Bock et Eberhard, 1993) comme étant en faveur de l'isolation du traitement syntaxique, ne permet pas de départager l'hypothèse minimaliste de l'hypothèse maximaliste, puisqu'elle ne porte pas sur des caractéristiques conceptuelles susceptibles d'influencer la relation d'accord, mais bien sur des facteurs sensés influencer l'assignation des fonctions syntaxiques.

Une autre caractéristique sémantique, intrinsèque à l'unité lexicale, mais directement liée à la notion de nombre concerne la propriété qu'ont certains noms de référer à une entité multiple. Ces noms, dits collectifs, sont grammaticalement singuliers mais désignent un ensemble, une collection d'êtres ou d'objets (Briet, 1996), tels que une équipe, un essaim, etc. Les deux interprétations (singulier ou pluriel) coexistent selon que la notion de globalité (" penser au tout ") ou celle de distribution (" penser aux parties ") est privilégiée. En utilisant le même paradigme de complètement de phrases, Bock et Eberhard (1993) ont étudié l'impact de la présence d'un nom local collectif sur l'accord du verbe chez des locuteurs américains. L'hypothèse maximaliste prédit qu'avec des sujets singuliers, une plus grande proportion d'erreurs d'attraction sera produite en présence de noms locaux singuliers collectifs (ex. **The condition of the fleet ARE...*) qu'en présence de noms locaux singuliers individuels (ex. **The condition of the ship ARE...*). Autrement dit, il est attendu qu'un nom local collectif entraîne un effet d'attraction puisque sa pluralité notionnelle est supposée être prise en compte, et cela même si l'information grammaticale qu'il porte est le singulier. Bock et Eberhard ne rapportent aucun effet d'attraction en présence de noms locaux collectifs

grammaticalement singuliers, et par conséquent aucune différence dans cet effet par rapport aux noms locaux individuels grammaticalement singuliers. Une tâche de jugement présentée aux locuteurs a pourtant confirmé que les collectifs étaient bien interprétés comme référant à une pluralité. Donc, malgré l'interprétation plurielle des collectifs, ils ne génèrent pas d'effet d'attraction sur l'accord du verbe. Néanmoins, les auteurs rapportent que la présence d'un nom local collectif grammaticalement pluriel tend à entraîner des erreurs d'attraction en plus grand nombre que celle d'un nom local individuel pluriel, ce qui semble indiquer que l'information conceptuelle peut perturber la relation d'accord, pour autant qu'elle soit combinée à une information grammaticale interférente.

Etude de l'input morphophonologique

Le rôle des variables morphophonologiques dans la réalisation de l'accord a reçu nettement moins d'attention que celui des facteurs conceptuels. Deux facteurs ont été étudiés par Bock et Eberhard (1993), il s'agit des pseudo pluriels et des pluriels irréguliers. Dans une première recherche, les auteurs ont manipulé la présence d'une marque morphophonologique de pluriel sur le nom local. En anglais, la forme régulière du pluriel du nom est représentée par le segment phonologique final [s] ou [z]. Trois conditions expérimentales ont été construites selon que le nom local était un pseudo pluriel (ex. *The player on the course*), un vrai singulier (ex. *The player on the court*) ou un vrai pluriel (ex. *The player on the courts*). Aucune erreur d'accord n'a été observée dans la condition pseudo pluriel, les erreurs ne surviennent que dans la condition vrai pluriel. Ce résultat indique que ce ne sont pas les corrélats phonologiques du nombre du nom local qui déclenchent l'effet d'attraction observé pour les vrais pluriels, mais bien l'information grammaticale qu'ils véhiculent.

Il reste à voir cependant si l'information grammaticale est à elle seule responsable du phénomène d'attraction ou si la présence d'information phonologique sur le nom local, concordant avec l'information grammaticale, peut augmenter l'apparition d'erreurs d'attraction. Autrement dit, il s'agit de déterminer dans quelle mesure la présence de régularités morphophonologiques liées au pluriel, présentes sur le nom local, contribue à l'effet d'attraction. C'est ce que Bock et Eberhard ont testé en comparant l'effet d'attraction avec des noms locaux dont le pluriel ne se marque pas régulièrement, c'est-à-dire dont la flexion pluriel [s] ou [z] est absente (ex. *The trap for the mice*, Le piège pour les souris), à des noms locaux pluriels régulièrement marqués (ex. *The trap for the rats*, Le piège pour les rats). Les résultats indiquent une absence d'effet de la régularité morphophonologique du pluriel sur l'effet d'attraction : les locuteurs n'ont pas commis plus d'erreurs d'accord lorsque le nom local était régulièrement marqué pour le pluriel que lorsqu'il possédait une forme irrégulière.

Meyer et Bock (1998) ont, quant à elles, étudié l'impact de facteurs morphophonologiques sur le processus d'accord en genre du pronom anaphorique

chez des locuteurs néerlandais. Des phrases contenant deux noms de genre identique ou différent étaient présentées (ex. *Kijk, daar ligt de/een aardappel bij het/een badpak*, Regarde, il y a une patate à côté du maillot de bain), chaque nom pouvant être précédé d'un déterminant défini (*de/het*), qui est marqué pour le genre, ou indéfini (*een*), qui est non marqué. La présentation de la phrase était suivie de la présentation d'un adjectif pouvant soit qualifier le premier nom (dans l'exemple cité ci-dessus, *gaar*, cuit/e), soit le second. La tâche consistait à produire une nouvelle phrase, en continuation avec la précédente, intégrant l'adjectif et commençant par un pronom anaphorique (Expérience 1, *Die/dat is gaar*, Elle/il est cuit/e). L'accord en genre entre le pronom et le référent constituait la variable dépendante de l'étude. Les résultats de cette première expérience vont dans le sens de l'hypothèse de l'isolation du processus d'accord : l'effet d'attraction du nom local n'est pas plus important lorsque le déterminant qui le précède est marqué que lorsqu'il est non marqué. Néanmoins, la relation d'accord étudiée ici concerne une source et une cible situées dans des phrases différentes, ce qui rend difficile la comparaison avec les recherches précédentes portant toutes sur une relation d'accord entre deux éléments au sein de la phrase.

Les travaux présentés jusqu'ici, plaidant en faveur d'un module syntaxique isolé des autres aspects du langage et de la cognition, ont presque tous été menés en anglais. L'anglais se distingue toutefois de bon nombre d'autres langues sur des propriétés structurales pouvant justement jouer un rôle prépondérant dans la régulation du flux d'information entre le niveau de formulation syntaxique et les autres niveaux de production. La première de ces propriétés est la pauvreté du système flexionnel anglais. En ce qui concerne le nombre, seuls le sujet et le verbe sont porteurs d'une information de nombre dans la phrase, aucune autre unité linguistique n'étant marquée pour le nombre en anglais ; par conséquent, l'accord en nombre se limite à l'accord sujet-verbe. Cependant, dans la plupart des cas, la marque de nombre n'apparaît pas sur le verbe. En effet, une telle marque ne se retrouve que dans la forme au présent de la troisième personne du singulier et dans quelques conjugaisons du verbe 'être' (*to be*). Outre la relative rareté des flexions de nombre, leur marque morphophonologique, nominale ou verbale, est également extrêmement pauvre. L'opposition singulier/pluriel se marque par la présence ou l'absence du /s/ final ; la présence du /s/ sur le nom indique la pluralité tandis qu'elle indique la singularité sur le verbe (à la troisième personne au présent uniquement). En ce qui concerne les oppositions de genre dans la langue anglaise, elles sont quant à elles fort limitées puisque non seulement, l'anglais ne comporte pas de genre strictement grammatical (ex. *le fauteuil, la chaise*), mais en plus le genre conceptuel n'est morphologiquement marqué que sur certains pronoms. Par conséquent, l'accord en genre se limite à un accord entre le pronom et le référent.

La seconde spécificité de l'anglais est, comme nous l'avons mentionné, sa rigidité structurale, probablement liée à la pauvreté de son système flexionnel. La pauvreté et la simplicité des systèmes formels d'opposition de nombre et de genre,

ainsi que la rigidité structurale, font sans doute de l'anglais un exemple privilégié pour postuler une relative isolation du système syntaxique. Nous verrons dans la section suivante, ainsi que dans les Chapitres 2, 3 et 5, que les résultats sont effectivement différents dans les autres langues.

1.2.2. Les recherches en faveur de l'hypothèse d'un input maximal

L'hypothèse d'un input maximal trouve ses sources dans les approches interactionnistes basées sur le principe de diffusion de l'activation entre les niveaux de traitement (ex. Bates et MacWhinney, 1989 ; Harley, 1993 ; Stemberger, 1985a). Dans un tel système, la construction de l'accord est réalisée en mobilisant un maximum d'informations disponibles aux différents niveaux de traitement. L'encodeur grammatical puise ses sources non seulement dans l'information lexico-syntaxique, mais également dans l'information fournie par les structures conceptuelle et morphophonologique pour mener à bien l'opération d'accord.

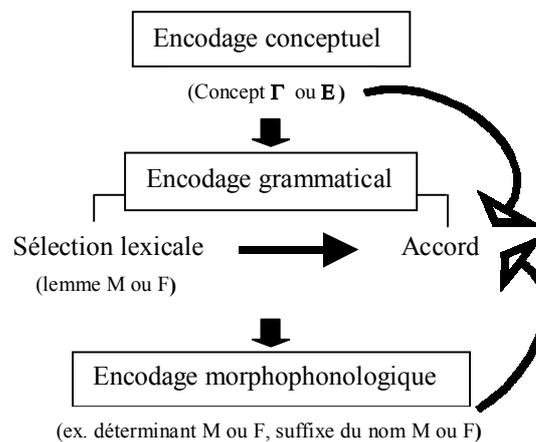


Figure 5. L'accord selon l'hypothèse maximaliste.

Les différentes sources peuvent coïncider, entraînant un renforcement du choix syntaxique, ou discorder, auquel cas elles sont amenées à entrer en compétition, au risque d'entraîner des erreurs d'accord. La Figure 5 illustre l'hypothèse d'un input maximal pour la réalisation de l'accord.

Etude de l'input conceptuel

Nous avons attiré l'attention du lecteur sur les spécificités structurales de l'anglais pouvant expliquer l'absence d'effet des facteurs conceptuels dans les différentes études passées en revue. Néanmoins, il semble cependant qu'il faille distinguer l'accord en nombre entre le sujet et le verbe de celui entre le référent et le pronom. Sur le plan linguistique, les unités impliquées dans ces deux relations d'accord ne sont pas situées de la même façon les unes par rapport aux autres dans la phrase : dans le cas de l'accord sujet-verbe, les deux éléments sont situés dans la même proposition, tandis que le référent et le pronom sont situés dans deux

propositions différentes. Par ailleurs, en situation conversationnelle, le pronom reçoit souvent sa marque de nombre sans qu'il n'ait d'antécédent en référence anaphorique dans une phrase précédente. Dans une telle situation, le nombre du pronom est attribué de façon similaire à celui du nom, à savoir par détermination conceptuelle sur base de l'intention communicative du locuteur. Le pronom doit donc dans certains cas pouvoir être sous contrôle direct de la représentation conceptuelle du message pour recevoir sa marque de nombre, et cela même en anglais. Quelques travaux ont effectivement mis en évidence des différences quant aux processus de production responsables d'assurer l'accord en nombre du pronom et celui du verbe. Il est donc possible que l'accord du verbe soit construit principalement sur base du nombre grammatical du nom, comme les études précédentes l'ont montré, tandis que la marque du pronom serait quant à elle influencée par la représentation conceptuelle du nombre, y compris dans des situations où une telle influence n'est pas réellement nécessaire (i.e. lorsqu'un référent linguistique est présent).

La toute première recherche sur l'utilisation des pronoms par les locuteurs anglophones remonte à Bock, Eberhard & Cutting (1992). La distributivité du sujet (non distribué vs. distribué) ainsi que la représentation notionnelle du nombre du nom sujet (unique vs. collectif) ont été manipulées. Les participants devaient produire soit un verbe, soit un pronom *tag* (ex. *The key to the cabinets vanished, didn't it?* La clé des cabinets a disparu, n'a-t-elle pas?). Contrairement aux résultats antérieurs obtenus pour l'accord du verbe, un effet significatif de la distributivité est rapporté pour la production du pronom, les locuteurs produisant plus d'erreurs (*didn't they?*) lorsque le sujet est distribué que lorsqu'il ne l'est pas. Par ailleurs, les noms sujets collectifs entraînaient plus d'erreurs d'accord du pronom que les noms sujets individuels, et un léger effet allant dans ce sens est également apparu sur la production des verbes.

Dans une recherche similaire menée récemment par Bock, Nicol et Cutting (1999) toujours auprès de sujets anglophones, l'accord du verbe a été étudié en parallèle avec celui de deux sortes de pronoms : le pronom *tag* et le pronom réflexif. Les noms sujets étaient soit collectifs soit individuels. Les résultats indiquent que le verbe et les pronoms se comportent de façon similaire lorsque le sujet est grammaticalement singulier, mais que les pronoms, *tags* et réflexifs, montrent une sensibilité accrue pour la pluralité notionnelle des sujets collectifs. Dans plus de 70% des cas, les locuteurs ont accordé les pronoms au pluriel, c'est-à-dire avec la représentation collective du nom sujet. Un effet de cette variable sur l'accord du verbe est également rapporté, bien qu'inférieur à celui sur l'accord des pronoms : 40% des verbes ont été mis au pluriel suivant un sujet collectif. Les auteurs en concluent que les caractéristiques notionnelles jouent un rôle prépondérant dans l'accord du pronom, mais qu'elles peuvent également influencer, dans une moindre mesure, l'accord du verbe.

Une récente étude menée par Eberhard (1999) a montré qu'il était possible d'obtenir un effet de la distributivité du sujet sur l'accord du verbe en anglais en faisant précéder les préambules par des images les illustrant. Toutefois, dans une seconde expérience, les mêmes préambules qui étaient précédés d'une image ont été présentés sans image, et l'auteur rapporte malgré tout un taux d'erreurs plus élevé dans la condition distribuée. La différence entre ces préambules et ceux utilisés par Bock et Miller (1991), qui eux, n'avaient pas conduit à un effet de distributivité, est que ceux utilisés par Eberhard pouvaient potentiellement être représentés par une image, ce qui n'était pas le cas dans l'étude de Bock et Miller. Autrement dit, un effet conceptuel de la distributivité apparaît sur le verbe en anglais à condition que le matériel soit suffisamment concret et imagé.

Une autre étude, conduite par Vigliocco, Verès & Franck (1997) a également mis en évidence un effet de la distributivité en anglais lorsque le nom sujet était précédé d'un adjectif (ex. *The smiling baby on the blankets*, Le bébé *souriant* sur les couvertures). Nous avons montré que la présence de l'adjectif contribuait à réduire considérablement le taux d'erreurs dans les items non distribués, faisant ainsi apparaître une différence significative avec les items distribués. La présence de l'adjectif pourrait favoriser une lecture singulière du sujet, en insistant sur une particularité de ce nom. Les études sur la distributivité n'ayant pas mis d'effet de cette variable en évidence semblent indiquer que l'absence d'effet est due à l'important taux d'erreurs dans la condition non distribuée, plutôt qu'à un faible taux d'erreurs dans la condition distribuée, ce qui va dans le sens de l'interprétation donnée au rôle de l'adjectif dans la présente expérience.

Ce sont néanmoins essentiellement les travaux menés dans les langues romanes qui ont apporté les données les plus claires à l'appui de la perspective maximaliste. Une série d'études menées par Vigliocco et ses collaborateurs ont porté sur l'effet de distributivité. La première recherche, conduite par Vigliocco, Butterworth et Semenza (1995), portait sur l'italien, caractérisé par deux propriétés structurales intéressantes. Premièrement, le sujet du verbe peut être omis de la phrase (pro-drop), bien que le verbe doive toujours s'accorder en nombre et en personne avec celui-ci ; dans ce cas, un recours à la représentation conceptuelle du sujet est indispensable pour déterminer les caractéristiques de nombre et de personne que le verbe devra porter. Deuxièmement, il est possible, en italien, de placer le sujet après le verbe dans la phrase ; la conjugaison du verbe doit donc être déterminée sur base d'un élément qui n'est pas encore spécifié linguistiquement. Vigliocco et al. (1995) ont mené deux études parallèles variant légèrement quant au dispositif. Dans la première étude, identique à celle de Bock et Miller, les participants devaient compléter librement des préambules composés d'un nom sujet et d'un nom local. Dans la seconde, les préambules étaient précédés d'un adjectif que les participants devaient utiliser pour compléter la phrase. Cette variante a permis d'augmenter les possibilités d'erreurs en contraignant les réponses des sujets. Des syntagmes nominaux non distribués (ex. *La strada verso i*

laghi – Pericolosa/e, La route vers les lacs – Dangereuse) étaient comparés à des syntagmes distribués (ex. *Il menu dei ristoranti – Complicato/i*, Le menu des restaurants – Complicé), en vue de tester l'effet de la distributivité sur l'accord du verbe. Les résultats montrent très clairement que les locuteurs italiens sont sensibles à la représentation conceptuelle plurielle des sujets distribués : près de trois-quarts des erreurs d'accord du verbe dans la condition où le nom sujet est singulier et le nom local pluriel se situaient sur les items dont la représentation était distribuée.

Une recherche similaire a été menée par Vigliocco, Butterworth et Garrett (1996) en espagnol, langue relativement proche de l'italien sur le plan structural puisqu'elle tolère, elle aussi, l'absence de sujet (pro-drop) ou encore la position post-verbale du sujet au sein de la phrase. A nouveau, les deux variantes du paradigme, avec ou sans présentation de l'adjectif, ont été testées. Des préambules distribués (ex. *La pasta de los libros*, La couverture des livres) étaient comparés à des préambules non distribués (ex. *La autora de las novelas*, L'auteur des nouvelles). Les résultats indiquent que les locuteurs espagnols se basent, eux aussi, sur l'information conceptuelle de nombre des syntagmes sujets, puisque sur la totalité des erreurs d'accord, environ deux-tiers se situaient sur les items à référence multiple.

La distributivité a également été étudiée en néerlandais par Vigliocco, Hartsuiker, Jarema et Kolk (1996). Comme l'italien et l'espagnol, le néerlandais est une langue fort riche sur le plan flexionnel et peu rigide en ce qui concerne l'ordre des mots (sujets post-verbaux), mais qui néanmoins nécessite la présence d'un sujet dans la phrase, contrairement aux deux autres langues. L'objectif de cette étude était de dissocier les facteurs causaux potentiels dans l'effet de distributivité ; l'obtention d'un tel effet en néerlandais viendrait infirmer l'hypothèse qu'il est causé par la possibilité de laisser tomber le sujet de la phrase, cette caractéristique n'étant pas présente en néerlandais. Des items à référence multiple (ex. *De handtekening op de cheques*, La signature sur les chèques) ainsi que des items à référence unique (ex. *De kerk bij de heuvels*, L'église près des collines) ont été présentés aux participants dans le cadre d'une tâche classique de complètement de phrases. Vigliocco et al. rapportent un important effet de distributivité chez les locuteurs néerlandais, 80% des erreurs ayant été produites sur les items référant à une pluralité d'objets. Bien que l'obtention de cet effet élimine la caractéristique 'pro-drop' comme facteur causal, puisque le néerlandais ne la possède pas, ce résultat ne permet pas de départager le rôle des deux autres facteurs, à savoir la richesse flexionnelle et la possibilité d'avoir des sujets post-verbaux.

Pour départager ces deux facteurs, Vigliocco et al. (1996) présentent, dans le même article, une recherche menée sur le français. Dans la majorité des phrases en français, le sujet précède le verbe. L'observation d'un effet de la distributivité dans

cette langue permettrait donc d'infirmar l'hypothèse selon laquelle ce serait la flexibilité de l'ordre des mots qui serait responsable de cet effet dans les autres langues. Les participants devaient compléter par un adjectif des préambules à référence multiple (ex. *L'uniforme des soldats - neuf*) ou unique (ex. *Le témoin des avocats - arrivé*). Les résultats indiquent que les locuteurs francophones tiennent compte de la notion plurielle véhiculée par les sujets distribués, deux-tiers des erreurs d'accord se situant sur ces items.

Enfin, une étude menée en français par Hupet, Fayol et Schelstraete (1998) (voir aussi Hupet, Schelstraete, Demaeght et Fayol, 1996) a mis en évidence un effet, sur l'accord sujet-verbe en production écrite, d'une variable conceptuelle d'une autre nature. Cette variable concerne la plausibilité de la relation sémantique entre le nom local et le verbe. Par une tâche de complétement de phrases, les auteurs ont constaté que les locuteurs produisaient plus d'erreurs d'accord du verbe lorsque le nom local était un sujet plausible de celui-ci (ex. *Dans la cabine remontent des mineurs fourbus*) que quand il ne pouvait pas constituer un sujet du verbe (ex. *Dans la cabine soupirent des mineurs fourbus*). Ce résultat indique que le lien sémantique entre le verbe et le nom local joue un rôle important dans l'apparition de l'effet d'attraction. Il attire également l'attention sur le fait que certaines erreurs d'attraction peuvent provenir non pas d'erreurs dans la réalisation de l'accord proprement dit, mais d'erreurs d'assignation des fonctions syntaxiques (cette dernière étant vraisemblablement rendue plus complexe dans des phrases où le verbe précède le sujet) ; si la fonction sujet est attribuée au nom local précédant immédiatement le verbe, ce dernier sera accordé avec le nombre de ce nom, sans qu'il s'agisse d'une erreur dans la réalisation de l'accord lui-même.

Etude de l'input morphophonologique

Nous avons cité précédemment une étude de Meyer et Bock (1998) dans laquelle les auteurs manipulaient la présence d'information morphophonologique de genre sur le déterminant. Les résultats de cette étude, dans laquelle un pronom anaphorique devait être produit dans une phrase autre que celle contenant son référent, plaidaient en faveur de la non prise en compte de cette information. Cependant, une seconde expérience menée par les mêmes auteurs remet en question cette conclusion. Dans l'Expérience 2, la phrase préambule (ex. *De/een aardappel bij het badpak*, La patate près du maillot de bain) était suivie d'un adjectif (ex. *gaar*, cuite) que les locuteurs devaient y intégrer par une relative (*De/een aardappel die gaar is bij het/een badpak*, La patate qui est cuite près du maillot de bain). Le pronom relatif, marqué pour le genre (*die* vs. *dat*), constituait la variable dépendante. Cette seconde expérience a l'avantage de tester l'impact de la représentation morphophonologique sur l'accord entre deux mots situés dans une même phrase, contrairement à leur première expérience où la source et la cible se situaient dans deux phrases contiguës. Les résultats indiquent cette fois une nette influence de la représentation morphophonologique présente sur l'antécédent (la

source), puisque près de quatre fois plus d'erreurs d'accord en genre du pronom (12,1% contre 3,4%) ont été produites dans la condition non marquée (*een*) que dans la condition marquée (*het/de*). Cependant, l'effet d'attraction rapporté n'est pas plus important en présence d'un nom local marqué que d'un nom local non marqué. Ces deux observations semblent indiquer qu'en néerlandais, l'information morphophonologique présente sur la source d'un accord (le déterminant) peut être prise en compte : combinée à l'information grammaticale, elle favorise la réalisation correcte de cet accord. Cette information morphophonologique n'aurait toutefois qu'un effet limité sur l'accord puisqu'elle apparaît ne pas influencer l'effet d'attraction ; autrement dit, sa présence sur le nom local semble insuffisante pour perturber l'opération d'accord.

Parallèlement à l'étude de la distributivité, Vigliocco, Butterworth et Semenza (1995) ont manipulé en italien la présence d'une marque morphophonologique de nombre sur le nom sujet. L'accord du verbe avec des sujets marqués (ex. *Il-S numero-S*, Le numéro, vs. *I-P numeri-P*, Les numéros) était comparé à l'accord avec des sujets non marqués (ex. *Il-S menu-Ø*, Le menu, vs. *I-P menu-Ø*, Les menus). Les auteurs ont observé une nette augmentation des erreurs d'accord du verbe avec les noms ne portant pas de marque morphophonologique comparativement aux noms marqués, deux tiers des erreurs se situant dans la condition non marquée. Une autre expérience conduite par les mêmes auteurs, manipulant conjointement la présence d'une marque sur le nom sujet et sur le nom local, a mis en évidence l'absence d'un effet morphophonologique du nom local : un nom local marqué n'entraîne pas un effet d'attraction supérieur à un nom local non marqué. Ces résultats vont dans le même sens que ceux obtenus par Meyer et Bock en néerlandais ; les locuteurs semblent utiliser l'information morphophonologique présente sur la source de l'accord lors de sa construction, mais l'information morphophonologique présente sur le nom local n'influence pas l'effet d'attraction.

Hartsuiker, Antón-Méndez, et van Zee (submitted) ont manipulé la présence d'information morphophonologique sur le nom local en néerlandais. Le nom local était constitué d'un pronom objet direct référant à un nom présenté précédemment, avant l'apparition du préambule. Ce nom pouvait être soit animé, auquel cas le pronom y référant était marqué pour le nombre et le cas (*hen*), soit être inanimé, auquel cas le pronom était non marqué pour le nombre et le cas (*ze*). Les erreurs d'attraction du nombre du pronom sur le verbe constituaient la variable dépendante. Globalement, les pronoms ambigus pour le nombre (et le cas) ont entraîné plus d'erreurs d'attraction que les pronoms portant une marque morphophonologique de nombre (et de cas) différente du nom sujet. Les auteurs en concluent que l'information morphophonologique influence la réalisation de l'accord en néerlandais. Cependant, cette recherche ne permet pas de dissocier l'impact de l'information de nombre de celui de l'information relative au cas : il se

pourrait que le pronom non marqué pour la fonction objet ait entraîné plus d'erreurs parce qu'il était un sujet potentiel de la phrase.

Dans une étude ultérieure, Hartsuiker, Schriefers, Bock & Kikstra (in prep.) ont manipulé la présence d'information morphophonologique sur le déterminant du syntagme nominal sujet. Des locuteurs néerlandophones devaient compléter des préambules contenant un nom sujet et un nom local, le nom sujet pouvant être précédé du déterminant de genre neutre (*het*), qui est marqué pour le nombre (*het-S* vs. *de-P*), ou du déterminant de genre commun (*de*) qui, lui, est ambigu quant au nombre (*de-S* vs. *de-P*). La quasi totalité des erreurs d'accord en nombre du verbe se sont concentrées dans la condition où le nombre était non marqué sur le déterminant du nom sujet. Ce résultat vient confirmer l'hypothèse suggérée dans la recherche précédente selon laquelle la représentation morphophonologique du nombre influence la réalisation de l'accord en nombre du verbe avec le sujet.

Citons enfin une étude de corpus oral de locuteurs canadiens français indiquant une forte tendance à la féminisation des noms commençant par une voyelle (Barbaud, Ducharme & Valdois (1982). Les auteurs rapportent que près de 12% des accords effectués dans ce corpus reflètent une assignation incorrecte du genre du nom, la majorité se faisant dans le sens d'une féminisation (ex. *Son âge est la même que lui*). Les noms commençant par une voyelle sont souvent précédés d'un déterminant ambigu pour le genre sur le plan morphophonologique : *l', son, [set]* (cet/cette, à l'oral). Cette observation suggère que les corrélats morphophonologiques des éléments liés au nom dans la structure syntaxique de la phrase influencent la représentation lexicale du genre du nom, et ainsi le processus d'accord grammatical⁷.

Comme l'illustre cette revue de la littérature sur l'accord, une seule recherche expérimentale a été menée sur la production orale de l'accord en français. Le français offre pourtant un terrain d'investigation intéressant compte tenu de la richesse de ses systèmes flexionnels de nombre et de genre. Par exemple, le français comporte une multitude de variations verbales de nombre, allant des alternances vocaliques (ex. [a] / [ō], [ε] / [ō]) à une multitude d'oppositions zéro/consonne (ex. ∅ / [t], ∅ / [s], ∅ / [v], etc.) (Dubois, 1965). En français, le nombre n'est pas seulement marqué sur le nom et le verbe, mais également sur les pronoms, les participes, les adjectifs et les déterminants. Enfin, une difficulté supplémentaire s'ajoute à la complexité du système flexionnel du nombre en français qui est que les oppositions, bien qu'elles apparaissent à l'écrit, ne sont pas toujours audibles, contrairement à l'anglais où une marque écrite est systématiquement représentée à l'oral. Cette spécificité de la langue est non

⁷ Ce point est discuté plus en détails dans le Chapitre 5.

seulement intéressante sur le plan théorique, elle l'est aussi sur le plan méthodologique puisqu'il est possible de manipuler la présence des marques à l'oral. En ce qui concerne le genre, le français possède un genre grammatical et un genre conceptuel, qui portent à la fois sur les substantifs, les pronoms, les participes, les adjectifs et les déterminants. Le genre conceptuel est marqué par des terminaisons morphologiques porteuses d'information, ce qui n'est le cas seulement que d'une partie des noms ayant un genre grammatical⁸.

Dans les chapitres 2 à 5, nous présentons une série de recherches visant à tester l'hypothèse d'une perméabilité de la procédure d'accord à des facteurs de nature non syntaxique. L'impact de variables relatives à l'information conceptuelle et à l'information morphophonologique est examiné. Les Chapitres 6 et 7 traitent de contraintes de nature proprement syntaxique pesant sur l'accord, et plus particulièrement de la pertinence de l'hypothèse d'une séparation du traitement syntaxique en un stade fonctionnel hiérarchique, et un stade positionnel linéaire, tel que le postule le modèle séquentiel. L'ensemble de ces recherches permettent de dégager une vision d'ensemble des différentes contraintes, i.e. conceptuelle, lexicale, syntaxique et morphophonologique à l'œuvre dans le traitement syntaxique en production. Une perspective globale et interlinguistique sur les résultats est avancée dans la Synthèse théorique du Chapitre 8.

⁸ Une description plus détaillée des types de genre se trouve dans l'Introduction du Chapitre 2.

Chapitre 2

When sex and syntax go hand in hand: Gender agreement in language production

In four experiments (two in French and two in Italian), we investigated whether the language production system uses conceptual information regarding biological gender in the encoding of gender agreement between a subject and a predicate. Both French and Italian have a nominal gender system that includes a distinction between nouns reflecting the sex of the referent (conceptual gender) and nouns for which the gender does not reflect the sex of the referent (grammatical gender). The experiments used a constrained sentence completion task (Vigliocco, Butterworth & Semenza, 1995). In Experiments 1 (Italian) and 2 (French) we found that errors in the agreement of gender between the subject and the predicate were more common when the subject head noun did not have any conceptual correlates. Experiments 3 and 4 established that the advantage for conceptual gender in the first two studies cannot be explained by the difference in animacy between nouns with conceptual gender (referring to humans and animals) and nouns with grammatical gender (referring to objects).

2.1 Introduction

Gender distinctions are crucial for every living being. Therefore, it is not surprising that languages have developed linguistic gender distinctions that parallel biological distinctions (Corbett, 1991). For example, most languages have different words for "mother" and "father" and "man" and "woman" and different pronouns for male and female referents (e.g., "he" and "she" in English; "il" and "elle" in French).

Languages differ in whether they include gender distinctions for words referring to entities without biological sex. Corbett (1991) distinguishes between languages with a semantic gender system and languages with a formal gender system. In the first category are languages such as English and Chinese where gender is encoded in linguistic elements only for referents having biological sex. In the second category are, for example, Romance languages. In these languages, all nouns are marked for gender, either masculine or feminine. For nouns referring to animate entities, there is often a transparent relation between the gender of the

noun and the sex of the referent (which we refer to as conceptual gender). However, for many animal names, and for a number of nouns referring to humans, the genders of the nouns do not bear any relation to the sex of the referent. The nouns in this latter class are gender marked, but the gender seems to be strictly a property of the lexical elements, as is the case for nouns referring to objects and abstract entities (referred to here as grammatical gender).

Regardless of the gender system a language possesses, gender marking is generally used to compute long-distance relations among elements in a sentence; that is, to compute agreement. For example, English subject nouns and pronouns agree in gender, as shown in example (Ex. 9).

Ex. 9 : **Mary Smith, the juror, left the court and then she went back to her room**

In richly inflected languages, such as Italian and French, gender agreement is mandatory between nouns and determiners, nouns and adjective modifiers, subject nouns and predicative adjectives, nouns and pronouns, and so on. Gender agreement is usually computed correctly during fluent language production. However, errors in gender agreement surface from time to time. Two examples from Italian are reported below.

Ex. 10 : ***Sono scese quattro persone, dei quali tre sono andati**
Are got-F down four persons-F, of which-M three are gone-M
Four persons got off, of the four, three went

Ex. 11 : ***Stanze che sono anni e anni che sono chiusi**
Rooms-F that are years-M and years-M that are closed-M
Rooms that have been closed for years and years

In (Ex. 10), the word "persone" is feminine. Therefore, the relative pronoun ("dei quali") and the past participle ("andati") should be feminine as well ("delle quali" and "andate", respectively), regardless of the gender of these four people. Interestingly, (Ex. 10) is an example of an error occurring for a noun referring to humans but with grammatical gender (that is, "persone" is feminine, regardless of the sex of the referent). In (Ex. 11) an error occurs for a noun referring to an object. Rooms, being objects, only have the formal property of feminine nouns, without any biological correlates. In this example, the presence of a masculine noun in the preverbal environment seems to have triggered the error.

Similar errors are observed in French. Two examples are provided in (Ex. 12) and (Ex. 13).

Ex. 12 : ***Le pays dans laquelle la recherche se développe le plus...**
The-M country-M in which-F research-F is most developed...

Ex. 13 : ***Il pense que la deuxième est faux.**
He thinks that the-F second-F is wrong-M

In (Ex. 12) the relative pronoun ("laquelle", F) should agree with the subject head noun ("le pays", M), but instead it agrees with the noun following the pronoun ("la recherche", F). In (Ex. 13), the predicative adjective ("faux", M)

should agree with the subject of the relative clause ("la deuxième", F), but instead it agrees with the subject of the matrix clause ("Il", M).

In this study, we looked at gender agreement errors in Italian and French to assess whether the system responsible for agreement in language production is sensitive to the distinction between grammatical and conceptual gender. Before reporting the experiments, we describe how language production models deal with agreement computation and present two contrasting hypotheses regarding the role of conceptual information during agreement processing. The literature on psycholinguistic investigations of number agreement is also reviewed. We then describe in more detail the gender systems of our two test languages.

2.2 Agreement in Models of Sentence Production

Language production entails four main components: a conceptualizer, a grammatical encoder, a phonological encoder and an articulator. The conceptualizer captures features of the speaker's intended meaning and communicative perspective and provides the input for grammatical encoding. During this stage, abstract lexical representations for words are selected. These representations, also referred to as lemmas, specify semantic and syntactic information about the word. According to a number of authors (e.g., Garrett, 1992; Levelt, 1989), the lemma includes the grammatical category of the word (i.e., noun, verb, adjective, etc.), the grammatical function it can take (i.e., subject, object, etc.), the kinds of syntactic structures it can be part of (i.e., noun phrase, verb phrase, etc.), and language-specific syntactic features (e.g., grammatical gender). Furthermore, according to some authors (Bresnan, 1982; Kempen & Hoenkamp, 1987; Levelt, 1989; Vigliocco, Butterworth & Garrett, 1996), the syntactic information specified in the lemmas guides the grammatical encoding of the sentence. A subsequent phonological encoding stage determines the sound structure and the pronunciation codes for the sentence. Lexemes, the phonological representations of words, are retrieved during this second step. Finally, the phonological representation is sent on for articulation.

In this general framework, agreement is computed during grammatical encoding, when a hierarchical structure representing the to-be-uttered sentence is constructed (Bock & Levelt, 1994; Kempen & Hoenkamp, 1987; Levelt, 1989; Vigliocco, Butterworth & Garrett, 1996). During the process of grammatical encoding, lemmas are first retrieved on the basis of the speaker's intentions. The selected lemmas are then assigned grammatical functions (e.g., subject, object, etc.) and the construction of the hierarchical frame for the sentence begins. Agreement is computed during the construction of a hierarchical frame, after grammatical functions are assigned to the different constituents but before words are placed in a linear order (Vigliocco & Nicol, 1998). According to a number of authors, frames are constructed as a sharing of agreement features between the

different sentential constituents (Kempen & Vosse, 1989; Vigliocco, Butterworth & Garrett, 1996), thereby rendering the study of agreement phenomena central to the study of grammatical encoding.

Agreement features are retrieved from conceptual representations and assigned to the lemmas when there is a conceptual motivation for them. For example, number features for nouns are usually assigned to the retrieved lemmas according to the number of entities the speaker wishes to discuss. For gender, since there is a clear distinction between conceptual and grammatical gender, we will postulate two different procedures. If the noun refers to an entity with biological sex, then the gender feature of the noun will depend on whether the speaker wants to talk about a male or female entity in a manner analogous to number features. However, for nouns referring to objects, for which gender is not a semantic property, the gender of the noun is not assigned on the basis of the speaker's intention but stored in the lexicon as an inherent property of the lemma (Vigliocco, Antonini & Garrett, 1997).

The question we address here is whether the encoder treats agreement features with a conceptual connotation (such as conceptual gender or number) differently from features with no conceptual connotation (such as grammatical gender). An answer to this question will provide us with important constraints regarding the type of input that the encoder receives from conceptual structures. In turn, the kind of input the encoder receives has consequences for how agreement computation is realized; that is, how agreement features are shared between the agreeing elements.

A first hypothesis concerning the interface between the conceptualizer and the grammatical encoder derives from theories of language production that emphasize the separation between these two levels of processing. In these theories, the impact of conceptual attributes on syntactic processes is strictly limited (Bock, 1987; Bock & Levelt, 1994; Garrett, 1976; Kempen & Hoenkamp, 1987; Levelt, 1989). According to this view, agreement is a purely syntactic operation that takes syntactic features such as the number and gender of an agreement controller (e.g., the subject of a sentence, if we consider subject-verb and subject-predicative adjective agreement) and copies them onto an agreement target (the verb or the predicative adjective). Conceptual connotations (e.g., male or female) help to establish the syntactic features (e.g., masculine or feminine) of the agreement controller (the noun). After these syntactic features have been established, no other information is retrieved from conceptual structures. We will refer to this hypothesis as the minimal input hypothesis. This hypothesis is compatible with most formal linguistic accounts of agreement in which agreement features are copied (Chomsky, 1965, 1981; Akmajian & Heny, 1975), or inherited (Gazdar, Klein, Pullum, & Sag, 1985) from a controller to a target. There are some general arguments in favor of this view of agreement processing. First, syntactic agreement is preferred cross-linguistically over conceptual agreement at least for certain

agreement relations (Corbett, 1983). In British English (and in some cases also in US English), for example, collective nouns may take a plural verb as in example (Ex. 14), denoting conceptual agreement. However, noun-specifier agreement has to be syntactic, and therefore (Ex. 15) is ungrammatical.

Ex. 14 : **The committee are voting themselves a raise**

Ex. 15 : ***These committee voted for a raise**

A second argument comes from the observation that children master agreement production well before they understand the significance of agreement features as expressed by nominal, and especially verbal, morphology of number (Keenay & Wolfe, 1972; Vigliocco & Fava, 1993). Additional support for the minimal input hypothesis comes from experimental investigations of subject-verb agreement in English reviewed below.

Alternatively, conceptual structures may exert a stronger control on the workings of the grammatical encoder. In this view, agreement processing is responsive to features of the conceptual representation beyond their syntactic features. We will refer to this possibility as the maximal input hypothesis. This position has been advocated in linguistic accounts of agreement (Barlow, 1993; Pollard & Sag, 1994) and is compatible with computational psycholinguistic models that use "feature-merging" (unification) as a mechanism for syntactic structure building (Kempen & Vosse, 1989; Kempen, 1997; McDonald, Pearlmutter and Seidenberg, 1994; Shieber, 1986; Stevenson, 1994; Vigliocco, Butterworth and Garrett, 1996).

Compatible with the maximal input hypothesis is the observation that across languages there are agreement relations for which conceptual agreement is strongly preferred over syntactic agreement (along the lines of the agreement hierarchy proposed by Comrie, 1975, and Corbett, 1983). In French, for example, the polite plural pronoun (vous, 2nd person plural used to politely refer to 2nd person singular) requires syntactic agreement (plural) with the verb while it requires conceptual agreement (singular) with the past participle (Ex. 16) or with the predicate (Ex. 17) (Comrie, 1975).

Ex. 16 : **Vous êtes venu**
You-P are-P come-S

Ex. 17 : **Vous êtes loyal**
You-P are-P loyal-S

Additional evidence for the maximal input hypothesis comes from the phenomenon referred to as notional concord (Quirk, Greenbaum, Leach & Svartvik, 1972) found both in oral and written spontaneous production. Notional concord refers to agreement with the "notion of number" expressed by a singular noun, instead of the (correct) agreement with the grammatical number. Errors like (Ex. 18) and (Ex. 19) are taken from spontaneous speech.

Ex. 18 : ***Siamo una squadra che giochiamo**
We are a team-S that play-P

Ex. 19 : ***C'est ce genre de texte qui sont très subtils**
It's this-S kind-S of text-S that are-P very subtle

A number of experiments have investigated agreement phenomena in language production by looking at errors in the agreement of number between the sentential subject and the verb, as in examples (Ex. 18) and (Ex. 19) (Bock & Miller, 1991; Bock & Cutting, 1992; Bock & Eberhard, 1993; Fayol, Largy & Lemaire, 1994; Eberhard, 1997; Vigliocco, Butterworth & Semenza, 1995; Vigliocco, Butterworth & Garrett, 1996; Vigliocco, Hartsuiker, Jarema & Kolk, 1996). Errors were induced by presenting speakers with sentential fragments such as "The king of the colonies..." and asking them to create a sentence from each fragment. In all of the published studies, the subject noun (also referred to as the head noun) and the verb were always separated by another noun (also referred to as the local noun) which either mismatched or matched in number with the head noun.

In experimental investigations concerning the role of conceptual factors in subject-verb agreement, the number of tokens in the conceptual scene denoted by a singular head noun was manipulated. This created contrasting cases in which a singular head noun either denoted a single entity, as in (Ex. 20), where the preferred reading of the preamble requires just one baby, or referred to a multiplicity of entities, as in (Ex. 21), where the preferred reading entails a number of bottles, each of which has its own label.

Ex. 20 : **The baby on the blankets**

Ex. 21 : **The label on the bottles**

Errors in agreement of number were found to be more common in Dutch, French, Italian and Spanish when the subject head noun was conceptually plural (as in Ex. 21) than when it was conceptually singular (as in Ex. 20), indicating that agreement can be influenced by conceptual numerosity in these languages (Vigliocco et al., 1995; Vigliocco, Butterworth & Garrett, 1996; Vigliocco, Hartsuiker, Jarema & Kolk, 1996). However, this effect has not been reported in English (Bock & Miller, 1991; Vigliocco, Butterworth and Garrett, 1996). Therefore, while the results obtained in English are compatible with the minimal input hypothesis, the results obtained in all the other languages are compatible with the maximal input hypothesis.

Two other general findings emerged in this series of studies. First, agreement errors were more common when the subject was separated from the verb by a local noun mismatching in number (as in "The king of the colonies") than when the local noun had the same number as the subject (as in "The king of the colony") (e.g., Bock and Miller, 1991; Vigliocco et al., 1995). These results indicate that features of the local noun can interfere with agreement computation. We will refer to this result as the effect of a local noun. Also, errors were more common when the head

noun was singular than when it was plural (Bock & Cutting, 1992; Eberhard, 1997). This latter result has been accounted for in terms of markedness: a marked (plural) local noun interferes more than an unmarked (singular) local noun in agreement computation.

To sum up, we described two ways in which information from conceptual structures could be retrieved during grammatical encoding. The minimal input hypothesis implies that the grammatical encoder only takes the "needed" information from conceptual structures. According to the maximal input hypothesis the encoder gets "all that it can" from conceptual structures. These two hypotheses, minimal input and maximal input, lead to different predictions about conceptual effects on agreement. Previous experimental studies focused on agreement in number between the sentential subject and the verb, and an effect of conceptual number was shown at least in some languages. These results provide limited support for the maximal input hypothesis.

The present study tests the two hypotheses introduced above using a different mapping between linguistic and conceptual features. We tested whether conceptual correlates can affect gender agreement in Italian and French. The case of gender agreement between the subject and a predicative adjective is interestingly different from subject-verb agreement in number. If conceptual information is taken into account in gender agreement, then for cases in which gender is conceptually specified there would be a match between the syntactic (masculine, feminine) and conceptual (male, female) features, while for grammatical gender, a lexical property of nouns' lemmas (Levelt, 1989; Vigliocco et al., 1997), features would only be specified at the syntactic level. Thus, we predict that errors in the agreement of gender would be less common when the syntactic gender has a conceptual connotation than when it does not. That is, conceptual information should help correct agreement since it provides redundant compatible information. The experiments described below also allowed us to investigate other factors, such as the effect of a gender mismatching modifier in pre-predicate position, and whether there is an asymmetry between masculine and feminine head nouns, parallel to that reported for number agreement.

2.3 Gender Systems in Italian and French

In Italian, nouns are always marked for gender (either masculine or feminine). Usually, the gender can be readily inferred from the word ending: masculine words end in *-o*, while feminine words end in *-a*. For most nouns referring to humans, the gender marking is strictly determined by the sex of the referent. The reference to male or female entities is achieved by changing the word ending; for example, *ragazzo* [boy] and *ragazza* [girl]. Nouns of this type are referred to as having conceptual gender. For nouns referring to objects and abstract entities, being masculine or feminine does not bear any conceptual force and the assignment of

these words to one or the other gender class is largely arbitrary (so that, for example, *pietra* [rock] is feminine while *sasso* [stone] is masculine.) Nouns in this category are referred to as having grammatical gender. For a large number of animal names, a single form (either feminine or masculine) is used to refer to both sexes (e.g., *scimmia* [monkey-F]; *gufo* [owl-M]). A few nouns referring to humans also belong to this latter class (e.g., *vittima* [victim-F]). It is important to note that at a morpho-phonological level, words with conceptual gender and those with grammatical gender can have identical gender features (e.g., *ragazza* and *pietra* both end in *-a*; *ragazzo* and *sasso* both end in *-o*). Both *ragazza* and *pietra* take the same form of the determiner (*la ragazza*, *la pietra*) and this is also the case for *ragazzo* and *sasso* (*il ragazzo*, *il sasso*).

Agreement with a predicate containing an adjective follows the same rule for nouns with conceptual and grammatical gender: if the noun is masculine, then the adjective must be masculine; if the noun is feminine, then the adjective must be feminine (Vincent, 1990). The sentences below illustrate this fact.

Ex. 22 : **La ragazza e' rossa**
The-FS girl-FS is red-FS (meaning: has red hair)

Ex. 23 : **La pietra e' rossa**
The-FS rock-FS is red-FS

Ex. 24 : **Il ragazzo e' rosso**
The-MS boy-MS is red-MS

Ex. 25 : **Il sasso e' rosso**
The-MS stone-MS is red-MS

The existence of such contrasts allows us to test whether errors in the agreement of gender are similarly distributed for subject Noun Phrases (NPs) in which the head noun has conceptual gender and for NPs in which the head noun has grammatical gender. Errors would result in sentences such as "*La pietra e' rosso*" or "*La ragazza e' rosso*".

In French, although the same general distinction between nouns with conceptual and grammatical gender applies, the situation is more complex. There is no general rule of morpho-phonological marking of gender for nouns with grammatical gender, although certain suffixes are associated with masculine forms (e.g., *-age*, *-isme*, *-ment*) while others are associated with feminine forms (e.g., *-tion*). In oral French, the feminine form of nouns with conceptual gender is determined on the basis of about 15 regular vocalic and/or consonantic variations from the masculine form (e.g., *-eur,M/-euse* or *-ice,F*, *-ien,M/-enne,F*, *-on,M/-onne,F*). In the written form, most feminine nouns tend to end in *-e*, whereas there is no general rule for masculine nouns.

The rule for predicative adjective agreement in French is the same for nouns with conceptual and grammatical gender. In the written modality, the feminine of the adjective is usually formed by adding *-e* to the masculine form (Harris, 1990),

with the orthographic adaptations this change requires (e.g., doubling of the final consonant). However, depending on the phonological properties of the adjective, gender marking does not always appear in the spoken modality (as illustrated in Ex. 26 and Ex. 27). For other adjectives, the same form is used for masculine and feminine in both the written and spoken modality (as illustrated in Ex. 28 and Ex. 29). Examples (Ex. 26) and (Ex. 27) also illustrate that subject-predicative adjective agreement is the same for nouns with conceptual and grammatical gender. In these examples, the marked words (^) are pronounced the same.

Ex. 26 : **Le serveur est ^gentil/ Le serveur est gros**
The-MS waiter-MS is kind-MS/ The-MS waiter-MS is fat-MS

Ex. 27 : **La serveuse est ^gentille/ La serveuse est grosse**
The-FS waitress-FS is kind-FS/ The-FS waitress-FS is fat-FS

Ex. 28 : **Le camion est ^rapide/Le camion est gros**
The-MS truck-MS is fast-MS/The-MS truck-MS is big-MS

Ex. 29 : **La voiture est ^rapide/La voiture est grosse**
The-FS car-FS is fast-FS/The-FS car-FS is big-FS

As mentioned above, there are interesting differences between the two test languages. A first difference concerns the morpho-phonological regularity of gender marking. Italian marking follows a few simple rules transparent both in the oral and written formats, whereas French marking is rather complex. Second, for Italian adjectives, the bound morpheme marking gender (and number) is part of the word (e.g., "malat-o" and "malat-a" [sick-M, sick-F]) for both feminine and masculine forms, while in French, the feminine form is derived from the masculine form through phonemic variations (e.g., "gros", "grosse" [big-M, big-F]). Using Italian and French as test languages may show whether these fine-grained differences have a role in grammatical encoding.

2.4 The Plan of the Study

We report four experiments eliciting gender agreement errors between the subject and a predicative adjective in Italian and French. In all of the experiments, participants were presented first with an adjective and then with a sentential fragment. Their task was to repeat the fragment and complete it with the adjective. All the sentential fragments included two nouns, the subject head noun and a local noun, embedded in a prepositional phrase that modified the subject head.

Experiments 1 (Italian) and 2 (French) assessed whether head nouns with conceptual or grammatical gender induce different numbers of agreement errors. Head nouns with grammatical gender are marked at the syntactic level and not at the conceptual level, while head nouns with conceptual gender are congruently marked at both levels. If conceptual information is taken into account by the encoder, agreement errors should be less common for head nouns with conceptual gender than for head nouns with grammatical gender. Experiment 2 (French) also

assessed the role of a local noun which did not match the head noun for gender. Experiments 3 (Italian) and 4 (French) assessed whether differences in animacy could account for any difference between nouns with conceptual and grammatical gender in the previous experiments. Because nouns referring to animate entities have conceptual gender while nouns referring to inanimate entities have grammatical gender, animacy was confounded with gender type in Experiments 1 and 2.

In terms of the general model of production presented above, we assume that nouns with grammatical gender have the gender feature set at the lexical (lemma) level. Nouns with conceptual gender receive specification for the gender feature from conceptual structures (e.g., in Italian, there is one lemma for "child" that could receive specification for male "bambino" or female "bambina").

2.5 Experiment 1: Italian

The materials consisted of sentential fragments (composed of a head noun and a local noun) and adjectives that could be used as plausible predicates for the fragments. We manipulated the gender type (conceptual vs. grammatical) of the subject head noun. The minimal input hypothesis predicts no difference between the number of agreement errors in these two conditions, since under this view only syntactic features are relevant for agreement computation. However, under the maximal input hypothesis, errors should be less common for subject head nouns with conceptual gender, since the syntactic and the conceptual features match and therefore can reinforce each other. The syntactic gender (masculine or feminine) of the head noun was also manipulated, while local nouns always mismatched the head noun gender. The rationale for including only mismatching local nouns was to attempt to induce the greatest number of errors possible. Some indication that the use of a gender mismatching local noun induces more errors than a matching local noun comes from observations of spontaneously occurring gender agreement errors (as in Ex. 10 and Ex. 11 above), and the well-documented effect of a local noun on errors in the agreement of number (e.g., Bock, 1995).

2.5.1. Method

Participants. Forty undergraduate students from the Department of Psychology, University of Trieste, volunteered for the study. They were all native speakers of Italian.

Materials. Materials consisted of sentence beginnings (preambles), composed of a subject head noun and a local noun, embedded in a prepositional noun phrase that modified the subject noun phrase and adjectives that could plausibly be used to complete the preambles. The variables experimentally manipulated were (1) the gender type of the subject head noun (conceptual gender vs. grammatical gender); (2) the gender of the subject head noun (masculine vs. feminine). For nouns with

conceptual gender, a pre-test was conducted in order to ensure that for each experimental form, there was no strong bias toward either the masculine or feminine form in the language. That is, nouns that can take both the masculine and feminine form are often biased toward one gender (e.g., "operaio" (worker-M) is more common than the corresponding feminine form "operaia") In order to exclude words of this kind, a list of 46 words with conceptual gender was presented to 20 students from the same population as in the main study. Their task was to judge how often they thought those words were used to refer to a man or to a woman in the following fashion. For each word, the two forms (masculine and feminine) were presented as end points of a continuum (eg., mago [wizard-M] 4-3-2-1-0-1-2-3-4 maga [wizard-F]). Participants were instructed to rate the relative frequency of usage of both forms. A judgement of "0" meant that speakers perceived both forms as equally common. Only those words that received a mean judgement of no greater than 1.5 in either direction were included among the experimental items.

In the experimental items, the local noun always mismatched the head noun gender, leading to two combinations - Masculine head noun, Feminine local noun and Feminine head noun, Masculine local noun. All nouns used in the experimental materials were singular. All the local nouns had grammatical gender. Pairs of experimental items were created in the following way. In each pair the same local noun was used with both a head noun with conceptual gender and a head noun with grammatical gender. In each pair, the head nouns were matched for number of syllables. The gender of the head nouns in a pair was also always the same, since only gender mismatching items were used. Examples of experimental preambles are provided in Table 1 (Tableau 1). Adjectives for each pair of preambles were also matched for number of syllables.

For each preamble, the two forms of the adjective (masculine and feminine) were presented, one above the other. The position of each form (above or below) of the adjective was counterbalanced in two lists. All the nouns used in the experimental materials were transparently marked for gender. That is, the word ending was *-a* for feminine nouns and *-o* for masculine nouns.

Two experimental lists were generated. Each list was composed of 64 experimental preambles and 64 fillers. Each pair of experimental preambles (i.e., preambles with the same local noun but different head nouns) was presented within the same list. There were eight items per condition in each list. The difference between the two lists was whether the masculine form of the adjective for each item was presented above or below the feminine form. Filler preambles also consisted of a noun phrase and a prepositional phrase. In 48 of the filler preambles, the gender of the head noun and of the local noun matched, while it mismatched in the remaining 16. Thirty-two of the filler preambles had a plural head and a plural local noun, and the remaining 32 had a plural head and a singular local noun. In 32

filler preambles the head noun referred to an animate entity, while it referred to an inanimate entity in the remaining 32 items⁹.

Tableau 1 (Table 1): Examples of experimental sentence preambles for Experiment 1 (Italian)

Conceptual Gender	
<u>Masculine Head Noun, Feminine Local Noun</u>	<u>Feminine Head Noun, Masculine Local Noun</u>
Lo sposo in chiesa <i>The-M groom-M in church-F</i>	La zingara nel parcheggio <i>The-F gypsy-F in the-M parking lot-M</i>
L'inquilino della casa <i>The tenant-M of the-F house-F</i>	La bidella nell'atrio <i>The-F janitor-F in the-M hall-M</i>
Grammatical Gender	
<u>Masculine Head Noun, Feminine Local Noun</u>	<u>Feminine Head Noun, Masculine Local Noun</u>
Il cero in chiesa <i>The-M candle-M in church-F</i>	La macchina nel parcheggio <i>The-F car-F in the-M parking lot-M</i>
Lo sgabuzzino della casa <i>The-M closet-M of the-F house-F</i>	La lampada nell'atrio <i>The-F lamp-F in the-M hall-M</i>

Procedure. Trials consisted of the presentation of the adjectives (masculine and feminine form) followed by the preamble. The participants' task was to repeat the preamble and complete it using the adjective. The two forms of the adjective, one above the other, were presented for 600ms at the center of the computer screen, followed by an interval of 400ms. The preamble was then presented for 1s at the center of the computer screen. The presentation of the materials was self-paced; participants were instructed to press the space bar on the computer keyboard in order to proceed from one trial to the next. Instructions emphasized rapid speech. Eight practice trials, performed before the experimental trials, ensured that the participants understood the instructions. The experimental sessions were tape recorded using an analog recording system (Marantz 201) and then transcribed.

Filler and experimental stimuli were presented in a pseudo-random order with the constraints that each experimental list started with at least four filler items, and no more than three experimental items were presented in succession.

Scoring Criteria. Produced utterances were transcribed and scored according to the following criteria. (1) Correct Responses were scored when the participant correctly repeated the preamble and completed the sentence using the correctly inflected form of the adjective. (2) Agreement Errors were scored when the

⁹ A full listing of the materials used in Experiments 1 - 4 can be found at this URL: <http://psych.wisc.edu/faculty/pages/gvigliocco/exps/appendix.html>.

completion met the criteria above but the adjective failed to agree in gender with the subject of the sentence. (3) Miscellaneous Responses were scored when the participant failed to repeat the preamble or the adjective (or parts of it), when he/she failed to repeat some words in the preamble, or when he/she produced a completion lacking the main verb. If two different utterances were produced in succession, only the first was scored, including those cases in which an agreement error was produced and immediately corrected.

Design and data analysis. Analyses of variance with both subjects and items as random factors were carried out using as the dependent measures the numbers of agreement errors and the numbers of miscellaneous responses. In each analysis, the following factors were orthogonally combined: (1) Gender type of the head noun (conceptual vs. grammatical), and (2) gender of the head noun (masculine vs. feminine)¹⁰. In this experiment, as in all the following experiments, the alpha level used for significance testing was 0.05.

2.5.2. Results

As Table 2 (Tableau 2) shows, there were 2253 (88%) correct responses, 51 (2%) gender agreement errors and 256 (10%) miscellaneous responses.

Distribution of agreement errors. Agreement errors were more common when the head noun had grammatical gender than when it had conceptual gender. No asymmetry between masculine and feminine nouns was observed. Analyses of variance revealed a main effect of the gender type of the head noun ($F(1,39) = 14.00$, $p = 0.001$; $F(1,124) = 18.42$, $p < 0.001$). There was no difference between masculine and feminine head nouns ($F(1,39) = 1.44$, $p = 0.238$; $F(1,124) = 2.05$, $p = 0.155$) and the interaction between the gender type and the gender of the head noun was not significant ($F_s < 1$).

Tableau 2 (Table 2) : Number of responses for each scoring category, Experiment 1

	Feminine head noun		Masculine head noun	
	Conceptual gender	Grammatical gender	Conceptual gender	Grammatical gender
Correct responses	569	556	558	570
Agreement errors	8	22	4	17
Miscellaneous	63	62	78	53

¹⁰ In this experiment and the following experiments we also considered the position of the adjective (above or below) as a factor in the analyses. This factor was introduced to assess whether participants paid more attention to one or the other position and therefore were affected when the incongruent adjective was in the preferred location. However, we consistently failed to observe any main effect or interaction related to this factor, and so these analyses are not reported.

Distribution of miscellaneous responses. None of the factors or interactions were significant: gender type of the head noun ($F(1,39) = 2.00, p = 0.165$; $F(1,124) = 2.71, p = 0.102$); gender of the head nouns ($F_s < 1$); interaction between the two ($F(1,39) = 2.40, p = 1.123$; $F(1,124) = 2.31, p = 0.131$).

2.5.3. Discussion

Experiment 1 showed that errors were more common when the head noun had grammatical gender than when it had conceptual gender. No difference was found between masculine and feminine head nouns. The effect of gender type lends some support to the maximal input hypothesis. The grammatical encoder does not use only the syntactic features of the subjects. When available, it also uses the conceptual correlates of these features.

We failed to observe an asymmetric distribution of errors for masculine and feminine head nouns. There are grounds for believing that the masculine form of nouns as well as adjectives is the unmarked form (this form is used in conjoined noun phrases in which there is a masculine and a feminine noun; it is used as the unmarked form of past participles; and when the reference is to a generic entity). However, this difference did not affect error production. Perhaps a tendency to consider the masculine as the unmarked form for nouns with conceptual gender was neutralized in the present experiment because we excluded all those nouns that presented a marked bias. However, this fact cannot account for the lack of asymmetry, since the large majority of errors occurred with nouns with grammatical gender. For these nouns, there is only one form, either masculine or feminine.

This first experiment did not include a control condition in which the gender of the local noun matched the subject head noun. It is therefore not clear whether the local noun played a role in increasing the number of errors. Moreover, if the local noun indeed affects the computation of agreement (increasing errors only in case of gender mismatch with the subject), it may be that this effect interacts with the effect of the gender type of the head noun. In order to investigate this possible interaction, Experiments 2 (French) and 3 (Italian) used a fully factorial design with all possible combinations of masculine and feminine gender between the head noun and the local noun.

2.6 Experiment 2: French

The main purpose of this experiment was to assess whether the effect of gender type found in Italian could also be found in French, using the same procedure as Experiment 1. This experiment also provided an additional control: we systematically varied the gender (masculine, feminine) of the local noun in order to explore whether the traditional mismatch effect could also be found in gender agreement. On the basis of previous findings regarding errors in the agreement of

number (e.g., Bock & Miller, 1991; Vigliocco et al., 1995) we expected that agreement errors would be more common when the head noun and the local noun mismatch in gender than when they match. Local nouns with conceptual gender and local nouns with grammatical gender were used in the present experiment, balanced between experimental conditions.

2.6.1. Method

Participants. Seventy undergraduate students from the Department of Psychology at the Université catholique de Louvain took part in the experiment. All were native French speakers. They received credit for their participation.

Materials. Experimental items consisted of sentence preambles and adjectives. Preambles contained a subject head noun and a local noun embedded in a prepositional phrase attached to the head noun. The manipulated variables were (1) the gender type of the head noun (conceptual vs. grammatical), (2) the gender of the head noun (feminine vs. masculine), and (3) the gender of the local noun.

Frequency of occurrence of the head nouns was kept constant (between 150 and 250 per 100 million in the Brulex database, Content, Mousty & Radeau, 1990), and all head nouns were bisyllabic. Adjectives were plausible continuations of the preambles.

There were 8 (2x2x2) experimental conditions according to the gender of the head and of the local nouns and the gender type of the head noun (conceptual, grammatical). Half of the experimental items had a local noun with conceptual gender and the other half had a local noun with grammatical gender. A total of 32 items were created, with 4 items in each condition. The lexical content was different for each item. Adjectives were always plausible modifiers of the subject head nouns. All items were part of one list given to all participants. Examples of experimental preambles with head nouns with conceptual and grammatical gender, and matching and mismatching local nouns are presented in Table 3 (Tableau 3).

The masculine and feminine forms of each adjective were presented one above and one below the center of the screen. The position of the two forms was counterbalanced: for half the items, the masculine form was above; in the other half it was below. All adjectives were marked for gender in the spoken modality.

There were 16 filler items: eight single head nouns and eight complex noun phrases consisting of a head noun followed by a relative clause. The list started with three filler items and was preceded by eight practice trials.

Procedure. Participants were tested individually. The two forms of the adjective were presented visually for 800 ms and replaced by a blank screen for 500 ms. The preamble was then displayed for 900 ms. Participants were asked to silently read the adjectives and the preambles. As soon as the preamble disappeared, they were

instructed to repeat the preamble and complete it with the adjective. The instructions emphasized rapid speech.

Tableau 3 (Table 3): Examples of Experimental Sentence Preambles for Experiment 2 (French)

Head nouns with conceptual gender	
<u>Feminine Head Noun, Feminine Local Noun</u>	<u>Feminine Head Noun, Masculine Local Noun</u>
La soeur de la pharmacienne <i>The-F sister-F of the-F chemist-F</i>	La présidente du jury <i>The-F president-F of the-M jury-M</i>
<u>Masculine Head Noun, Feminine Local Noun</u>	<u>Masculine Head Noun, Masculine Local Noun</u>
Le gardien de la prisonnière <i>The-M guardian-M of the-F prisoner-F</i>	Le délégué du groupe <i>The-M delegate-M of the-M group-M</i>
Head Nouns with grammatical gender	
<u>Feminine Head Noun, Feminine Local Noun</u>	<u>Feminine Head Noun, Masculine Local Noun</u>
La couleur de la robe <i>The-F color-F of the-F dress-F</i>	La sortie du tunnel <i>The-F exit-F of the-M tunnel-M</i>
<u>Masculine Head Noun, Feminine Local Noun</u>	<u>Masculine Head Noun, Masculine Local Noun</u>
Le travail de la couturière <i>The-M work-M of the-F dressmaker-F</i>	Le visage du voleur <i>The-M face-M of the-M robber-M</i>

The experimenter presented the following item as soon as the preamble had been completed by the participant. Experimental sessions were tape recorded using an analogue recording system (Marantz 210).

Scoring Criteria. The same scoring criteria were used as in Experiment 1.

Design and data analysis. The three experimental variables (gender of the head noun, gender of the local noun and gender type of the head noun) were part of a repeated measures design by subjects. The design was completely between items, as the lexical content of the items was different among the conditions. 2X2X2 ANOVAs by subjects (F1) and by items (F2) were performed. The gender type of the local noun was not included in the analysis of variance (being introduced to balance the type of local nouns). However, given its potential impact, we conducted a post-hoc t-test to assess whether there were differences between local nouns with conceptual and grammatical gender.

2.6.2. Results

There were 1941 (87%) correct answers, 56 (2%) agreement errors, and 243 (11%) miscellaneous responses. The distribution of responses is presented in Table 4 (Tableau 4).

Tableau 4 (Table 4) : Number of Responses for each scoring category, Experiment 2

Head nouns with conceptual gender				
	Feminine head noun		Masculine head noun	
	Feminine local noun	Masculine local noun	Feminine local noun	Masculine local noun
Correct responses	245	224	233	253
Agreement errors	3	8	5	2
Miscellaneous	32	48	32	25

Head nouns with grammatical gender				
	Feminine head noun		Masculine head noun	
	Feminine local noun	Masculine local noun	Feminine local noun	Masculine local noun
Correct responses	254	241	224	267
Agreement errors	9	20	9	0
Miscellaneous	17	19	47	13

Distribution of agreement errors. Agreement errors were more common when the head noun had grammatical rather than conceptual gender and when the head noun and the local noun mismatched in gender. Furthermore, errors were more common when the head noun was feminine. The gender type of the head noun was significant in both analyses ($F(1,69) = 5.17, p = 0.026$; $F(1,24) = 4.84, p = 0.038$). There was a main effect of the gender of the head noun, with more errors for feminine than masculine subjects ($F(1,69) = 13.59, p < 0.001$; $F(1,24) = 6.97, p = 0.014$). No effect of the gender of the local noun was found ($F_s < 1$). The interaction between the gender of the head noun and the gender of the local noun was significant ($F(1,69) = 14.10, p < 0.001$; $F(1,24) = 9.48, p = 0.005$), with more errors when the two nouns mismatched than when they matched in gender. The analysis by subjects also revealed a significant interaction between the gender of the head noun and the gender type of the head noun ($F(1,69) = 5.01, p = 0.028$), but this was not significant in the analysis by items ($F(1,24) = 3.10, p = 0.091$). The three-way interaction between the gender of the head noun, the gender of the local noun and the gender type of the head noun was not significant ($F(1,69) = 2.86, p = 0.091$; $F(1,24) = 1.74, p = 0.199$).

A post-hoc t-test, contrasting local nouns with grammatical gender to local nouns with conceptual gender, was conducted to assess whether the gender type of the local noun had an impact on error induction. Of the 56 agreement errors, 29 occurred on items with a local noun with grammatical gender while 27 occurred on items with a local noun with conceptual gender. The difference between these two conditions was not significant ($M = 0.03$, $SD = 0.82$, $t(69) = 0.29$, $p = 0.77$).

Distribution of miscellaneous responses. Miscellaneous responses showed a complex pattern of main effects and interactions that, however, were significant only in the subjects and not in the items analysis. The gender type of the head noun was significant in the subject analysis ($F(1,69) = 7.00$, $p = 0.01$) with more miscellaneous responses for nouns with conceptual than grammatical gender, but not in the item analysis ($F(1,24) < 1$). The gender of the head noun was also significant by subjects ($F(1,69) = 13.36$, $p < 0.001$), with more miscellaneous responses for nouns with masculine than feminine gender, but not by items ($F(1,24) = 1.89$, $p = 0.182$). All two-ways interactions were significant by subjects but not by items: the interaction between the gender of the head noun and the gender of the local noun ($F(1,69) = 17.98$, $p < 0.001$, $F(1,24) = 3.46$, $p = 0.075$); the interaction between the gender type of the head noun and the gender of the head noun ($F(1,69) = 6.44$, $p = 0.013$, $F(1,24) < 1$), and between the gender type of the head noun and the gender of the local noun ($F(1,69) = 5.80$, $p = 0.019$, $F(1,24) < 1$). The three-way interaction between the gender of the head noun, the gender of the local noun and the gender type of the head noun was not significant in either analyses ($F_s < 1$).

2.6.3. Discussion

We found a gender type effect in French, parallel to the results for Italian. Fewer agreement errors were produced when the subject head noun had conceptual gender than when it had grammatical gender. This result favors the maximal input hypothesis, according to which the conceptual correlates of gender are taken into account by the encoder in computing agreement. The fact that both languages showed an effect of the gender type indicates that the principle of maximal input is not a peculiarity of one isolated language but generalizes across different Romance languages. Figure 6 compares the percentages of errors in the gender mismatch condition (where the genders of the head noun and local noun differ) in French and Italian.

The design of this second experiment included a control condition in which both the local noun and the head noun had the same gender. We found a local noun effect: nearly four times more errors were produced in the gender mismatch condition than in the gender match condition. The magnitude of this effect was similar for feminine and masculine head nouns, and this did not interact with the effect of gender type. This result parallels what has been reported in the literature concerning errors in the agreement of number between the subject and the verb of a

sentence (Bock & Miller, 1991; Bock & Eberhard, 1993; Vigliocco et al, 1995; Vigliocco, Butterworth & Garrett, 1996).

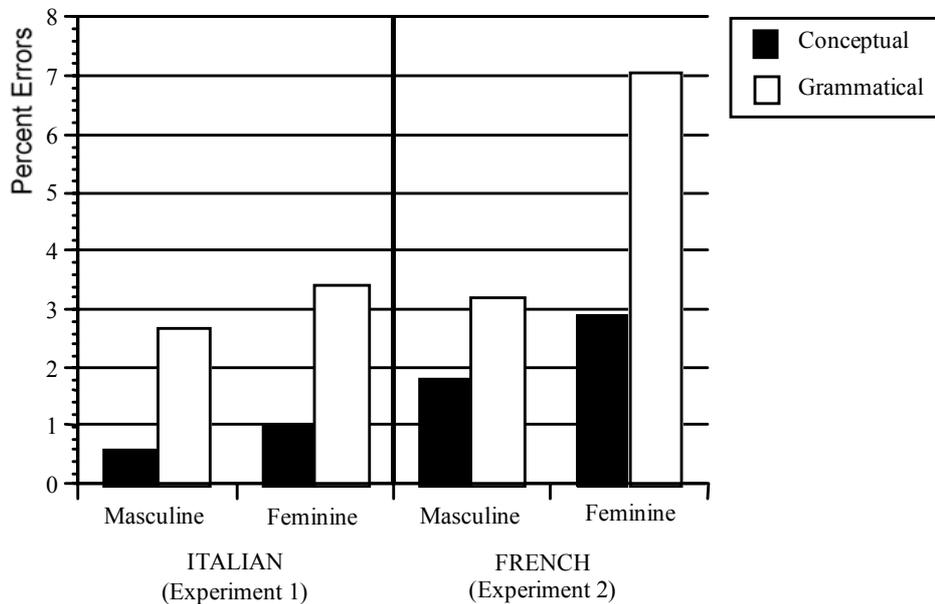


Figure 6. Percent agreement errors in Italian (Experiment 1) and French (Experiment 2). The percentages refer to errors occurring in the gender mismatching condition, to allow a comparison between the two languages, for masculine and feminine head nouns with conceptual and grammatical gender.

In contrast to Italian, it was found that the gender of the head noun was an important factor in inducing agreement errors. Errors were more common when the head noun was feminine than when it was masculine. In other words, speakers more often produced the masculine form of the adjective. As mentioned above, the masculine form of adjectives can plausibly be considered as the unmarked form. A possible explanation for this result is presented in the General Discussion.

We also observed a number of main effects and interactions in the miscellaneous responses. All these effects were only significant by subjects and not by items, raising the possibility that some idiosyncratic features of the items influenced the results of the experiment. However, the main contrast between nouns with conceptual and grammatical gender was replicated in Experiment 4.

These first two experiments showed an effect of gender type of the subject in both Italian and French. However, animacy co-varied systematically with gender type in both experiments. Errors might have been more common for head nouns with conceptual gender than for nouns with grammatical gender because the former refer to animate entities while the latter refer to inanimate entities. Experiments 3 (Italian) and 4 (French) assessed the role of animacy.

2.7 Experiment 3: Italian

Animacy clearly has an impact on language production. Animate entities tend to be agents and tend to be mapped into the subject position more often than inanimate entities (Bock & Warren, 1985). Does animacy play a role in inducing agreement errors? Bock and Miller's (1991) study of number agreement showed that the animacy of the head noun did not influence agreement errors when grammatical functions were unambiguous. However, with respect to gender agreement, animacy may be relevant because the type of gender (conceptual, grammatical) strongly correlates with it. Nouns with conceptual gender refer to animate entities and nouns with grammatical gender usually refer to inanimate entities.

Experiment 3 was conducted in order to test the effect of animacy of the subject head noun in Italian. Animate head nouns with grammatical gender (i.e., nouns referring to animals and humans that have a fixed grammatical gender and no conceptual correlates) were compared to inanimate nouns with grammatical gender. That is, animacy was manipulated, while the gender type of the subject was kept constant (purely syntactic with no conceptual features). We also manipulated the gender of the local noun to look for an effect of a mismatching local noun, as documented in Experiment 2.

2.7.1. Method

Participants. Sixty-four undergraduate students from the Department of Psychology at the University of Trieste, none of whom participated in Experiment 1, volunteered for the present experiment. They were all native speakers of Italian.

Materials. As in the previous experiments, materials consisted of preambles composed of a subject head noun and a local noun embedded in a prepositional phrase, and adjectives to be used in the completions. The manipulated variables were (1) the animacy of the head noun (animate vs. inanimate); (2) the gender of the head noun (masculine vs. feminine) and (3) the gender of the local noun (masculine vs. feminine). All the nouns in the experimental items had grammatical gender, and the animate head nouns referred to animate entities that have just one gender form in Italian (such as animal names: e.g., "scimmia" [monkey-F], "gufo" [owl-M]). All nouns used in the experimental materials were morphologically transparent with respect to gender, ending in *-o* for masculine nouns and *-a* for feminine nouns. All nouns in the experimental items were singular. As in Experiment 1, pairs of preambles were created in which the same local noun was used for both an animate and an inanimate head noun (as exemplified in Table 5 (Tableau 5)). Furthermore, the same head noun was also used for different (gender matched and gender mismatched) local nouns. Experimental preambles were rated for plausibility by an additional group of eight undergraduate students from the same population. There was no significant difference in plausibility between items

with animate and inanimate head nouns (3.8 and 3.6, respectively on a scale from 0 to 5).

Tableau 5 (Table 5): Examples of experimental sentence preambles for Experiment 3 (Italian) and their organization in the experimental lists (Lists 3 and 4 used the same items with the counterbalanced order of the adjectives).

List 1	List 2
<u>Masculine Head Noun, Masculine Local Noun</u>	<u>Masculine Head Noun, Feminine Local Noun</u>
<u>Animate</u>	
Il ghepardo nel villaggio <i>The-M cheetah-M in-the-M village-M</i>	Il ghepardo nella foresta <i>The-M cheetah-M in-the-F forest-F</i>
<u>Inanimate</u>	
Il capanno nel villaggio <i>The-M hut-M in-the-M village-M</i>	Il capanno nella foresta <i>The-M hut-M in-the-F forest-F</i>
<u>Masculine Head Noun, Feminine Local Noun</u>	<u>Masculine Head Noun, Masculine Local Noun</u>
<u>Animate</u>	
Il delfino nella vasca <i>The-M dolphin-M in-the-F tub-F</i>	Il delfino nell'acquario <i>The-M dolphin-M in-the aquarium-M</i>
<u>Inanimate</u>	
Il ciottolo nella vasca <i>The-M rock-M in-the-F tub-F</i>	Il ciottolo nell'acquario <i>The-M rock-M in-the aquarium-M</i>
<u>Feminine Head Noun, Feminine Local Noun</u>	<u>Feminine Head Noun, Masculine Local Noun</u>
<u>Animate</u>	
L'aquila sulla cima <i>The-F eagle-F on-the-F peak-F</i>	L'aquila nel cielo <i>The-F eagle-F in-the-M sky-M</i>
<u>Inanimate</u>	
La nuvola sulla cima <i>The-F cloud-F on-the-F peak-F</i>	La nuvola nel cielo <i>The-F cloud-F in-the-M sky-M</i>
<u>Feminine Head Noun, Masculine Local Noun</u>	<u>Feminine Head Noun, Feminine Local Noun</u>
<u>Animate</u>	
La talpa nell'orto <i>The-F mole-F in-the garden-M</i>	La talpa nella buca <i>The-F mole-F in-the-F hole-F</i>
<u>Inanimate</u>	
La zappa nell'orto <i>The-F hoe-F in-the garden-M</i>	La zappa nella buca <i>The-F hoe-F in-the-F hole-F</i>

Four lists were created in which each experimental condition was represented by four items. The order (above/below) of the masculine and feminine form of adjectives was counterbalanced. Table 5 (Tableau 5) reports examples of items in the different conditions for Lists 1 and 2 (the only difference between Lists 1-2 and Lists 3-4 concerned the position of the masculine and feminine form of the

adjective). Filler items were the same as used in Experiment 1. Each list consisted of 128 items (64 experimental and 64 fillers).

Procedure. The same procedure was used as in Experiment 1.

Scoring. The scoring was the same as in Experiments 1 and 2.

Design and Data Analysis. Analyses of variance with both subjects and items as random factors were carried out using as the dependent measures the numbers of agreement errors and the numbers of miscellaneous responses. In each analysis, the following factors were orthogonally combined: (1) Animacy of the head noun (animate vs. inanimate); (2) gender of the head noun (masculine vs. feminine) and (3) gender of the local noun (masculine vs. feminine). All factors were part of a repeated measures design by subjects. The design by items was treated as a mixed design as factors (1) and (2) were between items whereas factor (3) was treated as within items.

2.7.2. Results

There were 3630 (89%) correct responses, 60 (1%) gender agreement errors and 406 (10%) miscellaneous responses. The distribution of responses is presented in Table 6 (Tableau 6).

Tableau 6 (Table 6) : Number of Responses for each scoring category in Experiment 3.

Animate nouns				
	Feminine head noun		Masculine head noun	
	Feminine local noun	Masculine local noun	Feminine local noun	Masculine local noun
Correct responses	468	450	433	458
Agreement errors	1	15	15	2
Miscellaneous	43	47	64	52
Inanimate nouns				
	Feminine head noun		Masculine head noun	
	Feminine local noun	Masculine local noun	Feminine local noun	Masculine local noun
Correct responses	461	447	454	459
Agreement errors	1	15	10	1
Miscellaneous	50	50	48	52

Distribution of agreement errors. The interaction between the gender of the head noun and the gender of the local noun was significant ($F(1,63) = 48.31, p < 0.001$; $F(1,188) = 49.66, p < 0.001$). This interaction reflects an effect of the local

noun with more errors for preambles in which the gender of the head and local nouns mismatched. None of the other main effect or interactions achieved significance (all $F_s < 1$).

Distribution of miscellaneous responses. No difference between masculine and feminine head nouns was found ($F(1,63)=1.156$, $p = 0.216$; $F_2 < 1$). Similarly, there was no effect of the gender of the local noun ($F_s < 1$). The interaction between animacy and the gender of the head noun was not significant ($F(1,63)=1.45$, $p = 0.232$; $F_2 < 1$), nor the interaction between animacy and the gender of the local noun or the three way interaction (all $F_s < 1$).

2.7.3. Discussion

While no effect of animacy was found, there was a strong effect of the gender match or mismatch between the head and the local noun, replicating the results obtained in Experiment 2 (French). As in Experiment 1 (Italian), but in contrast to Experiment 2 (French), we did not observe an asymmetric distribution of agreement errors between masculine and feminine head nouns. The main conclusion from this study is that the difference reported in Experiment 1 between animate nouns with conceptual gender and inanimate nouns with grammatical gender cannot be attributed to animacy. Experiment 4 aimed to replicate these results in French, controlling for the animacy of the head nouns.

2.8 Experiment 4: French

In this last experiment, the gender type of the head noun was manipulated while keeping the animacy of the head noun constant. This was achieved by using nouns referring to animate entities with grammatical gender and contrasting them to nouns with conceptual gender (so that both types of nouns referred to animate entities). The local nouns always had grammatical gender which mismatched with the gender of the head nouns.

2.8.1. Method

Participants. Thirty students at the Department of Psychology of the Université catholique de Louvain took part in the experiment. None of them participated in Experiment 2. They were all native French speakers. They received course credit for their participation.

Materials. Experimental items consisted of sentence preambles and adjectives. The manipulated variables were (1) the gender of the head noun (masculine vs. feminine) and (2) the gender type of the head noun (conceptual vs. grammatical gender). In all of the experimental items, the local noun mismatched the head noun in gender.

All of the head nouns were animate nouns, while all of the local nouns were inanimate nouns. Frequencies of the head nouns were kept constant (between 150 and 250 per 100 million in the Brulex database, Content et al., 1990), and the head nouns were always bisyllabic. Adjectives were plausible continuations of the preambles.

Preambles were prepositional phrases attached to a subject head noun. There were 32 preambles divided into 4 experimental conditions according to the gender (feminine or masculine) and the gender type (conceptual or grammatical) of the head noun. The lexical content was different for each item. Both forms of the adjectives were presented on the screen, one above the other. The position of the two forms was balanced; the masculine form was above in half of the items and below in the other half. Examples of experimental items are presented in Table 7 (Tableau 7).

Tableau 7 (Table 7) : Examples of Experimental Sentence Preambles for Experiment 4 (French)

Conceptual Gender	
<u>Masculine Head Noun, Feminine Local Noun</u>	<u>Feminine Head Noun, Masculine Local Noun</u>
Le prieur dans l'église <i>The-M prior-M in the-F church-F</i>	La jument du manège <i>The-F mare-F of the-M riding school-M</i>
Grammatical Gender	
<u>Masculine Head Noun, Feminine Local Noun</u>	<u>Feminine Head Noun, Masculine Local Noun</u>
Le chaton sur la moquette <i>The-M kitty-M on the-F carpet-F</i>	La méduse sur le sable <i>The-F jellyfish-F on the-M sand-M</i>

Experimental items were part of one list given to all participants. There were 32 filler items: 12 single head nouns, 12 head nouns followed by relative clauses and 8 head nouns followed by prepositional phrases. Half of the fillers contained adjectives with no spoken gender inflection. The list started with three filler items and there were never more than three experimental items in a row.

Procedure. The procedure was the same as in Experiment 2.

Scoring. Scoring was the same as in the previous experiments.

Design and data analysis. The two experimental variables (gender and gender type) were part of a within-subjects experimental design. The design was between items, as the lexical content of the items was different in the different conditions. 2X2 ANOVAs by subjects (F1) and by items (F2) were performed on the data.

2.8.2. Results

There were 752 (78%) correct answers, 42 (5%) agreement errors, and 166 (17%) miscellaneous responses. The distribution of responses among the three scoring categories is presented in Table 8 (Tableau 8).

Tableau 8 (Table 8) : Number of responses for each scoring category, Experiment 4

	Feminine head noun		Masculine head noun	
	Conceptual gender	Grammatical gender	Conceptual gender	Grammatical gender
Correct responses	197	159	209	187
Agreement errors	6	21	3	12
Miscellaneous	37	60	28	41

Distribution of agreement errors. More agreement errors were produced when the head noun had grammatical gender than when it had conceptual gender ($F(1,29) = 20.78$, $p < 0.001$; $F(1,28) = 6.36$, $p = 0.018$). Although there were nearly two times more errors for feminine head nouns than for masculine ones, this difference was only marginally significant in the analysis by subjects, and was not significant in the analysis by items ($F(1,29) = 3.74$, $p = 0.063$; $F(1,28) = 1.59$, $p = 0.218$). There was no interaction between the gender of the head noun and its gender type ($F_s < 1$).

Distribution of miscellaneous responses. The analysis by subjects showed that significantly more miscellaneous responses were produced when the head noun was feminine ($F(1,29) = 8.07$, $p = 0.008$). However, this difference was not significant in the item analysis ($F(1,28) = 1.69$, $p = 0.204$). The effect of gender type, with more miscellaneous responses for nouns with conceptual than grammatical gender was also significant only in the subjects analysis ($F(1,29) = 14.11$, $p = 0.001$; $F(1,28) = 2.80$, $p = 0.105$). No significant interaction was observed (all $F_s < 1$).

2.8.3. Discussion

When animacy was kept constant, we still found an effect of the gender type of the head noun. Furthermore, we found a marginal tendency toward more errors for feminine than masculine nouns, in the same direction as we observed in Experiment 2. Therefore, animacy cannot account for the different pattern of results with nouns with grammatical and conceptual gender. This result is addressed further in the General Discussion.

In both Experiments 2 and 4, we found more miscellaneous responses for nouns with conceptual than grammatical gender (a result significant by subjects only). This trend, in the opposite direction to that found for agreement errors, can

be explained considering that the feminine and masculine forms can be confused for nouns with conceptual gender, creating opportunities for repetition errors such as "La présidente du jury..." [The-F president-F of the jury...], given "Le président du jury..." [The-M president of the jury...]. This situation does not apply to nouns with grammatical gender, where there is only one form: either feminine or masculine.

2.9 General Discussion

In this series of four experiments, we found that gender agreement errors between a subject and a predicative adjective were more common when the head noun did not have any conceptual connotation of gender than when it did. This was true for both French and Italian. We also found that this effect cannot be accounted for by the difference in animacy between nouns with conceptual and grammatical gender. These results are inconsistent with the minimal input hypothesis. Another important factor in inducing gender agreement errors was the presence of a gender mismatching local noun (Experiment 2, French and Experiment 3, Italian). This finding parallels previous cross-linguistic studies of number agreement showing that agreement errors are more common when the subject head noun is separated from the verb by a number mismatching local noun (e.g., Bock & Miller, 1991; Vigliocco et al., 1995). We also observed a tendency for an asymmetrical distribution of errors for masculine and feminine head nouns in French, with more errors for feminine than for masculine nouns (Experiment 2, and a trend in this direction in Experiment 4). However, errors were similarly distributed for masculine and feminine forms in Italian (Experiments 1 and 3). In the following section, these three aspects of the results are discussed in turn and compared to parallel results emerging from number agreement studies.

2.9.1. Comparisons Between Number and Gender Agreement

A first comparison concerns the role of conceptual features. Conceptual effects on subject-verb number agreement have been reported for different languages (see Vigliocco et al, 1995 for Italian; Vigliocco, Butterworth & Garrett, 1996, for Spanish, and Vigliocco, Hartsuiker, Jarema & Kolk, 1996, for French and Dutch). In these studies, syntactically singular head nouns referring to a multiplicity of objects (where there is a mismatch between syntactic and conceptual information) were compared to syntactically singular head nouns referring to single objects (where there is a match between syntactic and conceptual information). In those experiments, more errors occurred when conceptual and syntactic number mismatched than when they matched.

Here, conceptual effects on subject-predicate gender agreement are reported. In the present studies, the contrast lies between cases in which the conceptual information matched the syntactic gender and cases in which there was no

conceptual information, but only the lexically specified syntactic feature. When the conceptual features mismatched the syntactic features (as in the case of subject-verb agreement in number) agreement processes were hindered; when the conceptual features matched the syntactic features (as in the case of gender agreement between the subject and the predicate), agreement processes were facilitated by the congruent conceptual information. Both findings support the maximalist view presented here. Together, they show how a syntactic operation such as agreement can be disrupted or helped by conceptual correlates. They also show that the role of conceptual factors is not limited to a specific agreement relation and a specific agreement feature.

One may argue that if the grammatical encoder is sensitive to conceptual factors, then an effect of animacy should also be found. After all, animacy plays an important role in grammatical function assignment (Bock, 1990). However, animacy may not be relevant to processes concerned with syntactic features such as gender (and number). Bock and Miller (1991) and Hupet, Fayol and Schelstraete (in press) have shown that animacy affects the grammatical encoder at the stage in which grammatical functions are assigned (i.e., animate entities tend to be subjects, while inanimate entities tend to be assigned to other functions). This stage, however, appears to precede agreement computation, as evidenced by speech errors (Bock & Levelt, 1994; Vigliocco & Nicol, 1998). In our experiments, there was no ambiguity with respect to grammatical function assignment: the adjectives participants were required to use in their completions were always better modifiers of the head noun than of the local noun. Therefore, no animacy effect would be predicted.

The second comparison concerns the effect of a mismatching local noun in pre-predicate position. The most common type of number agreement error reported in the literature is the case in which the verb agrees with the noun phrase that immediately precedes it. In order to understand how the local noun effect comes about, a number of different properties of the modifying noun have been manipulated. Bock and Miller (1991) and Bock and Eberhard (1993) investigated whether the presence of subject-like semantic features in the local noun influenced attraction errors in English. They found that neither the animacy nor the conceptual number of the local noun affected the error pattern. Vigliocco and Nicol (1998) showed that the linear proximity between the local noun and the verb was not the relevant factor in inducing errors, but rather that it was the hierarchical relation between the subject and the local noun phrase.

In the present studies, we showed an effect of a local mismatching noun for gender agreement errors in both French and Italian. As in number agreement, this effect may be the result of interference by the gender feature of the local noun on the gender feature of the head noun during agreement processing. That is, assuming that agreement is computed during the construction of a hierarchical

frame for the sentence, the gender feature of the local noun would be erroneously taken as the gender feature of the subject.

Finally, most studies of number agreement have shown an asymmetrical distribution of errors with more errors when the subject head noun is singular than when it is plural (Bock & Miller, 1991; Bock & Eberhard, 1993; Vigliocco et al., 1995; Vigliocco, Butterworth & Garrett, 1996). This asymmetry has been accounted for in terms of the markedness of the head and local nouns. In particular, Bock and Cutting (1992) and Eberhard (1997) have proposed that the singular form is unmarked (the default form), while the plural is marked with respect to number (see Tiesma, 1982). A marked feature is more likely to interfere with correct agreement than an unmarked feature.

In our studies we found an apparent cross-linguistic difference. In Italian, errors were equally frequent for masculine and feminine head nouns. In French, errors were generally more common for feminine than for masculine head nouns. Although this result requires replication because of methodological differences between the experiments in the two languages, and because the asymmetrical distribution in French was only significant in Experiment 2, it may relate to differences between the two test languages. Italian and French differ in how gender is morpho-phonologically realized. In Italian, the feminine form of an adjective (e.g., "delicato" (delicate-M)) is obtained by changing the inflectional morpheme (e.g., "delicata" (delicate-F)). In French, the feminine form of an adjective (e.g., "delicat" (delicat-M)) is obtained by adding a morpheme to the masculine form (e.g., "delicate" (delicate-F)). Therefore, the asymmetry in French and the lack of asymmetry in Italian may reflect the fact that while erroneously producing the masculine or the feminine always implies changing the inflection in Italian, erroneously producing the masculine in French implies omitting the inflection, while producing the feminine implies adding an inflection. Adding an inflection might be computationally more expensive than omitting it. If we compare the findings for number agreement errors and the findings reported here for gender agreement errors in Experiment 2 (French), a puzzle arises. As discussed above, the asymmetrical distribution of errors in number agreement reported in a variety of languages has been accounted for by assuming that the unmarked form of the subject noun is the most affected from interference by a local (marked) noun. The same reasoning does not apply for gender. In Italian, both forms seem to produce interference. In French, the marked form of the head noun is most affected by the interference caused by an unmarked local noun. We suggest that the apparent asymmetry between masculine and feminine we found that in French can be described as a tendency to use the unmarked form of the adjective. It is not related to the properties of nouns, presumably because the majority of errors were observed for nouns with grammatical gender, for which there may not be a marked/unmarked dichotomy.

2.9.2. Relevance of the Present Studies for Theories of Sentence Production

According to the maximal input hypothesis, the grammatical encoder takes all of the information available from conceptual structures. Regarding the distinction between conceptual and grammatical gender, this hypothesis predicts that if conceptual information is redundant with the syntactic information (conceptual gender), then errors should be less common than when there is no such conceptual information (grammatical gender). Therefore, conceptual information on gender should enhance correct agreement. This is what we found in Experiments 1, 2 and 4. In the investigations of subject-verb agreement in number, on the other hand, conceptual information on number hampered correct agreement, because the conceptual number conflicted with the syntactic number.

What are the general implications of the hypothesis of maximal input for grammatical encoding? In the language production system, non-linguistic representations exert a fine control over grammatical encoding. According to this view, accuracy in production cannot be defended on the basis of insulation from non syntactic information, but can be explained in terms of having available information that most often is congruent from different sources (a "maximal input view"). This position is similar in nature to "constraint satisfaction accounts" put forward in the literature concerning sentence comprehension (e.g., Trueswell, Tanenhaus & Garnsey, 1994; McDonald et al., 1994). However, the accounts developed in the comprehension domain assume that all kinds of information could influence syntactic processing, including morphological and phonological information. For sentence production, while we believe there is evidence for a fine-grained control by conceptual structures on grammatical encoding, the evidence for effects of morphological and phonological factors is scant at present (but see Vigliocco et al., 1995).

Can we be more precise with respect to the mechanisms responsible for computing agreement? The results presented here are compatible with two different mechanisms, both allowing conceptual input. A first view, proposed by Vigliocco et al. (1995) for number agreement, is that the grammatical encoder retrieves agreeing features from conceptual representations for the different agreeing elements. For the case of number agreement between subject and verb, conceptual information about the number of participants would be independently retrieved for both the noun and the verb. This information, being redundant most of the time, contributes to accurate and efficient production. Independent retrieval of agreement features for the different agreeing elements would ensure that in case features are lost from one of the elements, they would still be available to the encoder on the other element, ensuring accuracy. Furthermore, from a processing standpoint, independent retrieval of agreement features may be important in order to allow simultaneous encoding of the different constituents. Simultaneous

encoding contributes to efficient production, since it implies that the encoder does not need to wait for a constituent to be fully encoded before encoding another constituent. Evidence for simultaneous encoding of phrases and clauses comes from speech errors (e.g., Garrett, 1976), showing that during grammatical encoding parallel elements in separate phrases and clauses can interact with each other, resulting, for example, in exchanges such as (Ex. 30).

Ex. 30 : (...) **examine the horse of the eyes (intended: the eyes of the horse)**
(Garrett, 1976, p. 136)

For the specific case of gender agreement, this view implies that when gender reflects the sex of the referent, the encoding of the sentence is, in principle, both more accurate (as our studies have shown) and also more efficient. In this case, the sex of the referent provides features for the different agreeing elements: the likelihood of losing the gender features is low (since the same feature is specified on both noun and adjective) and the noun phrase as well as the adjective can be fully encoded simultaneously. When the gender of the noun does not have conceptual connotations, however, there is more chance of losing the gender features (solely represented on the noun), and simultaneous assembly of the noun and adjective phrase would not be possible, since information from the noun phrase is necessary for the specification of the adjective phrase.

Alternatively, agreement could be computed only on the basis of the features of the noun. Nouns with conceptual gender are specified for gender both at the conceptual and at the syntactic level. Nouns with grammatical gender are specified only at the syntactic level. Therefore, nouns with conceptual gender would have a stronger gender marking than nouns with grammatical gender. The stronger the marking, the less likely it is to be lost, causing an error. We believe it is not possible to decide between these two alternatives solely on the basis of the present data. However, there are two observations that are not easily integrated with the "strength" view. First, the strength view predicts that a noun with "stronger" features should produce more interference as a local noun. If it is assumed, as discussed above, that the effect of a mismatching local noun comes about because the features of the local noun are erroneously taken for agreement computation, then the strength of the local noun's features should matter. Preambles with local mismatching nouns with conceptual gender (e.g., *Le travail de la couturière*, the work of the dressmaker) should induce more errors than mismatching local nouns with grammatical gender (e.g., *La couleur de la robe*, the color of the dress). The gender type of the local noun was not experimentally manipulated in our experiments, and a post-hoc test was possible in only a single experiment. However, no difference was found, suggesting that preambles with local nouns with conceptual gender may not show stronger interference than preambles with local nouns with grammatical gender.

Studies of number agreement errors mentioned earlier (see Vigliocco, Hartsuiker, Jarema & Kolk, 1996 for a review) are also incompatible with the "strength" idea. In those studies, the conceptual numerosity came about, not as a feature of the head noun, but as a feature of the complex subject noun phrase. That is, conceptual number was manipulated by contrasting sentential preambles with a preferred single token (example 11) or multiple token (example 12) interpretation. For single token items (e.g., "The baby on the blankets") there is only one baby, sitting on a number of blankets. However, for multiple token items (e.g., "The label on the bottles") there are multiple labels, each of which is attached to a separate bottle. In these preambles, the conceptual number is strictly a property of the complex noun phrase including the head and the local noun. Finding an effect (at least in French and Italian) of conceptual number first motivated the hypothesis of independent retrieval of features for the different agreeing elements (see Vigliocco, Butterworth & Garrett, 1996, for a detailed discussion). The strength hypothesis, focusing on features strictly related to a single lexical item, is silent with respect to the results concerning number agreement. We do not see any logical reason why conceptual effects should have a different mechanism for number and gender agreement; therefore, we favor an interpretation of the present findings in terms of independent retrieval of gender features for the two agreeing elements in the case of conceptual gender.

Throughout the paper, we have assumed that conceptual and syntactic correlates of gender are separate kinds of information in production. However, some have argued against this distinction. For example, Tawmoski-de Ryck and Verluyten (1982) claim that information about the grammatical gender of nouns is used together with information about the content in the selection of appropriate referring expressions. These authors point out that the gender of the noun that would naturally be used to refer to something determines the form of a pronominal reference when there is no pronominal antecedent (i.e., in deictic pronominal use). For example, in talking about something that could be named "table" (la-F table-F, in French), it is appropriate to say (Ex. 31).

Ex. 31 : Tu n'arriveras jamais à la faire entrer dans la voiture

You'll never manage to get it-F into the car (Tawmoski-de Ryck & Verluyten, 1982, p. 328)

However, for "a desk" (le-M bureau-M), the masculine pronoun would be used. These observations are taken to indicate that grammatical gender and semantic representations are intimately connected in determining the selection of pronouns (hence, agreement between nouns and pronouns) in language production (see Garnham, Oakhill, Ehrlich & Carreiras, 1995, for a similar view in comprehension). According to this view, both grammatical and conceptual gender are semantic/conceptual in nature. It would be difficult for such a position to explain why nouns with grammatical and conceptual gender induce different numbers of agreement errors. A distinction between conceptual and grammatical

gender is also supported in other investigations concerning subject-pronoun agreement (see Cacciari, Carreiras, & Barbolini-Cionini, 1997, for comprehension, and Meyer & Bock, 1998, for production).

2.10 Conclusion

We have addressed issues concerning the type of input the grammatical encoder receives from conceptual structures. We used errors in the agreement of gender between the subject of a sentence and a predicative adjective to test predictions derived from assuming minimal or maximal input to grammatical encoding. Results from our studies support a maximalist view according to which non-syntactic information, such as conceptual information concerning the sex of the referent, is used in computing gender agreement between the subject and a predicative adjective. Our findings suggest that syntactic features reflecting conceptual features are retrieved for the different sentential elements. This redundant information ensures accuracy and allows for a more efficient encoding of the sentence.

Chapitre 3

When Sex hits Syntax: Syntactic accuracy in language production

Sentences are produced by speakers on an average of 3-4 words per second. Beyond fulfilling communicative intentions, sentences have to be accurate at the level of syntax and phonological form (among others). Given the speed with which we produce speech, it is remarkable that errors in production are as rare as about one out of 1,000 words. This paper addresses the question of how this level of accuracy may be achieved, focusing specifically on syntactic accuracy. We contrast two broad hypotheses. In one view (the minimalist hypothesis), accuracy is ensured by limiting the possible interference on syntactic processes of other types of information. The alternative view (the maximalist hypothesis) is that accuracy is achieved using all information available (syntactic and non-syntactic). We present parallel experiments in Italian and French, assessing errors in gender agreement between a noun and an adjective, whose results falsify a minimalist view. By showing that conceptual information is used to ensure syntactic accuracy, we provide evidence against a strong division of labour between conceptual and syntactic processes in language production.

3.1 Introduction

Universally, languages use different words to refer to certain humans of different sex, for example man and woman; father and mother; husband and wife. According to typological classifications (Corbett, 1991), the languages of the world tend to represent gender in the linguistic form either as a semantic property or as a formal property. English is an example of a language with an almost pure semantic gender system in which gender marking is used for animate (mainly human) referents and in which there is a transparent relation between the sex of a referent and the gender of nouns and pronouns (with a few exceptions such as ship referred to as «she»). Romance languages are examples of formal gender systems (Corbett, 1991). All nouns are marked for gender (masculine or feminine). While there is a transparent relation between the gender of a noun and the sex of a referent for certain nouns referring to humans and some animals, there is no conceptual basis for gender distinctions for nouns referring to objects and abstract entities, as well as for a number of nouns referring to animate entities (Harris & Vincent, 1990).

Regardless of whether the gender of a noun is a semantic or a formal property, it needs to be readily available during language production, because this information is necessary for computing dependencies (agreement) between a noun and other words in a sentence, such as determiners, adjectives, past participles and pronouns. The rule for establishing agreement, for example, between a noun and a predicative adjective in languages such as Italian and French, is very simple: if the noun is feminine, the adjective must be feminine. If the noun is masculine, the adjective must be masculine. This rule is the same for cases in which the gender of the noun is based on the sex of the referent, as in *signore*, *monsieur* [gentleman], *signora*, *madame* [lady], and when the gender of the noun is only a formal linguistic property without a transparent (or any) relation to the sex of a referent, for example *vittima*, *vittime* [victim-fem].

Given that gender agreement follows the same rule when the gender of the noun is conceptually motivated and when it is not, and given that agreement can be considered as a paradigmatic example of syntactic dependency (Bock, 1995), it provides us with an excellent tool for assessing central issues in language production. At a general level, it allows us to gain insight into how accuracy in performance is achieved. Speakers are remarkably accurate in their production. Bock (1991) estimated that slips of the tongue occur in speech approximately every 1,000 words, with errors in realizing the syntactic form being even more rare (in a corpus of approximately 15,000 sentences, Deese (1984) found only 77 structurally unacceptable utterances). This level of accuracy is outstanding if we consider that on average speakers produce 2-3 words per second (Levelt, 1989) and therefore there are ample opportunities for derailment.

How is this level of syntactic accuracy achieved? It is generally assumed that producing a sentence entails multiple levels of processing (Bock, 1982; Garrett, 1976; Levelt, 1989). Syntactic processes operate at the level of grammatical encoding, during which stage abstract lexical representations for words would be retrieved, and a hierarchically organized structure for the sentence first worked out and then linearized (see Vigliocco & Nicol, 1998 for a descriptions). Gender agreement, being a syntactic operation, is encoded during this stage. A subsequent phonological encoding stage determines the sound structure and the pronunciation codes for the sentence.

Within this general framework, syntactic accuracy could be ensured by insulating the operations of the grammatical encoder from potentially interfering information from other levels.

Although the processes at this level receive conceptual information as input; they are assumed to be autonomous from conceptual influences in their operation (see for e.g., Garrett, 1999). In the case of gender, conceptual information about the sex of a referent would be used in order to establish the gender of the noun (e.g., for cases such as *amico*, *ami* [friend-masc]/*amica*, *amie* [friend-fem]), or to

establish the word form to be used (e.g., for cases such as *marito*, *mari* [husband]/*moglie*, *femme* [wife]). Agreement with a predicative adjective, then, is realized as a purely syntactic operation, the same whether the gender of the noun is conceptually motivated or is just a formal property of the noun (and therefore is oblivious to the conceptual information). We have labeled this view the minimalist view in previous work (Vigliocco and Franck, 1999). With respect to gender agreement this view is clearly parsimonious (i.e., it assumes that the same processes apply for nouns with and without conceptual connotations of sex). This position is adopted by models of sentence production such as the one developed by Garrett (1976, 1999) and Levelt (1989) that embrace the principles of modularity of mind (Fodor, 1983).

A second possibility, that we have labeled the maximalist view (Vigliocco & Franck, 1999), assumes that accuracy is achieved by using information from other levels. The use of additional information would protect accuracy providing additional sources of information, available in case the primary information is lost. With respect to gender agreement, it assumes that accuracy is achieved by the use of conceptual information beyond establishing the gender of a noun, or the word form. Since this information, when present, is usually congruent with the syntactic information, it provides an additional source, in case the syntactic information is lost. In this view conceptual information penetrates the agreement process, giving an advantage when there is a congruent mapping between the sex of the referent and the gender of the noun, but also creating difficulties when there is an incongruent mapping between the two types of information. This view is similar in spirit to other proposals that have been put forward to account for the influence of non-syntactic factors in sentence comprehension, such as constraint satisfaction views (e.g., McDonald, Pearlmutter & Seidenberg, 1994; Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995) and the competition model (Bates & McWhinney, 1989) developed to account for cross-linguistic differences in the use of linguistic and conceptual cues in language processing.

In previous studies (Vigliocco & Franck, 1999), we have assessed whether the presence of congruent conceptual information is advantageous compared to the situation in which only the syntactic information is available. Consider for example the sentences in (Ex. 32) and (Ex. 33) in Italian.

Ex. 32 : **La gatta sul divano e' nera**
The-fem cat-fem on the couch is black-fem

Ex. 33 : **La giacca sul divano e' nera**
The-fem jacket-fem on the couch is black-fem

In both sentences, the predicative adjective (black) has to agree in gender with the subject head noun. In (Ex. 32) the gender of the noun *gatta* is assigned on the basis of the speaker's intention (i.e., if the speaker wants to talk about a female cat, then the feminine form *gatta* is to be used; if the speaker wants to talk about a male

cat, then the masculine form *gatto* is to be used). In (Ex. 33), the gender of the noun is a lexical property (i.e., it is not conceptually motivated). The question concerns whether the sex of the referent will be used beyond deciding the gender of the noun. If conceptual information is used in agreement, errors such as (Ex. 34) should be less common than errors such as (Ex. 35). This is what we found in both Italian and French.

Ex. 34 : **La gatta sul divano e' nero*

The-fem cat-fem on the-masc couch-masc is black-masc

Ex. 35 : **La giacca sul divano e' nero*

The-fem jacket-fem on the-masc couch-masc is black-masc

Therefore, these previous studies lend some support to the maximalist hypothesis by demonstrating that having congruent conceptual information is better than not having conceptual support. Vigliocco and Zilli (in press) extended the investigation to include two language-impaired speakers of Italian whose main difficulty concerned the production of syntactic structures and markers (i.e., agrammatic patients). The same materials used in Vigliocco and Franck (1999) were presented to the aphasic speakers. It was found that although the patients produced significantly more errors than age and education matched non-language impaired speakers for sentences in which the noun had grammatical gender, they did not differ from the non-language impaired speakers for nouns with conceptual gender. This result was also interpreted as indicating that the redundant and congruent conceptual information is used in the encoding by the patients.

In these previous studies the effect of conceptual information was assessed by contrasting cases in which the gender was assigned on the basis of the speaker's intentions (i.e., nouns with conceptual gender) to cases in which the gender was a lexical property of the noun (i.e., nouns with grammatical gender). However, the crucial test of the maximalist hypothesis requires showing that the effect of conceptual information is general, and therefore is not limited to a class of nouns (i.e., the nouns with conceptual gender) and, more strongly, to show that the direction of the conceptual effect differs depending upon whether the conceptual information is congruent or incongruent with the syntactic information. That is, to show that when conceptual and syntactic information are incongruent, accuracy is jeopardized. Italian and French provide us with an excellent opportunity of performing this crucial test by investigating gender agreement between the subject of the sentence and a predicative adjective for subject nouns such as *vittima*, *vittime* [victim-fem] or *prodigio*, *prodige* [prodigy-masc]. These nouns, referred to as "epicenes" (Marcantonio & Pretto, 1988), are functionally ambiguous in that they have a fixed grammatical gender but can refer to either a male or female entity. Therefore, epicene words allow two gender assignment possibilities: one based on the syntactic marker; the other based on the sex of the referent, which may or may not be the same as the syntactic gender. For these words, agreement between a noun, subject of a sentence and a predicative adjective is with the

grammatical gender of the noun, regardless of whether it refers to a male or female entity. By introducing a male or female actor co-referring to such nouns we can contrast cases in which the conceptual information is congruent with the gender of the noun to cases in which the conceptual information is incongruent. According to the minimalist view, agreement errors should be similarly common in these two different cases. Because the maximalist view predicts that errors are modulated by the conceptual information, they should be more common when the conceptual information is incongruent with the gender of the noun.

We report four experiments (two in Italian and two in French). In all of the experiments speakers were required to repeat and complete a sentence fragment with a (provided) predicative adjective. In the first two experiments (one in Italian and one in French) speakers were first presented with a sentential context as in (Ex. 36), which they were asked to read silently.

Ex. 36 : **Un camion ha investito Fabio/Fabiola che correva in bicicletta ascoltando musica**
A truck hit Fabio/Fabiola who was riding the bike while listening to music

Un camion a percute Fabien/Fabienne qui roulait à vélo et ne l'avait pas entendu
A truck hit Fabien/Fabienne who was riding a bike and had not heard it

Then they were presented with a sentence fragment such as (6) to complete.

Ex. 37 : **La vittima dello scontro**
La victime de l'accident
The-fem victim-fem of-the-masc crash-masc

In (Ex. 37) the noun *vittima*, *victime* refers back to a man (Fabio/Fabien) or to a woman (Fabiola/Fabienne). The sentential context introduces a main actor whose sex is congruent or incongruent with the gender of the noun (in the example above, Fabiola/Fabienne represents a case of congruity; Fabio/Fabien represents a case of incongruity). If the discourse information concerning the sex of the referent is taken into account in the encoding of subject-predicative adjective agreement, as it is predicted by the maximalist view, errors in the agreement of gender should be more common in the incongruent than in the congruent condition. Note that in contrast to previous studies (Vigliocco & Franck, 1999; Vigliocco & Zilli, in press) the conceptual information here is a property of the discourse model and not a property of the noun itself. Therefore, an effect of the sex of the referent cannot be accounted for in terms of differences in the assignment of gender to the nouns (the same nouns are used in both conditions). In fact, the use of epicenes allows us to assess the role of conceptual information within items, since the same target noun (e.g., *vittima*) is used with a sex congruent and a sex incongruent referent.

Experiments 3 and 4 are follow-up experiments in which we used the same sentence fragments but no introductory sentential context. The participants' task was the same as in Experiments 1 and 2. In these experiments, speakers completed the sentences without having conceptual information about the sex of the referent.

By assessing the impact of conceptual information in both Italian and French we ensure that any effect found is not an idiosyncratic property of just one language.

3.2 Experiments 1(Italian) and 2 (French)

In Experiments 1 (Italian) and 2 (French) we manipulated the congruency between the sex of a referent introduced in a context sentence and the gender of the noun with which the adjective had to agree. In both experiments, the sentence fragments were composed of a subject noun phrase (embedding a subject head noun which was either masculine or feminine and always referred to humans) and by a modifying noun phrase referring to an inanimate entity (embedding a local noun which was also either masculine or feminine) (see Table 1 (Tableau 9) below). Introducing a local noun with either same or different gender than the head noun in the experiments allows us to combine a conceptual and a syntactic manipulation in the same study. A higher error rate for fragments in which the local noun had different gender (mismatch) than for fragments in which it had the same gender (match) is a general finding that has been reported previously (Vigliocco & Franck, 1999). A higher incidence of errors for number mismatching local nouns is also a common finding in studies that have investigated number agreement between subject and verb (e.g., Bock & Miller, 1991; Bock & Eberhard, 1993; Vigliocco, et al., 1995; Vigliocco, Butterworth & Garrett, 1996; Vigliocco & Franck, 1999). The preferred interpretation of the mismatch effect is in terms of interference between the syntactic information carried by the subject head noun and the syntactic information carried by the local noun during the construction of a phrasal structure (Bock and Cutting, 1992; Vigliocco & Nicol, 1998). By orthogonally combining the manipulation of the sex of the referent and the manipulation of the gender match/mismatch between the subject head and the local noun we can assess the relative contribution of these two factors.

3.2.1. Method

Participants. Sixty-four undergraduate students in the Department of Psychology, University of Trieste participated in Experiment 1 (Italian). Forty undergraduate students in the Department of Psychology, University of Louvain participated in Experiment 2 (French). All participants were native Italian and French speakers and they volunteered or received credits for their participation.

Materials. Materials for each experiment consisted of sentential contexts, introducing either a male or female actor; sentence preambles composed of a subject head noun and a local noun embedded in a prepositional phrase and adjectives to be used in the completions.

Manipulated variables were the sex of the referent introduced by the sentential context (male, female); the gender (masculine or feminine) of the subject head

noun and the gender (masculine or feminine) of the local noun. Head nouns were all epicenes (i.e., nouns referring to humans with fixed gender). Since for nouns of this type there may be biases in the language with respect to whether the noun is used more commonly to refer to a man or to a woman, we conducted a rating experiment with 16 additional native speakers of Italian and French, who did not participate in any other part of the study. In the rating task, speakers were asked to indicate on a seven-point scale whether the word referred more often to a male (3) or female entity (-3). 32 words were used in the experiment. In both languages, the masculine nouns used in the experiments were more heavily biased toward male referents ($M = 1.75$ for Italian; $M = .74$ for French), then the feminine nouns toward female referents ($M = .09$ for Italian; $M = .40$ for French). T-tests revealed that both were significant ($t(30) = 4.93$, $p < .0001$ for Italian; $t(30) = 2.87$, $p = .008$ for French).

Given that it was impossible to select only words that were judged as equally likely to refer to a man or to a woman because in both languages the number of epicenes is limited, the ratings obtained were used as a covariate in statistical analyses performed in the production experiments to assess the impact of stereotypical gender (Carreiras, Garnham, Oakhill & Cain, 1996). Each head noun was used with two different local nouns, one matching and one mismatching in gender. All local nouns were nouns with grammatical gender, referring to inanimate entities. All the nouns used in the experiments were singular. The adjectives used in the experiments were all plausible predicates of the subject head noun. Pairs of adjectives matched in number of syllables were used for each pair of preambles with the same head noun and different local noun. For each preamble, the two forms of the adjective (masculine and feminine) were presented one above the other. The position of each form (above or below) was counterbalanced across items. Finally, it is important to note here that Italian and French differ with respect to how predictive of a given gender the word ending of a word is. For approximately 80% of Italian nouns, the word ending predicts of the gender of the noun (words that end in *-o* are masculine, words that end in *-a* are feminine). The situation is quite different in French where there is not a systematic relation between word endings and gender (although certain endings predict a given gender better than others, Tucker, Lambert and Rigault, 1977). In Italian, all nouns used in the experiment were overtly marked for gender in the morpho-phonological form (i.e., they ended in *-o* for the masculine, and they ended in *-a* for the feminine). In French, it was impossible to systematically control or manipulate this variable given the relatively small number of epicenes that could be used in the experiments.

Sentential contexts were created that introduced a main actor and provided a plausible context for the experimental sentence preambles. A pre-test was conducted to ensure that the main actor introduced by the sentential context and the subject head noun used in the sentence preambles plausibly co-referred.

In Experiments 1 and 2, thirty-two experimental items and twenty-two filler items were used. In the fillers, all subject head nouns had conceptual gender (i.e., the noun unambiguously referred to either a male or female entity). In half it was masculine, and in half feminine. The sentential contexts for the filler items were created as for the experimental items. For the fillers, the sex of the actor introduced in the context and the gender of the head noun were always congruent. Four lists were created to represent the different experimental conditions. In each list, each condition was represented by four experimental items; filler items were the same across the lists. The organization of the filler and experimental items in each list was pseudo-random, with the only constraint that the list started with two filler items. Examples of experimental materials are reported in Table 1 (Tableau 9).

Procedure. Participants were tested individually. Materials were visually presented at the centre of a computer screen. Trials consisted of the self-paced presentation of the sentential context that participants were instructed to read silently. Next, a sentence preamble was presented for 900ms and finally, adjectives were presented for 500ms. The presentation times were the same as used in other experiments in Italian by Vigliocco, Butterworth and Garrett (1996) and were sufficiently long to allow full comprehension of the materials. The participants' task was to complete the sentence preamble using the adjective in their completions. After completing a sentence, participants moved to the next item by pressing the spacebar in the majority of trials (95%). However, in a small proportion of filler trials they were presented at this point with a question concerning the sentential context and were instructed to respond. For example, in the Italian experiment, a filler trial was as follows: Sentential Context: "Per imbrogliare sul valore del terreno da stimare, Matteo si e' fatto pagare la tangente" [In order to cheat on the value of the land, Matteo received a bribe]; Sentential Preamble: "Il consulente del notaio" [The assistant of the notary]; Adjective: "corrotto/corrotta" [corrupt] and finally Context Question: "Perche' Matteo ha preso la tangente?" [Why did Matteo receive a bribe?]. These questions were introduced to ensure that participants were paying attention to the sentential context. Completed sentences were recorded and then transcribed.

Scoring. Completed sentences were assigned to one of the following scoring categories: correct responses when the participant correctly repeated and completed the preamble; agreement errors when the participant correctly repeated the preamble but produced an adjective that did not agree with the subject; and miscellaneous responses when s/he failed to repeat the preamble, parts of it or the adjective.

Design and data analyses. Statistical analyses of agreement errors and miscellaneous responses (both with subjects and items as random factors) were performed. Experimental factors were context (congruent vs., incongruent); gender of the head noun (masculine vs., feminine) and gender of the local noun (masculine

vs. feminine). Context was manipulated within subjects and items. Gender of the head and gender of the local nouns were manipulated within subjects and between items. The ratings reflecting biases in the language were used as a covariate in the item analyses.

Tableau 9 (Table 1) : Example of Experimental Materials used in Experiments 1 (Italian) and 2 (French).

Italian	
Context	
Anche se ha solo sei anni Giovanni/Caterina dà concerti in tutto il mondo <i>Although (he/she) is only six, Giovanni/ Caterina performs concerts around the world</i>	Le località che Luigi/Luisa ci ha fatto vedere erano tutte molto suggestive <i>The places that Luigi/Luigia showed us were all very spectacular</i>
Sentence Preambles	
<u>Masculine Subject Noun, Masculine Local Noun</u>	<u>Feminine Subject Noun, Feminine Local Noun</u>
Il prodigio del conservatorio <i>The prodigy of the music school-masc</i>	La guida per la gita <i>The guide for the tour-fem</i>
<u>Masculine Subject Noun, Feminine Local Noun</u>	<u>Feminine Subject Noun, Masculine Local Noun</u>
Il prodigio dell' accademia <i>The prodigy of the-fem academy-fem</i>	La guida per il viaggio <i>The guide for the journey-masc</i>
French	
Context	
Les recherches de Roland/Rachel en biologie moléculaire comportent des enjeux considérables <i>The research of Roland/Rachel in molecular biology involves very important stakes</i>	Les locataires craignent Louis/Louise qui peut les mettre dehors quand bon lui semble <i>The tenants fear Louis/Louise who can force them to leave whenever he (she) wants to</i>
Sentence Preambles	
<u>Masculine Subject Noun, Masculine Local Noun</u>	<u>Feminine Subject Noun, Feminine Local Noun</u>
Le génie du département <i>The genius of the department-masc</i>	L'autorité de la maison <i>The authority of the house-fem</i>
<u>Masculine Subject Noun, Feminine Local Noun</u>	<u>Feminine Subject Noun, Masculine Local Noun</u>
Le génie de la faculté <i>The genius of the faculty-fem</i>	L'autorité de l'immeuble <i>The authority of the building-masc</i>

3.2.2. Results and discussion

In Italian, we obtained 1783 (87.0%) correct responses, 169 (8.3%) agreement errors and 96 (4.7%) miscellaneous responses. In French, we found 1084 (84.7%) correct responses, 180 (14.0%) agreement errors and 16 (1.3%) miscellaneous

responses. The distribution of agreement errors and miscellaneous responses in the different experimental conditions is reported in Tables 2 (Italian, Tableau 10) and 3 (French, Tableau 11).

Tableau 10 (Table 2) : Distribution of agreement errors and miscellaneous responses (% in parenthesis) in the different experimental conditions in Experiment 1 (Italian).

	Context Congruent		Context Incongruent	
	Masculine Head Noun	Feminine Head Noun	Masculine Head Noun	Feminine Head Noun
Agreement errors				
Masculine Local Noun	0 (0)	13 (5)	23 (9)	40 (15.6)
Feminine Local Noun	6 (2.3)	5 (2)	31 (15.2)	49 (19.1)
Miscellaneous Responses				
Masculine Local Noun	10 (4)	24 (9)	11 (4)	7 (3)
Feminine Local Noun	10 (4)	11 (4)	6 (2)	17 (7)

Tableau 11 (Table 3) : Distribution of agreement errors and miscellaneous responses (% in parenthesis) in the different experimental conditions in Experiment 2 (French).

	Context Congruent		Context Incongruent	
	Masculine Head Noun	Feminine Head Noun	Masculine Head Noun	Feminine Head Noun
Agreement errors				
Masculine Local Noun	1 (0.6)	16 (10)	22 (13.8)	48 (30)
Feminine Local Noun	5 (3)	20 (12.5)	24 (15.6)	44 (27.5)
Miscellaneous Responses				
Masculine Local Noun	1 (0.6)	2 (1.2)	5 (3.1)	5 (3.1)
Feminine Local Noun	1 (0.6)	0 (0)	1 (0.6)	1 (0.6)

In both languages, agreement errors were most common for preambles introduced by an incongruent context. Errors were also more common for preambles with a feminine subject head noun. The main effect of context was significant in Italian ($F(1,63) = 41.47, p < .0001$; $F(1,59) = 20.15, p < .0001$) and in French ($F(1,39) = 31.11, p < .001$; $F(1,59) = 57.4, p < .0001$). The main

effect of gender of the head noun was also significant in both languages ($F(1,63) = 12.53, p = .001$; $F(1,59) = 18.3, p < .0001$, in Italian; and $F(1,39) = 24.1, p < .0001$; $F(1,59) = 20.14, p < .0001$, in French). The main effect of the bias was significant in Italian ($F(1,59) = 4.6, p = .041$) but not in French ($F < 1$). The effect of bias did not interact with context ($F < 1$ in both languages). No other main effects or interactions were significant in either language (all $F_s < 1$, except for the interaction between context and gender of the head noun: $F(1,63) = 1.41, p = .23$; $F(1,59) = 2.2, p = .13$; in Italian; $F(1,39) = 1.27, p = .31$; $F(1,59) = 1.8, p = .18$, in French).

Miscellaneous responses, in Italian, were most common for preambles with feminine head nouns and masculine local noun when the context was congruent. As for agreement errors, we also found an effect of the bias. The analysis of variance revealed a three way interaction ($F(1,63) = 8.2, p = .006$; $F(1,59) = 3.4, p = .07$). The main effect of bias was significant in this analysis ($F(1,59) = 5.7, p = .02$). All other main effects and interactions were not significant (all $F_s < 1$). In French, miscellaneous responses were somewhat more common for incongruent contexts, however, this factor did not reach a significant level ($F(1,39) = 2.1, p = .16$; $F(1,59) = 3.4, p = .07$), no other main effect or interaction reached significance (all $F_s < 2$; $p > .15$).

By manipulating the congruency between the sex of the referent and the gender of the noun, we obtained different error rates. Errors were most common when the sex of the referent and the gender of the noun were incongruent, this was the case in both languages. This finding lends support to the maximalist view indicating that both conceptual and syntactic information are taken into account in agreement, contrary to the prediction of the minimalist view. Together with our previous studies, these experiments indicate that conceptual information can help or hinder grammatical encoding. In the experiments by Vigliocco & Franck (1999) it was shown that when conceptual information was present and congruent, errors were less common than when it was not present. In the current experiments we showed that conceptual information hinders agreement when it is incongruent with the syntactic information.

There are three other interesting findings. First, the effect of gender mismatch between the head and local noun was not significant in either language. This finding is at odds with previous experiments in which an effect of a mismatching local noun was largely attested in different languages and in different forms of agreement relations (e.g., Bock & Miller, 1991; Bock, Nicol & Cutting, in press; Vigliocco, Butterworth & Semenza, 1995). However, none of the previous studies introduced a discourse context. In our studies, it seems that the conceptual information given by the context plays such a large role that it overshadows the potential impact of mismatching syntactic information. This possibility is further assessed in Experiments 3 and 4. Second, as in previous experiments by Vigliocco

& Franck (1999), errors in French (14.0%) are generally more common than in Italian (8.3%), a finding which may be related to the degree of overt marking of gender for nouns in the two languages. As mentioned above, while in French there is not a clear and transparent relationship between the gender of a noun and the word termination, the situation is very different in Italian. An hypothesis concerning how morpho-phonological form may affect grammatical encoding is presented in the general discussion. Third, in both languages we found an effect of the gender of the head noun, with more errors for feminine than masculine head nouns. In previous studies, such an effect was reported for French, but not Italian (Vigliocco & Franck, 1999). This finding will also be addressed in the general discussion.

3.3 Experiments 3 (Italian) and 4 (French)

The main purpose for conducting these two follow-up studies was not only to obtain some indication of error rates in the absence of a context, but also and foremost to assess whether the lack of a gender mismatch effect in the previous studies is due to the greater influence of conceptual over interfering syntactic information.

3.3.1. Method

Participants. Sixty-four undergraduate students in the Department of Psychology, University of Trieste participated in Experiment 3 (Italian); and 60 undergraduate students in the Department of Psychology, University of Louvain participated in Experiment 4 (French). They were all native Italian and French speakers and they volunteered or received credits for their participation. None of them took part in Experiments 1 and 2.

Materials. Experimental materials consisted of the same sentence preambles and adjectives used in Experiments 1 and 2. Since the sentential contexts were not used, only two lists were created combining the following manipulated variables: gender of the subject head noun (masculine, feminine); gender of the local noun (masculine, feminine). Each condition was represented by eight items.

Procedure. The procedure was the same as in the previous experiments, with the exception that no sentential context was presented. Therefore, participants were first presented with the sentence preambles and then the adjectives. As in the previous experiments, their task was to produce a full sentence aloud using the preamble and the adjective.

Scoring. The same criteria used in Experiments 1 and 2 were used in Experiments 3 and 4.

Design and data analysis. The same as used in the previous experiments, excluding the context factor.

3.3.2. Results and discussion

In Italian, we obtained 1922 (93.9%) correct responses, 68 (3.3%) agreement errors and 58 (2.8%) miscellaneous responses. In French we found 1713 (89.3%) correct responses, 141 (7.3%) agreement errors and 66 (3.4%) miscellaneous responses. Table 4 (Tableau 12) reports the distribution of agreement errors and miscellaneous responses in the different experimental conditions in both languages.

Tableau 12 (Table 4) : Distribution of agreement errors and miscellaneous responses (% in parenthesis) in the different experimental conditions in Experiments 3 (Italian) and 4 (French).

	Experiment 3 (Italian)		Experiment 4 (French)	
	Masculine Head Noun	Feminine Head Noun	Masculine Head Noun	Feminine Head Noun
Agreement errors				
Masculine Local Noun	3 (0.6)	32 (6.3)	2 (0.4)	73 (15.2)
Feminine Local Noun	20 (3.9)	13 (2.5)	14 (2.9)	52 (10.8)
Miscellaneous Responses				
Masculine Local Noun	14 (2.7)	15 (2.9)	15 (3.1)	19 (4)
Feminine Local Noun	13 (2.5)	14 (2.7)	19 (4)	13 (2.7)

In both languages, agreement errors were most common when the subject head noun and the local noun mismatched in gender. As in the previous experiments, errors were also more common for feminine than masculine head nouns.

In the analyses, the main effect of the gender of the head noun was significant in Italian and French ($F(1,63) = 6.5, p = .013$; $F(1,59) = 14.5, p = .001$ in Italian and $F(1,59) = 82.2, p < .0001$; $F(1,59) = 39.5, p < .0001$, in French). The interaction between the gender of the head noun and the gender of the local noun was also significant ($F(1,63) = 14, p < .0001$; $F(1,59) = 13.4, p = .001$ in Italian; and $F(1,59) = 7.3, p = .009$; $F(1,59) = 3.4, p = .07$, in French). Bias was significant in both languages (Italian: $F(1,59) = 8.3, p = .006$; French: $F(1,59) = 4.3, p = .042$).

In the analysis of miscellaneous responses, no main effect or interaction was significant both in Italian and in French (all $F_s < 2$; $p > .15$).

A main finding from these two experiments is the gender mismatch effect. When no biasing context was presented, errors were modulated by the presence of matching or mismatching gender information in the subject noun phrase. This finding is consistent with previous studies regarding number and gender agreement

(e.g., Bock, 1995; Vigliocco & Franck, 1999) and is consistent with the hypothesis that the lack of such an effect in Experiments 1 and 2 might be related to the fact that the conceptual factor overshadowed the effect of the local noun.

Overall, errors in these two experiments were less common than in Experiments 1 and 2 when the context was incongruent with the gender of the subject head noun (3.3% vs. 14.7%, in Italian and 7.1% vs. 21.7% in French). Errors in Experiments 3 and 4 were also somewhat more common than in Experiments 1 and 2 when the context was congruent with the gender of the head noun (2.3% vs. 3.3% in Italian; and 6.5% vs. 7.3% in French). However, a direct comparison among the different conditions is not warranted because in the no context experiments we cannot exclude the possibility that speakers developed a discourse model in which the referent was either a man or a woman depending on the grammatical gender of the noun or the speakers' idiosyncratic biases.

As in the previous experiments, agreement errors were overall more common in French (7.3%) than in Italian (3.3%), and also more common for feminine than masculine nouns. These two findings will be discussed below.

3.4 General discussion

The main result from this series of studies is finding that syntactic accuracy (operationalized as agreement accuracy) was modulated by the type of conceptual information regarding the sex of a referent. The results have been replicated in two languages, indicating that the findings generalize across Romance languages. These results falsify a minimalist view, according to which conceptual information does not affect syntactic operations. Before discussing the implications for models of language production, we address the other findings from these experiments.

First, a basic finding consistently reported in both naturally occurring and experimentally induced agreement errors (both agreement in gender and in number) is the effect of a mismatching local noun (e.g., Bock, 1995). While we replicated previous results in Experiments 3 and 4, we did not in Experiments 1 and 2. These two experiments differ from all other studies in the literature in that a sentential context was introduced, providing a more detailed discourse model for the sentence preamble. This more detailed discourse model may have increased the weight of the conceptual information over the syntactic information shadowing any syntactic interference from the gender mismatching local noun in the experimental situation.

Second, in both series of experiments, errors were systematically more common in French than in Italian. A similar finding was reported by Vigliocco & Franck (1999). Although the present experiments are parallel in terms of procedures, the materials used for each of the two languages are not translation equivalents and therefore we cannot completely exclude that the different error

rates may be related to differences in materials and other, uncontrolled, experimental differences. However, a plausible alternative explanation for the difference is in terms of a maximalist view in which not only conceptual, but also morpho-phonological information, when available, is used to ensure accuracy. While gender is marked in a systematic manner in the word endings in Italian; the relation between gender and word ending is far from systematic in French. In particular for our experiments, while morpho-phonological cues of gender were present in Italian; they were not in French. That is, while in all the experiments in Italian the word ending used were highly predictive of the gender of the word, this factor was not controlled in French. Compatible with this hypothesis, in another series of studies, Vigliocco, Hartsuiker, and Franck (1999) found that error rates in both languages were influenced by the manipulation of morpho-phonological predictability and that in both languages, when morpho-phonological cues to gender were maximized, error rates were similar.

Finally, we found an asymmetrical distribution of errors for subject head nouns with masculine and feminine gender. Whereas in French this result replicates previous studies (Vigliocco and Franck, 1999), such an effect has not previously been reported for Italian. The higher error rate for feminine nouns can be explained as a tendency to overuse the masculine form of adjectives, which can be considered as the "unmarked" form (Tiesma, 1982), since it is used to refer to a mixed groups of men and women, when the sex of the referent is unknown, and as the form of past participles). The tendency to overuse the unmarked masculine form of the adjectives may have been reinforced by the fact that the epicenes used in the experiments were generally biased toward a masculine stereotypical gender (and therefore masculine words were more strongly biased toward a male referent than feminine words toward a female referent).

The effect of context on agreement accuracy has important implications for models of language production. In previous studies we reported that conceptual information enhanced accuracy in agreement construction when it was present and congruent (Vigliocco and Franck, 1999). Here we show that conceptual information hinders agreement construction when it is incongruent. These results help explain how syntactic accuracy is achieved. The use of additional and unnecessary conceptual information may protect from information loss, or more generally strengthen syntactic information (contrary to the idea of a strict separation of labor between conceptual and syntactic encoding to ensure that unnecessary information will not interfere).

This seems plausible since, in general, conceptual and syntactic information are congruent, and incongruency can be considered more as an exception in Romance languages, and probably in Indo-European languages, more generally. For gender, when the referent is human, in the vast majority of cases the gender of the noun is determined on the basis of the sex of the referent (more than 90% in

our estimate for Italian and French). Words that refer to either sex are rare (and the words used in the present experiments are an almost exhaustive list); and to our knowledge, the number of cases in which the gender of a noun and the sex of a referent are always conflicting is extremely small (two such cases In Italian: *sentinella* [sentry-fem]; *vedetta*, [lookout-fem]). Furthermore, it is interesting to note that as professions previously considered typically "for men" have become more popular among women, these languages changed (and are changing) the gender of the nouns accordingly, creating new feminine equivalents of masculine nouns (e.g., *senatore/senatrice*, *sénateur/sénatrice* [senator, masculine and feminine]).

To say that the language production system uses conceptual information when available to ensure accuracy is not to say that only conceptual information is used. If this were the case, gender agreement would always be with the sex of the referent regardless of the gender of the noun. This is clearly not true. First, as mentioned above, for nouns such as *vittima*, subject-predictive adjective agreement has to be with the gender of the noun, regardless of the sex of the referent. Second, in Experiments 1 and 2, although we found a substantial number of errors when the sex of the referent and the gender of the noun were incongruent, in the majority of cases (84% in Italian and 77% French) agreement was with the grammatical gender of the nouns.

For the same class of words agreement between the subject and an anaphoric pronoun can be either syntactic or conceptual. That is, both sentences (Ex. 38) and (Ex. 39) are acceptable in Italian when referring to a man.

Ex. 38 : *La vittima sosteneva che lo avrebbero dovuto rimborsare...*

The-fem victim-fem maintained that he should-masc be reimbursed...

Ex. 39 : *La vittima sosteneva che la avrebbero dovuta rimborsare...*

The-fem victim-fem maintained that she should-fem be reimbursed...

This fact indicates how both types of information play a role in the encoding process. The observation that different types of agreement relationship may use syntactic and/or conceptual information has been described by Comrie (1975) and Corbett (1983) in terms of agreement hierarchy according to which the further an agreement controller (e.g., the subject) is from an agreement target (e.g., the predicative adjective, or the anaphoric pronoun), the more likely conceptual factors will control the agreement relationship (see Vigliocco, Butterworth & Garrett, 1996, for a discussion). In processing terms, these facts seem to indicate a strictly local domain (presumably clause internal) for syntactic control, and a broader domain (sentential and multi-sentential) for conceptual control. Evidence indicating that (at least for the production processes concerned with agreement) the clause may define the domain of syntactic control comes from studies by Bock and Cutting (1992) and Bock, Nicol & Cutting (1999). However, what is crucial is that even in the local domain of syntactic control, conceptual information is taken into

account. In a complementary manner, syntactic information is not completely neglected in the broader domain of conceptual control. Cacciari, Carreiras & Barbolini-Cionini (1997) have investigated the role of syntactic and conceptual factors in the comprehension of anaphoric relations involving epicenes in Italian. In their Experiment 3, they used materials such as (Ex. 38) and (Ex. 39) above and (Ex. 40) below. In (Ex. 40) the antecedent "amante" [lover] is not an epicene but what they refer to as an "ungendered" word, since lacking any morpho-phonological marking of syntactic gender, its gender is solely under conceptual control.

Ex. 40 : *L'amante sosteneva che la/lo avrebbero dovuta/dovuto rimborsare...*
The-0 lover-0 maintained that s/he should-fem/should-masc be reimbursed...

Ungendered words (Ex. 40) were used as a baseline to which to compare incongruent (Ex. 38) and congruent (Ex. 39) epicenes. In a series of experiments they found that while there was no processing cost associated with incongruency between the grammatical gender of the noun and the conceptual referent; there was a processing advantage for having a congruent relation. That is, while reading times did not significantly differ between (Ex. 38) and (Ex. 40); reading times in (Ex. 39) were significantly shorter than in (Ex. 38) and (Ex. 40). These results indicate that also when subject-anaphoric pronoun is considered, both types of information (syntactic and conceptual) are taken into account, although the weight carried by each of them is reversed for subject-predicative adjective and subject-anaphoric pronoun agreement.

According to the maximalist view, considering the different types of information, when available, helps accurate production, which, as discussed in the introduction, is one fundamental property of linguistic performance.

Are the conceptual effects we discussed generalizable, or are they limited to gender agreement in Italian and French? In previous studies concerning subject-verb agreement in number, conceptual effects, parallel to the conceptual effects reported in the current experiments have been shown for a number of languages (Bock, et al., 1999; Eberhard, 1999; Vigliocco, Butterworth & Garrett, 1996; Vigliocco, Butterworth & Semenza, 1995; Vigliocco, Hartisuiker, Jarema & Kolk, 1996), and using different manipulations. For example, Bock et al (1999) showed that plural verbs were erroneously more commonly used for collective nouns (e.g., jury) than for nouns referring to individuals (e.g., juror). Eberhard (1999) and Vigliocco and colleagues, instead showed that plural verbs were erroneously used more often for so called "distributed" noun phrases, such as (Ex. 42) than for non-distributed noun phrases, such as (Ex. 41).

Ex. 41 : *The baby on the blankets*

Ex. 42 : *The label on the bottles*

For non-distributed materials, the preferred interpretation implies only one instance of "baby" in the example, sitting on a number of blankets. For distributed

materials, instead, the preferred interpretation entails many labels, each of which is attached to a (different) bottle. If the number of participants in the conceptual scene (one baby, many labels) is taken into account during phrasal construction, errors in the agreement of number between the subject and the verb should be more common for (Ex. 42) than for (Ex. 41), and this is the finding reported in all these studies. Note that this contrast is particularly interesting because the conceptual number (plural) of participants comes about only when the entire complex noun phrase is considered and in this respect it contrasts with the grammatical number of the subject noun phrase (which is singular). Therefore, together these findings and the results reported here provide support for the maximalist view presented above according to which conceptual information is used during grammatical encoding beyond establishing syntactic features.

3.5 Conclusion

We have contrasted two general views concerning the interplay of syntactic and conceptual information during language production. These two views account in very different ways for a fundamental property of fluent production, that is its accuracy, and make different predictions with respect to whether or not non-syntactic factors can affect grammatical encoding. The experiments we conducted support the predictions of the maximalist view indicating that non-syntactic conceptual information can hinder grammatical encoding when the conceptual information is incongruent with the syntactic information, thus falsifying a strict division of labour between conceptual preparation and grammatical encoding.

Chapitre 4

Intermezzo : l'hypothèse lexicale

L'hypothèse maximale d'un contrôle conceptuel sur la réalisation de l'accord en genre a été avancée dans les Chapitres 2 et 3 pour rendre compte des effets observés en français et en italien. A l'issue de ces deux chapitres, nous suggérons un mécanisme explicatif de l'effet conceptuel selon lequel les différents segments de la phrase sont spécifiés directement par la représentation conceptuelle du sexe de l'expérienceur. Le présent chapitre rapporte une recherche menée en français par laquelle nous montrons qu'un nom local de genre conceptuel entraîne plus d'erreurs qu'un nom local de genre de surface. L'interprétation théorique d'un effet conceptuel du nom local sur l'accord est discutée. Nous suggérons qu'un tel effet plaide en faveur d'une représentation lexicale de la propriété sémantique de genre, et non en faveur de l'hypothèse notionnelle comme il est classiquement supposé dans la littérature. L'hypothèse lexicale est avancée comme alternative à l'hypothèse notionnelle, pouvant à la fois rendre compte de l'effet conceptuel sur le nom local et de celui sur le nom sujet. Ces nouveaux résultats nous amènent à reconsidérer l'hypothèse notionnelle et les arguments qui la soutiennent.

4.1 Introduction

Les études visant à déterminer l'influence de facteurs conceptuels sur l'accord peuvent être divisées en deux groupes : les unes ont manipulé l'information conceptuelle véhiculée par des propriétés du nom local (Bock & Miller, 1991 ; Bock & Eberhard, 1993 ; Hupet, Fayol & Schelstraete, 1998), les autres ont manipulé les propriétés conceptuelles du nom sujet (Bock, Eberhard & Cutting, 1992 ; Bock, Nicol & Cutting, 1999 ; Vigliocco & Franck, 1999, voir Chapitre 2 ; Vigliocco & Franck, submitted, voir Chapitre 3). Bien que de fortes différences soient apparues entre les recherches quant à l'effet de la distributivité du sujet, relatif à l'interprétation plurielle de l'ensemble du syntagme nominal¹¹, il semble qu'il y ait une relative homogénéité dans les résultats concernant les propriétés conceptuelles propres aux noms. Globalement, ces études ont mis en évidence soit une tendance, soit un effet significatif des propriétés conceptuelles étudiées. Dans l'Introduction générale, nous avons synthétisé ces résultats publiés dans la littérature. Nous les repassons brièvement en revue avant d'aborder la question de leur interprétation.

¹¹ Une perspective critique sur les effets de distributivité est avancée dans la Synthèse.

Les études menées en anglais manipulant les propriétés conceptuelles du nom local montrent une tendance des locuteurs à prendre en compte cette information, pour autant qu'elle s'accompagne de l'introduction d'une difficulté syntaxique. Ainsi, un nom local animé entraîne plus d'erreurs d'accord qu'un nom local inanimé, à condition que la structure syntaxique de la phrase soit suffisamment complexe (Bock & Miller, 1991). Par ailleurs, un nom local collectif a tendance à générer plus d'erreurs d'accord qu'un nom local individuel, en tout cas lorsqu'ils sont également grammaticalement pluriels et le sujet de la phrase singulier (Bock & Eberhard, 1993). Autrement dit, un effet conceptuel du nom local apparaît dans cette dernière recherche quand les locuteurs doivent déjà traiter une situation de mismatch entre le nombre grammatical du sujet et celui du nom local.

La recherche de Hupet, Fayol & Schelstraete (1998) indique que la production écrite de l'accord du verbe en français est, elle aussi, influencée par une caractéristique conceptuelle du nom local. Les auteurs ont rapporté significativement plus d'erreurs d'accord lorsque le nom local était un sujet sémantiquement plausible du verbe que lorsqu'il ne pouvait pas être sujet du verbe.

Les propriétés conceptuelles du nom sujet ont aussi été étudiées en anglais. Deux recherches ont mis en évidence un effet hautement significatif de la référence notionnelle du nom sujet : plus d'erreurs d'accord du pronom-tag (Bock, Eberhard & Cutting, 1992) et du pronom réflexif (Bock, Nicol & Cutting, 1999) ont été rapportées lorsque le nom sujet était collectif que lorsqu'il était individuel. Une tendance à produire plus de verbes pluriels après un nom sujet collectif singulier est également apparue dans la recherche de Bock et al. (1999), similaire à celle rapportée par Bock & Eberhard pour les noms locaux.

Dans le Chapitre 2, nous avons rapporté un effet significatif de la présence d'un corrélat conceptuel de genre sur le nom sujet, en français et en italien. Les locuteurs de ces deux langues ont produit plus d'erreurs d'accord lorsque le nom sujet avait un genre grammatical que lorsqu'il avait un genre conceptuel. Par ailleurs, des manipulations de contrôle ont confirmé que cet effet ne dépendait pas du caractère animé vs. inanimé des noms ayant un genre conceptuel vs. grammatical, puisque le même effet a été trouvé en comparant des noms animés ayant un genre conceptuel à des noms animés ayant un genre de surface. Les recherches présentées dans le Chapitre 3 ont, quant à elles, révélé qu'une discordance entre l'information grammaticale de genre portée par le nom sujet et l'information contextuelle relative au sexe du référent, constitue une source significative d'erreurs. Autrement dit, la référence contextuelle de genre du nom sujet s'est également avérée être un facteur conceptuel déterminant dans la réalisation de l'accord en français et en italien.

Toutes ces recherches, qu'elles manipulent les propriétés du nom local ou du nom sujet, visaient à déterminer la nature de l'information sur base de laquelle l'accord est construit. En manipulant les propriétés du nom local, comme l'ont fait

les premières études, les auteurs se proposaient d'étudier dans quelle mesure des propriétés conceptuelles peuvent *perturber* le processus syntaxique d'accord. Autrement dit, cette démarche examine la nature des informations, portées par des éléments qui ne sont pas impliqués dans la réalisation de l'accord, générant des interférences. Par contre, en manipulant les propriétés du nom sujet, l'opération d'accord est directement étudiée puisque ces propriétés sont celles *normalement* prises en compte par le processus. Les auteurs, d'un côté comme de l'autre, postulent que ces deux types de manipulations, du nom local et du nom sujet, mettent à l'épreuve la même hypothèse : celle d'un contrôle conceptuel sur l'opération syntaxique d'accord. Comme l'écrivent Bock & Eberhard (1993) à propos de la manipulation de la collectivité du nom local :

The assumption behind this work, as in other studies of speech errors, was that the factors responsible for the errors are closely related to the factors that affect normal, error-free language production. We are therefore supposing that when a local noun attracts agreement to itself, it does so because it possesses properties that normally control subject-verb number agreement. (Bock & Eberhard, 1993, p. 67)

Ce postulat, selon lequel un effet conceptuel du nom local 'signifie' la même chose qu'un tel effet du nom sujet, semble effectivement acceptable si l'hypothèse se limite à prédire un effet conceptuel sur l'occurrence des erreurs d'accord. Si toutefois l'on s'interroge sur le mécanisme par lequel cet effet a lieu, il nous semble que les implications théoriques des effets conceptuels sur le nom local doivent être différenciées de celles des effets conceptuels sur le nom sujet.

4.2 Interprétation de l'effet conceptuel

Dans le Chapitre 2, une description du mécanisme psycholinguistique pouvant rendre compte de l'effet conceptuel sur l'accord en genre est brièvement proposée. Le mécanisme, également avancé pour expliquer l'accord en nombre, repose sur l'hypothèse que l'ensemble des éléments qui codent le nombre ou le genre dans la phrase reflètent des propriétés notionnelles liées à la quantité ou au sexe des objets dont il est question. Selon cette hypothèse, tous ces éléments sont codés directement à partir de la représentation notionnelle ou conceptuelle du message. Prenons par exemple l'accord en nombre entre le sujet et le verbe. Si un locuteur veut exprimer l'idée que « les tomates rougissent », le concept sous-jacent à *tomates* est multiple, et cette multiplicité communique à l'ensemble du message une représentation multiple. Le concept sous-jacent à *rougissent* est, par conséquent, lui aussi multiple ; il reflète le fait que l'action est effectuée par une multiplicité d'agents. En ce qui concerne l'accord en genre, le principe est le même : dans l'idée « son mari est galant », l'expérienceur¹² est sexué et cette

¹² Par *expérienceur*, nous entendons l'argument porteur du cas *Experier*, i.e. [+Participatif] en tant que [+Affecté] (Godin, 1976, cité dans Costermans, 1980).

information se propage sur l'entièreté du message à exprimer qui devient notionnellement <de sexe masculin>. Autrement dit, l'ensemble de la structure conceptuelle est considérée comme une notion masculine ou féminine. L'adjectif véhicule donc lui aussi cette information puisqu'il baigne dans la notion de masculin conférée au message par l'expérimenteur. Le nom sujet et l'adjectif prédicatif reçoivent donc leur spécification linguistique de la même manière : par détermination de la structure conceptuelle.

Cette *hypothèse notionnelle*, originellement soutenue par des linguistes tels que Barlow (1993, voir aussi Pollard & Sag, 1994), a reçu l'attention de psycholinguistes qui l'ont implémentée dans des modèles de production de la phrase (Kempen & Hoenkamp, 1987 ; Kempen & Vosse, 1989 ; De Smedt, 1990). La perspective développée par Kempen & Hoenkamp se situe dans le cadre des modèles de production séquentiels décrits dans l'Introduction générale (ex. Garrett, Levelt). L'originalité de leur modèle (IPG, Incremental Procedural Grammar) est qu'il postule que l'encodage syntaxique procède par assemblage de segments. Chaque segment est formé de deux nœuds et d'un arc les reliant. Les nœuds représentent les catégories syntaxiques, les arcs représentent les fonctions syntaxiques (ex. S-sujet-SN, SN-tête-N). Lorsqu'un concept d'agent est activé, il sélectionne un lemme qui lui correspond et est ensuite placé dans un segment syntaxique contenant les propriétés syntaxiques du lemme (ex. nombre, genre). Ces paramètres syntaxiques sont transmis à l'entièreté du segment du syntagme nominal. Lorsque deux segments s'assemblent, par exemple le segment syntagme nominal et le segment syntagme verbal, les propriétés de nombre et de genre sont copiées sur ce dernier au niveau du nœud central P de l'arbre syntaxique.

Un tel modèle ne peut toutefois pas rendre compte des effets conceptuels observés puisque l'accord se fait par copiage de l'information lexico-syntaxique de la source sur la cible. Reprenant les mécanismes de base du modèle IPG, selon lesquels la production se fait par segments, De Smedt a mis au point un système de formulation du langage (IPF, Incremental Parallel Formulator¹³) dans lequel la réalisation de l'accord se fait par *unification* des segments syntaxiques (voir aussi Kempen & Vosse, 1989). L'hypothèse sous-jacente à cette procédure d'accord est que les segments à accorder sont tous les deux spécifiés directement à partir de la représentation conceptuelle de la phrase, et indépendamment l'un de l'autre, pour être ensuite unifiés au niveau du nœud P de l'arbre syntaxique. La procédure de copiage postulée dans l'IPG, selon laquelle seul le nombre ou le genre du segment source est déterminé conceptuellement et copié sur la cible, est remplacée par une procédure d'unification des informations portées par la source et la cible. Le même processus d'unification est celui qui assure d'une part l'assemblage des segments,

¹³ Le terme 'parallèle' ne réfère pas aux théories interactionnistes de la production mais plutôt à l'idée d'un encodage en parallèle des différentes portions syntaxiques de la phrase.

et donc la construction de l'arbre syntaxique, et d'autre part l'accord, i.e. la vérification que les informations portées par les deux segments sont compatibles.

Pour revenir aux effets conceptuels observés sur le plan empirique, voyons comment ce modèle IPF peut en rendre compte. En ce qui concerne les effets rapportés pour le nom sujet, prenons par exemple celui que nous rapportons dans le Chapitre 2 : moins d'erreurs surviennent lorsque le nom sujet a un genre conceptuel que lorsqu'il a un genre grammatical. Si la représentation conceptuelle contient une information quant au sexe de l'expérienceur, cette information est supposée déterminer à la fois le genre du segment correspondant au nom sujet et celui du segment correspondant à l'adjectif prédicatif. La procédure d'unification au stade de l'encodage syntaxique a dans ce cas deux informations à sa disposition ; celle sur le nom et celle sur l'adjectif. Si, par contre, la structure conceptuelle ne comporte pas d'information de genre, comme c'est le cas pour les noms inanimés¹⁴, l'adjectif ne peut pas être spécifié directement par ce niveau. L'accord consiste alors en une procédure unidirectionnelle de copiage puisqu'il s'agit d'unifier le segment nominal, spécifié au niveau lexico-syntaxique, au segment adjectival, non spécifié. Une seule information est disponible dans ce cas, i.e. l'information lexico-syntaxique inhérente au nom sujet. L'avantage des noms ayant un genre conceptuel par rapport à ceux ayant un genre grammatical réside dans la redondance de l'information de genre, portée par le segment source et le segment cible lors de l'unification.

Pour expliquer l'effet conceptuel du nom sujet qu'ils rapportent pour les pronoms, Bock et ses collaborateurs reprennent l'hypothèse de l'influence notionnelle et suggèrent le même principe de fonctionnement (sans entrer dans les détails du mécanisme). Le pronom, comme le nom sujet, est supposé recevoir sa marque de nombre directement de la notion conceptuelle de nombre véhiculée par le message :

Because English pronouns in conversational speech often refer deictically, (...) pronoun number should be accessible directly from the features of a message representation. (Bock et al., 1999, p. 341)

Les auteurs expliquent donc l'effet conceptuel sur l'accord du pronom par le fait que les deux éléments impliqués dans l'accord, i.e. le nom sujet et le pronom, sont spécifiés conceptuellement sur la base de la notion de nombre propagée dans le message par l'agent. Le fait que les verbes ne se montrent que faiblement sensibles à la représentation conceptuelle du message est expliqué par le caractère abstrait de la quantification d'une action, par opposition au caractère concret de la

¹⁴ Dans le cas des noms animés ayant un genre de surface, le locuteur a normalement en tête le sexe de la personne à laquelle ce nom réfère. Dans la situation expérimentale que nous avons utilisée, les locuteurs ne disposent toutefois pas de cette information puisque la phrase est imposée et qu'aucune information conceptuelle n'est procurée. Cette situation est donc assimilable à celle des noms ayant un genre grammatical : la spécification du genre est strictement syntaxique et non conceptuelle.

notion de nombre sur les pronoms, assimilables aux noms. Par conséquent, le verbe, plutôt que de puiser sa marque de nombre de la représentation notionnelle du nombre, la copierait de l'information syntaxique portée par le nom sujet.

Suivant notre interprétation du cadre théorique du modèle IPF, l'occurrence d'une erreur d'accord dépend de deux facteurs. Premièrement, elle dépend de la présence d'autres segments pouvant interférer avec la procédure d'unification entre le segment cible et le segment source. Une erreur peut ainsi survenir si le segment cible (ex. le syntagme verbal) est unifié non pas avec le segment source (le syntagme nominal sujet) mais avec un autre segment disponible à ce moment de la production (ex. le syntagme nominal local). Cette erreur dans l'unification est l'explication du phénomène d'attraction du nom local et relève du stade de l'encodage syntaxique de la phrase. Deuxièmement, la redondance de l'information au niveau de l'encodage syntaxique, i.e. la présence d'information sur le segment source et sur le segment cible, assure une plus grande résistance aux erreurs d'attraction. L'unification incorrecte d'un segment cible non spécifié par le niveau conceptuel (ex. le syntagme adjectival si la structure conceptuelle contient un expérienceur inanimé) avec le segment du nom local entraînera automatiquement le copiage de la marque de ce dernier sur le segment cible, et donc une erreur d'attraction. Si, par contre, le segment cible est spécifié au niveau conceptuel (l'expérienceur est un animé), une erreur d'unification avec le segment du nom local générera moins facilement une erreur dans la spécification finale du segment cible, puisque celle-ci peut éventuellement l'emporter lors de l'unification. Ce second facteur, i.e. la spécification du segment cible par le niveau conceptuel, n'aura toutefois un effet réducteur des erreurs d'attraction qu'à condition que l'information portée par le segment source soit identique à celle du segment cible (ex. genre conceptuel). Si le locuteur a sélectionné un nom sujet dont la spécification syntaxique de nombre est différente de sa représentation conceptuelle (ex. un collectif ou un nom ayant un genre de surface opposé au sexe du référent), alors l'information portée par le segment cible, au lieu d'être en redondance avec celle du segment source, est discordante avec cette dernière. Dans ce cas, une erreur d'attraction sera d'autant plus susceptible de survenir que le segment cible, portant une information différente du segment source, porte par conséquent la même information que le segment du nom local¹⁵.

Quel est par conséquent le lien entre l'effet d'attraction et l'effet conceptuel du nom sujet ? Les erreurs d'attraction surviennent au stade de l'encodage syntaxique, au cours de la procédure d'unification. La spécification conceptuelle du verbe ou de l'adjectif peut influencer le processus syntaxique d'attraction dans deux directions différentes : (1) si elle concorde avec celle du nom sujet (redondance),

¹⁵ Pour rappel, les erreurs d'attraction surviennent lorsque le nom sujet et le nom local sont de nombre ou de genre différents. Si le segment cible a une marque opposée au segment du nom sujet, et que ce dernier a lui aussi une marque opposée au segment local, le segment cible a la même marque que le segment local.

elle contribue à *réduire le risque d'attraction* en s'imposant, au niveau de l'unification, sur la marque portée par le nom local avec laquelle elle est en discordance; (2) si elle est en discordance avec celle du nom sujet, elle contribue à *accroître le risque d'attraction* en renforçant l'unification avec la marque du nom local avec laquelle elle concorde. Le phénomène d'attraction peut donc être conçu comme un phénomène relevant du niveau syntaxique mais dont l'ampleur est modulée par le corrélat conceptuel du nom sujet de la phrase.

Qu'en est-il de l'interprétation de l'effet conceptuel sur le nom local ? Les recherches menées par Bock & Miller (1991) et par Hupet, Fayol & Schelstraete (1998) se distinguent de celle de Bock & Eberhard (1993) quant à l'origine des erreurs d'accord qu'elles rapportent. Dans les deux premiers cas, les auteurs ont manipulé des caractéristiques relatives au 'degré de subjectivité' du nom local (*subjecthood*, Keenan, 1976) : son caractère animé / inanimé et la relation sémantique qu'il entretient avec le verbe. Il est plus que probable que le caractère animé du nom local (chez Bock & Miller), ainsi que la possibilité qu'il soit l'agent du verbe sur le plan sémantique (chez Hupet et al.), aient indirectement augmenté les risques d'erreur d'accord par le biais d'erreurs dans l'assignation des fonctions syntaxiques. Le nom local ayant été assigné à la fonction sujet de la phrase, il s'en suit que les éléments devant normalement s'accorder avec le sujet s'accordent avec ce nom local. Autrement dit, ces erreurs d'accord ne feraient que témoigner de dysfonctionnements ayant eu lieu à un stade antérieur à la construction de l'accord, lorsque les unités lexicales sélectionnées reçoivent le rôle syntaxique qu'elles joueront dans la phrase. Par conséquent, nous ne prendrons pas en considération ces résultats dans la discussion des interprétations possibles de l'effet conceptuel, puisqu'ils ne portent pas sur l'opération d'accord mais bien sur l'assignation des rôles syntaxiques, laquelle, on le sait, opère sous haut contrôle conceptuel (ex. Bock & Warren, 1985).

La seule étude rapportant une tendance envers un effet conceptuel direct du nom local sur l'accord est donc celle de Bock & Eberhard. Comment les auteurs interprètent-ils ce résultat ? Ils présentent l'absence d'effet significatif de la collectivité du nom local comme un argument en défaveur de l'hypothèse d'un contrôle notionnel de la marque du verbe :

The results are somewhat less favourable to arguments which emphasise the grounding of verb agreement in the semantics of number (Pollard & Sag, 1988). Though nominal number marking has obvious notional correlates, the control which that marking seems to have over verb number is not necessarily conferred by its meaning. (Bock & Eberhard, 1993, p. 89)

Il nous semble que le raisonnement théorique suivi par les auteurs qui se basent sur l'hypothèse d'un contrôle notionnel de la marque du verbe est incorrect. L'absence d'un effet du nom local chez Bock & Eberhard ne permet pas, selon nous, de conclure que la représentation conceptuelle du message n'a pas influencé l'accord du verbe puisqu'ils n'ont fait que manipuler une propriété du nom local.

Plus important, si les auteurs avaient effectivement trouvé un tel effet, nous pensons que cet effet ne constituerait pas non plus un argument en faveur de l'hypothèse notionnelle, contrairement à l'interprétation que Bock et al. (1999) font de la tendance observée par Bock & Eberhard dans l'accord du verbe¹⁶ :

Notional features must play a role in verb agreement, even if it is a weaker or more indirect role than in pronoun agreement (see Bock & Eberhard, 1993). (Bock et al., 1999, p. 342)

Autrement dit, Bock et al. considèrent que l'effet de la propriété conceptuelle du nom sujet qu'ils rapportent *signifie* la même chose que la tendance envers un effet de la propriété conceptuelle du nom local rapportée par Bock & Eberhard. Tous deux témoigneraient d'une détermination notionnelle, au niveau du message, du nombre du verbe. Le problème central de cette interprétation théorique est qu'*il n'y a aucune raison théorique de prédire que la représentation conceptuelle du nombre du nom local aura un effet sur l'accord*, autrement dit, ni l'absence ni la présence d'un effet conceptuel du nom local n'éclaire l'hypothèse notionnelle. L'hypothèse notionnelle avance que la notion de pluriel au niveau du message provient de *l'agent* de l'action qui est multiple, et que par conséquent cette action est, elle aussi, multiple sur le plan notionnel puisqu'effectuée par plusieurs agents. Dans ce cadre, on ne voit pas pourquoi le concept qui sous-tend le nom local, parce qu'il serait multiple, influencerait la multiplicité de l'action, à moins de postuler également que les relations entre ce concept et l'action sont incorrectes, i.e. que ce concept serait considéré comme l'agent de l'action. Dans ce cas, l'effet conceptuel refléterait tout au plus une erreur dans l'élaboration des relations entre les concepts au niveau de la structure message, autrement dit, le fait que le locuteur n'a pas les idées claires sur ce qu'il veut réellement exprimer. Notons par ailleurs qu'en précisant que le contrôle n'est pas *nécessairement* notionnel (voir citation), Bock & Eberhard semblent sous-entendre que leur résultat ne permet effectivement pas d'écarter l'hypothèse notionnelle.

Pour démontrer un effet conceptuel sur l'accord du verbe, les auteurs auraient dû, selon nous, manipuler la collectivité du nom sujet, et non celle du nom local. Cette démarche est celle que Bock et al. ont suivie, ainsi que celle que nous avons adoptée pour étudier l'accord en genre. Toutefois, comme nous l'avons brièvement suggéré dans la discussion du Chapitre 2, l'effet rapporté ne permet pas non plus de conclure de façon décisive que la représentation conceptuelle contrôle l'accord de l'adjectif prédicatif. Il est possible que l'information *lexicale* de genre portée par un nom animé soit plus résistante à l'erreur parce qu'elle est double : le lemme de *mari* comporte à la fois la propriété syntaxique <masculin> et la propriété sémantique <homme>. Le genre linguistique, ou lexical, d'un nom animé serait donc plus 'fort' que celui d'un nom inanimé, et donc plus résistant à l'erreur. Le

¹⁶ Notons toutefois que cette interprétation des auteurs repose avant tout sur l'effet qu'ils rapportent des propriétés notionnelles du nom sujet, mais ils semblent vouloir étayer cette interprétation par les résultats de Bock & Eberhard, ce que nous semble incorrect.

moindre taux d'erreurs d'accord avec les noms sujets ayant un genre conceptuel ne relève, dans ce cadre théorique, plus du niveau de la représentation conceptuelle du message mais bien du niveau de la représentation lexicale des propriétés sémantiques du nom¹⁷. Cette *hypothèse lexicale*, contrairement à l'hypothèse notionnelle, prédit que le nom local aura plus d'impact sur les erreurs d'attraction s'il possède un genre conceptuel puisque, ayant une propriété sémantique en plus, il serait plus 'fort' et aurait donc d'autant plus de chances d'influencer l'accord.

En résumé, l'observation de *l'absence d'un effet* conceptuel du nom local ne constitue pas un argument contre l'hypothèse notionnelle, mais l'observation de la *présence d'un tel effet* n'est pas non plus un argument en faveur de l'hypothèse notionnelle, pour les raisons exposées ci-dessus. Au contraire, l'observation d'un tel effet constitue un argument qui devrait nous amener à nuancer l'interprétation notionnelle proposée dans le Chapitre 2. Si nous trouvons un effet de la 'force' du genre du nom local, cette observation pourrait indiquer que l'effet rapporté dans le Chapitre 2, que nous avons interprété en faveur de l'hypothèse notionnelle, peut également être interprété par l'hypothèse lexicale. Selon cette dernière, l'effet conceptuel observé témoignerait du fait que c'est la propriété sémantique du genre au niveau lexical qui exerce un contrôle sur l'accord, et non pas sa représentation conceptuelle.

Nous avons conduit une expérience visant à déterminer l'impact de l'information de genre du nom local sur les erreurs d'accord en français. La même expérience est actuellement en cours en italien. Le type de genre du nom local a été manipulé : il pouvait soit posséder un corrélat sémantique (genre conceptuel), soit ne pas en posséder (genre de surface). Dans les deux cas, les noms locaux étaient des noms animés, pour nous assurer que tout effet obtenu ne pourrait pas être attribué à un effet relatif au degré de subjectivité du nom local.

4.3 Expérience

4.3.1. Méthode

Participants. 80 étudiants de l'Université catholique de Louvain âgés de 18 à 25 ans ont participé à l'expérience. Le français était leur langue maternelle et ils ont été payés 100 FB pour leur participation.

¹⁷ Un certain flou réside autour des termes *conceptuel* et *sémantique*. L'un et l'autre sont souvent utilisés pour désigner une même réalité, à savoir l'aspect relatif au sens d'une unité. Nous suggérons d'utiliser le terme général *conceptuel* pour désigner les caractéristiques de sens au niveau de la représentation non linguistique du message et le terme *lexico-sémantique* pour désigner les caractéristiques de sens propres aux unités linguistiques, stockées de façon permanente dans le lexique.

Matériel. Le matériel expérimental était constitué de 16 préambules composés d'un nom sujet (N1) et d'un nom local (N2). Le nom local était inclus dans un syntagme prépositionnel rattaché au nom sujet. Les variables manipulées étaient : (1) le genre du nom sujet (Masculin vs. Féminin), le genre du nom local (Masculin vs. Féminin), et le type de genre du nom local (Conceptuel vs. Surface). Le type de genre du nom sujet était maintenu constant : tous avaient un genre grammatical, pour maximiser la proportion d'erreurs d'accord.

Un pré-test visant à établir la polarité de genre des noms locaux ayant un genre de surface a été conduit. 9 personnes ont évalué la polarité des noms sur une échelle de 1 à 7 ; 1 lorsque le nom réfère exclusivement à un homme, 7 lorsqu'il réfère exclusivement à une femme. 16 noms masculins et 16 noms féminins ont été choisis parmi ceux ayant la plus faible polarité. Rappelons toutefois que la polarité a été introduite comme facteur covariant dans l'analyse des deux expériences présentées dans le Chapitre 3 et qu'aucune des deux analyses n'a mis en évidence un quelconque biais de cette variable sur l'occurrence des erreurs d'accord.

Chaque préambule était accompagné d'un adjectif, sélectionné de telle façon que la marque de genre s'entende à l'oral. Un pré-test a toutefois été conduit sur ces adjectifs pour estimer dans quelle mesure ils pouvaient à la fois qualifier un nom masculin et un nom féminin, ou s'ils entretenaient une préférence pour l'un ou l'autre genre. Septante personnes, autres que celles participant à l'expérience, ont ainsi évalué sur une échelle de 1 à 7 la polarité de genre de 150 adjectifs. Seuls des adjectifs ayant un score situé entre 5 et 7, se situant donc au centre des pôles masculin – féminin, ont été sélectionnés pour l'expérience.

Quatre versions de chaque préambule ont été créées. La longueur des quatre versions a été contrôlée pour qu'il y ait au maximum une syllabe de différence entre chacune d'entre elles. Le Tableau 13 illustre un item dans les quatre conditions expérimentales. L'ensemble des items se trouve dans l'Annexe 7.

Tableau 13 : Exemple d'un item dans les différentes conditions expérimentales. Dans la colonne de gauche, la première lettre réfère au genre du nom sujet, la deuxième au genre du nom local.

Condition	Item
FM, nom local avec genre conceptuel	La démarche du comédien - IRRITANTE
FF, nom local avec genre conceptuel	La démarche de la comédienne - IRRITANTE
FM, nom local avec genre de surface	La démarche du personnage - IRRITANTE
FF, nom local avec genre de surface	La démarche de la fripouille - IRRITANTE

Les items expérimentaux se répartissaient dans quatre listes de présentation différentes, une seule version de chaque item n'apparaissant dans chaque liste. Chaque liste contenait seize items expérimentaux répartis dans les quatre

conditions expérimentales. Les listes contenaient également quarante items de remplissage. Ceux-ci comportaient parfois un nom sujet avec un genre conceptuel, introduisaient une certaine variabilité dans les structures syntaxiques présentées (noms simples, relatives) ainsi que des adjectifs non marqués pour le genre. Les items expérimentaux étaient arrangés dans un ordre semi-aléatoire de sorte que jamais plus de deux d'entre eux ne se suivaient. Les listes commençaient par cinq items de remplissage.

Procédure. Les participants étaient testés séparément, en présence de l'expérimentateur qui régula lui-même l'apparition des items. Le matériel était présenté visuellement. Les deux formes de l'adjectif étaient présentées l'une au-dessus de l'autre pendant 800 ms, suivies d'un écran blanc d'une durée de 500ms. Le préambule était alors affiché pendant 900 ms. On demandait aux participants de lire silencieusement l'adjectif suivi du préambule, puis de répéter la totalité en complétant le préambule le plus rapidement possible dès que celui-ci avait disparu. Les sessions expérimentales étaient enregistrées sur micro-CD.

Catégories d'erreurs. Les réponses des participants ont été réparties en trois catégories différentes : (1) les réponses correctes, (2) les erreurs d'accord en genre entre le nom sujet et l'adjectif prédicatif ; seules les réponses ne contenant qu'une erreur en genre sans aucune autre malformation de l'énoncé étaient placées dans cette catégorie, et (3) les erreurs miscellanées : lorsque les participants ne répétaient pas l'adjectif ou l'un des deux noms présentés, lorsque l'un des mots de la phrase était changé ou remplacé, ou encore lorsque rien n'était produit. Si deux phrases étaient produites successivement, seule la première était catégorisée dans l'une de ces trois sections, y compris si le participant s'auto-corrigeait.

Design et analyse des données. Deux analyses de variance prenant les participants (F1) et les items (F2) comme facteurs aléatoires ont été conduites sur les erreurs d'accord et miscellanées comme mesures dépendantes. Dans chaque analyse, les facteurs suivants ont été combinés de façon orthogonale : le genre du nom sujet, le genre du nom local et le type de genre du nom local.

4.3.2. Résultats

Au total, 976 réponses correctes ont été produites (76.25%) ainsi que 30 erreurs d'accord (2.34%) et 274 erreurs miscellanées. La distribution des erreurs d'accord se trouve dans le Tableau 14.

Distribution des erreurs d'accord. Significativement plus d'erreurs ont été produites lorsque le nom local avait un genre conceptuel que lorsqu'il avait un genre de surface ($F(1,79)=7.58$, $p=.007$; $F(1,14)=5.83$, $p=.02$). Aucun effet significatif de mismatch n'a été observé, comme l'indique l'absence d'interaction entre le genre de N1 et le genre de N2 ($F(1,79)=2.31$, $p=.13$; $F(1,14)=2.97$, $p=.09$). Ni le genre de N1 ni celui de N2 n'ont entraîné d'effet principal sur

l'occurrence des erreurs d'accord (respectivement : $F_s < 1$ et $F_1(1,79)=2.03$, $p=.16$; $F_2(1,14)=1.9$, $p=.18$).

Tableau 14 : Distribution des erreurs d'accord en genre dans les différentes conditions expérimentales. Dans la colonne de gauche, la première lettre réfère au genre du nom sujet, la deuxième au genre du nom local.

	Nom local de genre de surface	Nom local de genre conceptuel	Total
MM	1	2	3
MF	2	10	12
FM	3	5	8
FF	2	5	7
Total	8	22	30

Distribution des erreurs miscellanées. Il s'est produit plus d'erreurs de ce type lorsque le nom local était féminin que lorsqu'il était masculin ($F_1(1,79)=9.67$, $p=.003$; $F_2(1,14)=6.09$, $p=.018$). Aucun autre effet ni interaction n'est à rapporter.

4.4 Discussion

Les résultats de cette expérience indiquent qu'un nom local de genre conceptuel est un meilleur 'attracteur' qu'un nom local de genre de surface. La différence entre ces deux types de noms est que le premier fait référence à une entité dont on connaît le sexe, tandis que le second ne porte pas cette information relative au sexe de l'entité à laquelle il réfère. Certes, cette information est potentiellement présente dans le discours en situation d'élocution spontanée, mais ce n'était pas le cas dans la situation expérimentale manipulée.

Deux questions se posent quant à ce résultat. Premièrement, que signifie-t-il et quelle est sa portée théorique pour l'interprétation des résultats précédents ? Deuxièmement, comment peut-on expliquer la différence observée entre le français et l'anglais ?

4.4.1. Remise en question de l'hypothèse notionnelle

Dans le Chapitre 2, nous avons proposé une interprétation de l'effet conceptuel observé sur le nom sujet en termes d'une détermination de la structure conceptuelle du message sur l'accord. Cette hypothèse, que nous avons appelée l'hypothèse notionnelle, considère que l'ensemble des concepts liés à l'expérimenteur (peu importe son rôle thématique exact) dans la représentation conceptuelle du message reçoivent la notion de genre que celui-ci véhicule en tant qu'être sexué. Une telle hypothèse suppose un contrôle direct de la représentation notionnelle de genre sur les différents segments de la phrase. Dans les phrases que nous avons manipulées, l'adjectif prädicatif est dans ce cadre spécifié directement par la notion de genre

impulsée par l'expérienceur dans le message, pour autant que celui-ci soit sexué. Dans le cas d'un concept non sexué, cette information ne se trouve pas dans la représentation conceptuelle et ne peut donc pas déterminer directement, notionnellement, le genre de l'adjectif. Celui-ci est donc spécifié au sein du système linguistique par l'information lexico-syntaxique que porte le nom sujet.

L'effet conceptuel du nom sujet que nous avons mis en évidence dans le Chapitre 2 est réapparu lorsque nous avons manipulé le nom local : un plus grand nombre d'erreurs ont été produites lorsque ce dernier était de genre conceptuel que lorsqu'il était de genre de surface. Sur le plan théorique, il n'est pas pertinent de recourir à l'hypothèse notionnelle pour rendre compte de cet effet, comme nous l'avons expliqué dans l'introduction. La raison est que cela n'a pas de sens de supposer que le concept sous-jacent au nom local ait déterminé la notion de genre du message puisqu'il n'occupe pas la position d'expérienceur dans la structure du message. L'effet conceptuel du nom local ne peut donc s'expliquer que par le fait que les noms ayant un genre conceptuel possèdent une propriété sémantique, dans le lexique, que ceux ayant un genre de surface n'ont pas. La présence de cette propriété renforcerait le genre du nom local et augmenterait ainsi le risque d'attraction. Autrement dit, la différence entre les noms locaux ayant un genre conceptuel et ceux ayant un genre de surface, qui est celle pertinente pour rendre compte de l'effet observé, se situe au niveau lexical, linguistique, et non au niveau conceptuel extra-linguistique. Un nom local avec un genre conceptuel est doublement spécifié pour le genre dans le lexique : par une propriété syntaxique et par une propriété sémantique. Un nom local avec un genre de surface n'est spécifié que par sa propriété syntaxique. Par conséquent, cette double information au niveau lexical rendrait le genre du nom local plus 'fort', et donc plus susceptible de générer une erreur d'attraction. Le même principe peut expliquer pourquoi un nom sujet de genre conceptuel a entraîné moins d'erreurs qu'un nom de genre de surface : parce qu'étant doublement spécifiée, l'information de genre qu'il véhicule est moins susceptible d'être perdue ou incorrectement traitée.

Comment l'hypothèse lexicale peut-elle être implémentée dans le modèle IPG/IPF présenté dans l'introduction ? Le mécanisme responsable de la réalisation de l'accord dans ces modèles est l'unification des segments syntaxiques. Dans ce cadre, l'information lexico-sémantique de genre sur un segment nominal devrait contribuer à renforcer sa probabilité d'unification avec le segment adjectival. Autrement dit, la seule façon de rendre compte de cet effet dans la perspective de l'unification est de considérer que l'information lexico-sémantique sur le segment du nom local viendrait renforcer l'information syntaxique et ainsi augmenter le risque qu'il soit unifié avec le segment adjectival. De même, la présence de cette information lexico-sémantique sur le nom sujet contribuerait à réduire les risques que le segment adjectival s'unifie avec un autre segment que celui du nom sujet. Le bénéfice des noms ayant un genre conceptuel ne provient plus de la redondance entre l'information portée par le segment source et le segment cible, comme il est

postulé dans l'hypothèse notionnelle, il provient du fait que le segment source, lui-même, possède une double information de genre. Cette force le rendrait plus susceptible d'être unifié au segment cible, augmentant le risque d'attraction s'il s'agit du nom local, réduisant ce risque s'il s'agit du nom sujet.

Ce mécanisme que nous avançons pour rendre compte de l'effet lexicosémantique du nom local dans le cadre du modèle IPF reste toutefois hypothétique. Les modèles interactifs de la production (ex. Dell, 1986) permettent également de rendre compte de cet effet sur l'accord. Comme nous l'avons déjà dit, ces modèles, essentiellement construits pour implémenter les processus de traitement au niveau lexical, ne traitent pas de l'accord mais avancent des principes intéressants quant à l'organisation de ce niveau. Comme l'écrivent MacDonald, Pearlmutter & Seidenberg (1994) :

Knowledge of the grammar (...) can largely be encoded in the structure of the lexicon itself, e.g., by weights on connections between units representing different types of information. (MacDonald & al., 1994, p. 696)

Comme l'hypothèse lexicale consiste justement à expliquer un pattern d'erreurs syntaxiques par la façon dont les unités impliquées sont représentées au niveau lexical, ces modèles procurent un cadre de référence à cette hypothèse. Le principe de base est que l'activation peut se propager entre les différents niveaux de traitement (lexical, syntaxique, phonologique) et que donc le lexique assure un contrôle important sur le traitement syntaxique. Dans un modèle interactif, les unités lexicales (lemmes) sont reliées aux différentes propriétés syntaxiques et sémantiques qu'elles possèdent. Lorsqu'une unité lexicale est activée par le niveau conceptuel, l'activation se propage à l'ensemble des nœuds auxquels elle est connectée dans le réseau. Le genre d'un nom sujet aura d'autant plus de poids dans un tel modèle qu'il est à la fois représenté comme propriété sémantique et comme propriété syntaxique de ce nom. Il activera par conséquent d'autant plus facilement la forme correcte de l'adjectif. Un nom local, même s'il n'entretient pas de relation syntaxique spécifique avec l'adjectif, entraîne lui aussi une certaine activation de ses propriétés. S'il active à la fois une propriété sémantique et une propriété syntaxique de genre, son genre aura d'autant plus de chances d'activer la forme de l'adjectif que s'il n'active qu'une propriété syntaxique.

La Figure 7 illustre les activations en jeu dans la phrase *La victime du beau-frère est morte*. Le nom sujet ne possède qu'une propriété syntaxique de genre, il n'a pas de propriété sémantique. Le nom local possède les deux propriétés, ce qui augmente le risque de production d'une erreur d'attraction, d'autant plus que le nom sujet ne possède pas de propriété sémantique. Les deux premières lignes de nœuds font partie du niveau lexical, la dernière, du niveau phonologique.

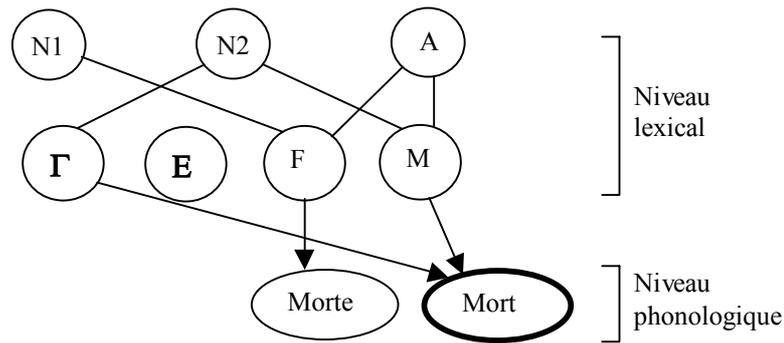


Figure 7. Illustration partielle des activations dans le réseau lexical de la phrase : *La victime du beau-frère (est) morte*. La double activation des propriétés du nom local augmente le risque d'erreur d'attraction.

La situation inverse est représentée dans la Figure 8 pour la phrase *Le beau-frère de la victime est mort*. Dans ce cas, la double propriété de genre du nom sujet limite les risques d'erreurs d'attraction, d'autant plus que le nom local n'est que simplement marqué.

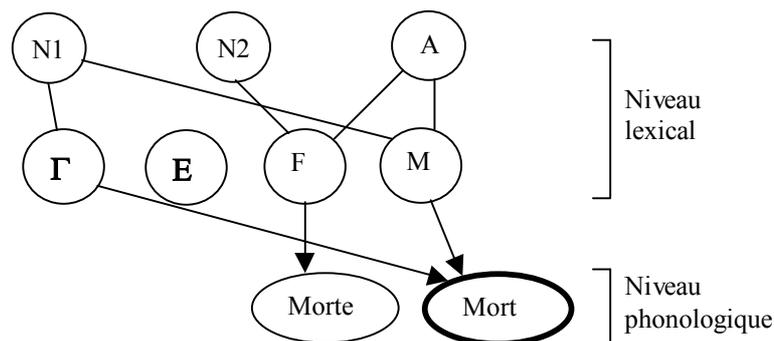


Figure 8. Illustration partielle des activations dans le réseau lexical de la phrase : *Le beau-frère de la victime (est) mort*. La double activation des propriétés du nom sujet limite le risque d'erreur d'attraction.

Notons que la relation conceptuelle et syntaxique particulière qui unit le nom sujet et l'adjectif entraîne une plus forte activation entre le genre du nom sujet et la forme correspondante de l'adjectif qu'entre celui du nom local et la forme de l'adjectif qui lui correspond. Cela n'est pas représenté dans les figures.

Il arrive toutefois que l'information lexico-sémantique ne concorde pas avec l'information lexico-syntaxique, surtout en ce qui concerne le nombre : certains noms sont syntaxiquement pluriels alors qu'ils réfèrent à des entités uniques (ex. les ciseaux, les lunettes), tandis que d'autres sont syntaxiquement singuliers mais sémantiquement pluriels (ex. la foule, le groupe). L'accord du verbe avec de tels noms doit se faire sur la base de l'information syntaxique, sous peine de générer des erreurs (ex. les ciseaux est coupant). Dans ce cas, l'activation de la propriété lexico-sémantique, au lieu de renforcer le choix syntaxique, diminue son impact, augmentant ainsi le risque d'erreur s'il s'agit du nom sujet, le réduisant s'il s'agit

du nom local. Pour reprendre la terminologie du modèle de compétition proposé par Bates & MacWhinney (1989), une compétition s'engage entre l'indice sémantique et l'indice syntaxique dans leur lutte pour le contrôle de l'accord du verbe, lorsque les deux indices sont opposés. On remarque effectivement que même s'il arrive que le locuteur se trompe, et qu'il se trompe même plus souvent lorsqu'il y a discordance entre les informations syntaxique et sémantique (comme l'indiquent les tendances observées par Bock & Eberhard et par Bock & al. et les résultats du Chapitre 3), la plupart du temps l'accord se fait correctement. Les locuteurs favorisent donc l'information syntaxique, une des raisons étant certainement que c'est le meilleur indice pour former une phrase correcte (voir Synthèse théorique pour une discussion de ce point). Autrement dit, il semble que le système responsable de construire l'accord est, dans une certaine mesure, *contraint* par l'information lexico-sémantique, mais *dominé* par l'information lexico-syntaxique.

Cette prédominance de l'information syntaxique peut expliquer pourquoi la tendance rapportée par Bock & Eberhard n'apparaît que dans la condition où les noms locaux collectifs sont également syntaxiquement pluriels : la contrainte lexico-sémantique est trop faible pour générer à elle seule une erreur (collectif au singulier), par contre, son effet est révélé lorsqu'elle est combinée avec la contrainte syntaxique (collectif pluriel). La prédominance de la propriété syntaxique dans la détermination de la forme de l'adjectif peut être implémentée dans un modèle interactif en renforçant le poids de la relation entre la propriété syntaxique du lemme et la forme de l'adjectif, comparativement à sa propriété lexico-sémantique.

Doit-on abandonner l'hypothèse notionnelle ? Il est vrai qu'elle ne peut pas rendre compte de l'effet conceptuel du nom local sur l'accord, tandis que l'hypothèse lexicale peut, à elle seule, rendre compte de l'effet conceptuel du nom local et du sujet. Des auteurs tels que Levelt et al. (1999) suggèrent que dans une telle situation, la préférence doit être donnée à l'hypothèse la plus 'minimaliste', en vertu du principe méthodologique du rasoir d'Ockham. Toutefois, bien que l'hypothèse lexicale soit effectivement suffisante pour rendre compte de cet effet, cela ne signifie pas qu'il n'y ait pas, en plus du contrôle de l'information lexico-sémantique, un contrôle de l'information conceptuelle venant de la représentation du message. Il est possible que les deux facteurs influencent la réalisation de l'accord, eux aussi chacun dans une certaine mesure, et peut-être en interaction l'un avec l'autre¹⁸.

Mais alors, comment mettre à l'épreuve l'hypothèse notionnelle si, ni l'étude des propriétés conceptuelles du nom sujet, ni celle des propriétés conceptuelles du nom local ne sont pertinentes ? Un premier type de recherches consiste à manipuler

¹⁸ Le rôle des différents facteurs dans l'accord (conceptuel, lexico-sémantique, lexico-syntaxique, morphophonologique) et leurs relations sont abordés dans la Synthèse théorique (Chapitre 8).

la 'notion' au niveau du message sans qu'elle n'aie de représentation lexicale. Les études sur la distributivité procurent un tel paradigme puisque les phrases distribuées (ex. *Le timbre sur les enveloppes*) sont notionnellement plurielles bien que les propriétés lexicales sémantiques et syntaxiques des noms sujets soient singulières. Un effet de la distributivité a ainsi été rapporté dans différentes langues (Vigliocco & al., 1995, en italien ; Vigliocco & al., 1996a, en espagnol ; Vigliocco et al., 1996b, en français et en néerlandais ; Eberhard, 1999, en anglais). Ces études ont montré que les locuteurs produisent significativement plus de verbes au pluriel lorsque le sujet est distribué que lorsqu'il ne l'est pas, malgré que ses propriétés lexico-sémantique et lexico-syntaxique soient singulières. Ce résultat ne peut pas s'interpréter par l'hypothèse lexicale puisque dans ces recherches, l'information lexicale est au singulier. Seule l'information présente au niveau de la représentation conceptuelle de la phrase est multiple et peut donc avoir déterminé ce résultat, comme le suggère l'hypothèse notionnelle¹⁹.

Deuxièmement, il s'agira de comparer, dans des conditions de contrôle satisfaisantes, les effets supposés conceptuels et ceux supposés lexicaux. L'hypothèse notionnelle prédit que la présence d'une triple information, i.e. conceptuelle, lexico-sémantique et lexico-syntaxique, entraînera un effet plus important que la présence d'une double information, lexico-sémantique et lexico-syntaxique, limitée au lexique. Les effets que nous rapportons sur le nom sujet et le nom local ne sont effectivement pas tout à fait identiques, bien qu'ils aient été testés en manipulant la même variable (type de genre ; conceptuel vs. surface) : l'effet d'un corrélât conceptuel sur le nom sujet (9 vs. 33 erreurs) a tendance à être plus net que celui d'un tel corrélât sur le nom local (8 vs. 22).

L'effet de la distributivité semble par ailleurs être de moindre ampleur que celui de la collectivité du nom sujet. Tandis que la pluralité des noms distribués n'est représentée qu'au niveau notionnel, elle est représentée à la fois au niveau notionnel et au niveau lexical pour les noms collectifs. La seule langue dans laquelle les deux variables ont été manipulées est l'anglais. Les études classiques sur la distributivité n'ont indiqué aucun effet de cette variable (Bock & Miller, 1991 ; Vigliocco et al., 1996a et 1996b). Toutefois, en modifiant le paradigme et en présentant des images illustrant les préambules, i.e. en renforçant la représentation conceptuelle de la phrase, Eberhard (1999) a rapporté un effet de distributivité. L'obtention de cet effet repose donc entièrement sur l'insistance placée sur la représentation conceptuelle du message, ce qui ne semble pas être le cas en ce qui concerne la collectivité du nom sujet, puisque pour cette variable, une nette tendance a été rapportée sans présentation d'images (Bock & al., 1999).

Ces deux tendances restent bien entendu à vérifier dans des recherches mieux contrôlées. Par exemple, il est possible de créer des items directement comparables (ex. *la victime de l'accident* vs. *l'accident de la victime*). Elles vont néanmoins

¹⁹ Le concept et les effets de distributivité sont discutés plus en détails dans la Synthèse.

dans la direction attendue si on fait l'hypothèse qu'une information double, i.e. lexicale et notionnelle (comme c'est le cas pour les noms ayant un genre conceptuel ou les collectifs), assure une réalisation plus correcte de l'accord qu'une information simple (uniquement lexicale, ex. noms locaux, ou uniquement notionnelle, ex. noms distribués).

Enfin, les recherches présentées dans le Chapitre 3 semblent indiquer que la référence à un personnage dans le contexte influence l'accord de l'adjectif avec des noms sujets de genre de surface. Les locuteurs ont produit un nombre considérable d'erreurs d'accord lorsque le sexe du référent et l'information syntaxique de genre portée par le nom sujet étaient opposés (ex. *la victime*, pour référer à un homme). L'information conceptuelle concernant le sexe du référent est entièrement déterminée par l'intention communicative du locuteur, elle est donc variable et ne peut par conséquent pas être stockée comme propriété sémantique permanente du nom. L'observation d'un effet de cette variable suggère que de l'information conceptuelle extra-linguistique assure un certain contrôle de la procédure d'accord. Un examen attentif du matériel expérimental révèle toutefois qu'en introduisant un nom propre dans la phrase contexte, nous n'avons pas seulement manipulé un facteur conceptuel, mais aussi un facteur relatif aux propriétés morphophonologiques du nom propre et un facteur relatif à ses propriétés syntaxiques. Les noms propres, comme les noms communs, sont en effet porteurs d'information morphophonologique sur la terminaison du nom (ex. *-ine* dans *Céline* ou *-ette* dans *Bernadette* sont des terminaisons typiquement féminines). Ils comportent également de l'information syntaxique relative à leur utilisation dans la phrase ; par exemple, un nom propre féminin nécessite un accord féminin de l'adjectif qui le qualifie. Ces deux propriétés, morphophonologique et syntaxique, internes au système linguistique, peuvent-elles expliquer l'effet observé ?

Nous avons testé l'hypothèse d'une influence des propriétés morphophonologiques du nom propre dans les expériences menées en français. Pour ce faire, nous avons calculé la corrélation entre l'indice de prédictibilité du genre porté par la terminaison du nom propre (estimé d'après les tables de Tucker, Lambert & Rigault, 1977, voir Chapitre 5) et l'occurrence des erreurs d'accord. Aucune corrélation n'a été trouvée lorsque le sexe du référent et le genre du nom sujet étaient en discordance ($r = 0.075$). Autrement dit, l'important taux d'erreurs dans cette condition ne peut pas être expliqué par une discordance entre la propriété lexico-syntaxique de genre du nom sujet et l'information morphophonologique de genre véhiculée par le nom propre²⁰.

Il reste toutefois que la propriété lexico-syntaxique de genre des noms propres coïncidait toujours avec sa propriété conceptuelle (le sexe). Nous ne pouvons donc

²⁰ Notons également qu'aucune corrélation n'a été trouvée lorsque le sexe du référent et le genre du nom sujet étaient en concordance ($r = (-)0.039$), ce qui indique qu'une concordance entre l'information morphophonologique du nom propre et l'information syntaxique du nom sujet ne constitue pas non plus un facteur favorisant une réalisation correcte de l'accord.

pas exclure que l'effet conceptuel rapporté soit lié à la discordance entre la propriété syntaxique du nom propre et la propriété syntaxique du nom sujet, les deux noms étant en relation de dépendance anaphorique. Les recherches futures pourront se pencher sur cette question en présentant par exemple non plus une phrase contexte mais une image représentant un personnage de sexe masculin ou féminin, afin de limiter la manipulation à de l'information conceptuelle, sans plus manipuler d'information syntaxique. L'interprétation des résultats rapportés dans le Chapitre 3, en termes d'influence conceptuelle extra-linguistique sur l'accord, restera en suspens tant que les effets de ces deux facteurs n'auront pas été dissociés sur le plan expérimental.

4.4.2. Interprétation des différences entre le français et l'anglais

Notre recherche rapporte un effet significatif de l'information conceptuelle de genre sur le nom local en français. La manipulation de la collectivité du nom local par Bock & Eberhard n'a, par contre, pas révélé d'effet significatif de l'information conceptuelle de nombre sur le nom local. Comment expliquer cette différence ?

Tout d'abord, Bock et al. n'ont pas non plus rapporté d'effet significatif de la collectivité du nom sujet. Si les auteurs avaient observé un effet significatif de cette variable conceptuelle sur le nom sujet, alors l'absence d'effet du nom local rapportée par Bock & Eberhard aurait pu plaider en faveur de l'hypothèse notionnelle pour expliquer l'effet du nom sujet. Mais ni l'information conceptuelle sur le nom sujet, ni celle sur le nom local, n'ont entraîné d'effet significatif. Autrement dit, les locuteurs anglophones semblent faire peu état de l'information relative à la notion sous-jacente au nombre syntaxique, que cette notion se situe au niveau de la représentation conceptuelle ou comme propriété lexico-sémantique du mot.

L'absence d'effet conceptuel en anglais n'est toutefois pas nouvelle, plusieurs études se sont penchées sur la question, et diverses explications théoriques ont été avancées (ex. Vigliocco et al., 1996b, voir aussi Synthèse). Bock et al. rapportent toutefois une forte tendance des locuteurs à considérer la multiplicité notionnelle des noms sujets collectifs pour accorder le verbe. Cette tendance peut refléter l'impact de la représentation notionnelle multiple du message sur la spécification du nombre du verbe ; elle peut aussi être expliquée par le fait que l'information lexico-sémantique plurielle portée par le nom sujet collectif l'a emporté sur l'information syntaxique singulière. L'observation d'une tendance similaire (bien que moindre) en ce qui concerne le nom local suggère que l'information prise en compte est bien la propriété lexico-sémantique du nom et non l'information notionnelle extra-linguistique²¹. Toutefois, Eberhard (1999) a montré que,

²¹ Nous faisons ici l'hypothèse que le nombre des noms collectifs est, comme le genre conceptuel, représenté comme une propriété lexico-sémantique du lemme. Un nom collectif, qu'il soit syntaxiquement singulier ou pluriel, réfère toujours à une entité plurielle. Sa numérosité est donc fixe et ne dépend pas du contexte communicatif. La distinction entre les aspects conceptuel (extra-

moyennant certaines conditions, les locuteurs anglophones pouvaient également considérer l'information notionnelle de nombre, i.e. lorsque cette information était rendue suffisamment concrète et imagée (effet de distributivité). Ce dernier résultat ne peut être interprété comme un effet lexico-sémantique puisque la pluralité des items à référence multiple ne peut être dégagée que sur la base de l'information contextuelle et ne peut donc pas être stockée comme propriété lexico-sémantique du nom.

En résumé, les résultats semblent indiquer que les locuteurs anglophones et francophones peuvent tous deux utiliser, dans une certaine mesure, l'information conceptuelle extra-linguistique et l'information lexico-sémantique pour construire l'accord. Des variations entre les deux langues existent toutefois en ce qui concerne l'importance donnée aux différents indices conceptuel, lexico-sémantique et syntaxique dans la réalisation de cette opération syntaxique. Ces variations, pouvant être liées à des différences entre le nombre et le genre ou encore au statut fonctionnel de l'accord dans la langue, sont discutées dans la Synthèse théorique (Chapitre 8).

4.5 Conclusion

Dans ce chapitre, nous n'avons plus discuté les résultats sous l'angle des hypothèses maximale vs. minimale, comme nous l'avions fait pour les chapitres précédents. Les hypothèses mises en regard ont été dénommées hypothèse notionnelle et hypothèse lexicale. Ce changement dans la terminologie reflète le fait qu'il ne s'agit plus ici d'opposer une conception selon laquelle seule l'information lexico-syntaxique détermine l'accord à une conception selon laquelle l'information conceptuelle peut, elle aussi, exercer un contrôle sur la réalisation de l'accord. L'opposition porte sur la détermination de la *nature* de la représentation ayant causé un effet conceptuel : est-elle conceptuelle, i.e. extra-linguistique, relevant du niveau de l'intention communicative comme le suggère l'hypothèse notionnelle ou est-elle sémantique-lexicale, i.e. intra-linguistique, comme le suggère l'hypothèse lexicale ?

linguistique) et lexico-sémantique de la représentation du nombre nous semble pertinente en ce qu'elle permet de rendre compte de la différence rapportée entre les effets de distributivité et de collectivité du nom. Tandis qu'une nette tendance envers un effet de la collectivité (le nom est pluriel sur le plan lexico-sémantique) est systématiquement rapportée, ce n'est pas le cas de l'effet de distributivité (le nom est singulier sur le plan lexico-sémantique). Nous ne nous avançons toutefois pas quant à l'existence d'une propriété lexico-sémantique de nombre pour les pluriels réguliers. Leur sémantique dépend entièrement de la représentation conceptuelle du message, et l'on voit mal comment une propriété sémantique lexicale fixe pourrait être encodée dans le lexique, étant donné l'existence de différents types de singuliers et de pluriels (ex. singulier générique comme dans *le chien est un animal fidèle*, singulier quantificatif comme dans *il vaut mieux avoir un chien que deux*, ou encore singulier indéfini comme dans *un chien a mordu le voisin*) (voir Dominguez, Cuetos & Segui, 1999, pour un support expérimental à l'hypothèse que la forme plurielle régulière est récupérée à partir de la forme singulière, tandis que deux entrées lexicales distinctes sont réservées aux noms opposés par le genre conceptuel, voir aussi Synthèse théorique).

L'hypothèse notionnelle est incompatible avec une vision séquentielle et modulaire de la composante syntaxique, comme le soutiennent encore récemment Levelt (Levelt et al., 1999) ou Garrett (in press). Selon celle-ci, les opérations syntaxiques sont calculées sur la base d'un input strictement limité aux propriétés syntaxiques véhiculées par les lemmes. Leurs propriétés sémantiques ne pénètrent pas le traitement syntaxique au-delà de l'assignation des fonctions syntaxiques. Cette vision repose elle-même sur la décomposition du traitement lexical en deux étapes successives : la première, responsable du traitement des propriétés sémantiques, la seconde, responsable du traitement des propriétés syntaxiques (voir Schriefers, 1993 et Van Berkum, 1997, pour des arguments expérimentaux). L'hypothèse lexicale est, elle aussi, incompatible avec cette vision modulaire stricte qui suppose une 'hypermodularisation' de la production. Cette hypothèse n'est toutefois pas incompatible avec une vision plus large de la modularité, selon laquelle les opérations syntaxiques se font sur la base de représentations strictement linguistiques, y compris lexico-sémantiques, mais sont aveugles aux informations conceptuelles extra-linguistiques. Le cloisonnement dans cette perspective se fait quant aux informations contextuelles non linguistiques, et l'opération syntaxique d'accord peut être guidée par l'information sémantique présente dans le lexique. Cette vision de la modularité est, selon nous, proche de celle développée par Fodor pour rendre compte des mécanismes de compréhension du langage ; elle suggère que le cloisonnement se fait essentiellement par rapport aux informations contextuelles, variables, et non par rapport aux informations linguistiques, stockées de façon permanente dans le lexique.

Chapitre 5

Quand morphophonologie et syntaxe collaborent

Dans les deux premiers chapitres, nous avons étudié la perméabilité du processus d'accord à l'information conceptuelle de genre. Cependant, pour soutenir l'hypothèse que les traitements syntaxiques au niveau fonctionnel opèrent sur la base d'un input maximal au niveau fonctionnel, il ne suffit pas de démontrer son ouverture sur l'information conceptuelle, encore faut-il observer qu'ils intègrent toutes les informations pertinentes, et donc aussi les informations relatives à la morphophonologie des mots. Ce chapitre présente la manipulation de deux variables relatives à la morphophonologie du syntagme nominal sujet et l'analyse de leur effet sur la réalisation de l'accord en genre de l'adjectif prédicatif. La première variable a trait à l'information morphophonologique de genre présente sur la terminaison du nom. La seconde variable concerne l'information morphophonologique de genre sur l'article précédent le nom. Nous montrons que la présence de ces deux informations facilite une réalisation correcte de l'accord en français. Une étude parallèle menée en italien est ensuite brièvement présentée, dans laquelle nous rapportons également un effet morphophonologique sur la réalisation de l'accord. Une discussion théorique des deux recherches est avancée pour rendre compte des mécanismes potentiels responsables de cet effet et des différences interlinguistiques mises en évidence.

5.1 Introduction

La perspective modulaire de la production du langage, développée par des auteurs comme Garrett (1988, 1993), Bock (ex. 1986, 1987) et Levelt (ex. 1989, Bock & Levelt, 1994, Levelt, Roelofs & Meyer, 1999), domine encore largement les théories psycholinguistiques de nos jours. En ce qui concerne la production de phrases, les relations entre les niveaux de traitement de la syntaxe et de la forme phonologique sont conçus comme strictement séquentiels, le premier précédant le second, sans possibilité d'une rétroaction. Si l'on se réfère au modèle proposé par Garrett et illustré dans la Figure 1 de l'Introduction générale, les processus de traitement syntaxique responsables d'élaborer la structure fonctionnelle de la phrase au stade de l'encodage grammatical opèrent avant la récupération de la morphophonologie des mots qui a lieu au stade de l'encodage phonologique de la production. Les modèles élaborés par Levelt et Bock adoptent la même hypothèse :

le flux d'information entre les niveaux de production se fait de haut en bas dans le système, sans influence possible du bas vers le haut. Une telle restriction sur le flux d'information n'empêche toutefois pas un fonctionnement en parallèle des différents niveaux ; par exemple, une unité de production, une fois traitée au niveau de l'encodage grammatical, peut être transférée au niveau de l'encodage phonologique qui la traitera en même temps que l'unité d'encodage grammatical suivante (Bock, 1987).

5.2 L'interface entre phonologie et syntaxe : les observations

L'adoption de la perspective théorique modulaire repose principalement sur l'observation des erreurs spontanées. L'argument avancé par Garrett (1993) est que les erreurs d'échange, portant sur des unités fonctionnelles et supposées avoir lieu au stade de l'encodage grammatical, ne présentent aucune similarité phonologique entre les mots impliqués. Il semblerait donc que le traitement phonologique n'ait pas encore eu lieu au stade de l'encodage grammatical, sinon il aurait dû transparaître dans la survenue des erreurs fonctionnelles. Les études menées sur l'état du 'mot sur le bout de la langue' indiquent aussi que certaines propriétés sémantiques et phonologiques du mot peuvent être accessibles au locuteur sans qu'il puisse le dénommer, autrement dit sans qu'il puisse en récupérer la forme phonologique (ex. Vigliocco, Antonini & Garrett, 1997). Sur le plan expérimental, une étude sur l'accord menée par Bock & Eberhard (1993) en anglais semble également indiquer que les erreurs d'attraction ne sont pas sensibles à la présence d'une (fausse) marque morphophonologique de pluriel (pseudo-pluriel) sur le nom local (voir Introduction générale pour plus de détails). Les auteurs concluent de ce résultat que le processus d'accord au stade de l'encodage grammatical n'est pas perméable à l'information morphophonologique de pluriel.

Ces observations plaident en faveur d'une séquentialité stricte des encodages grammatical et phonologique. Par contre, l'analyse faite par Dell & Reich (1987) des erreurs de substitution sémantique a montré que ces dernières impliquent deux unités proches phonologiquement dans des proportions supérieures au niveau de chance.

Deux études expérimentales (Levelt & Maasen, 1981 et Bock, 1986) ont également montré une tendance envers un effet phonologique sur les choix syntaxiques. Levelt & Maasen (1981) ont conduit une recherche dans laquelle les participants devaient décrire des animations de formes géométriques. Les formes pouvaient soit voyager dans la même direction, soit dans des directions opposées. Les auteurs ont tout d'abord constaté que la structure syntaxique utilisée pour décrire ces animations variait selon la direction : lorsque les formes allaient dans des directions opposées, les participants avaient tendance à coordonner deux

phrases entières (ex. *Le carré va vers le haut ET le cercle va vers le bas*), tandis que lorsque les deux formes allaient dans la même direction, la coordination se faisait au sein du syntagme nominal (ex. *Le carré ET le cercle vont vers le haut*). Les auteurs ont également manipulé l'accessibilité lexicale des différentes formes. Ils ont ainsi montré que lorsque les formes étaient moins accessibles, les participants avaient tendance à utiliser plus de phrases coordonnées que lorsque les formes étaient plus accessibles. L'utilisation de phrases, plutôt que de syntagmes, coordonnées permet en effet au locuteur de disposer de plus de temps pour récupérer le mot correspondant à la seconde forme. Autrement dit, la difficulté d'accès à la phonologie du mot semble avoir influencé le choix de la structure syntaxique. Les auteurs rapportent par ailleurs que cette 'révision' syntaxique prend du temps, comme en témoigne l'augmentation du temps pris par les locuteurs pour commencer à décrire l'image dans cette condition.

L'étude de Bock (1986) visait à tester l'effet de l'amorçage phonologique sur le choix de la structure syntaxique active ou passive. Les participants devaient décrire des images d'actions impliquant un agent et un patient (ex. la foudre qui frappe une église). Les images étaient précédées d'un mot amorce pouvant être proche sur le plan phonologique soit de l'agent, soit du patient de l'action. Bock rapporte que les participants ont tendance à produire plus de structures actives lorsque l'amorce est phonologiquement proche du patient (ex. *la foudre frappe l'église*) tandis que plus de phrases passives sont produites lorsque l'amorce est phonologiquement proche de l'agent (ex. *l'église est frappée par la foudre*)²². Autrement dit, l'inhibition phonologique semble avoir influencé le choix de la structure syntaxique au niveau de l'encodage grammatical. En contrôlant mieux les délais de présentation de l'amorce par rapport à l'image à décrire, Bock (1987) a répliqué l'effet rapporté en 1986, atteignant cette fois un seuil de signification satisfaisant pour conclure à une influence phonologique sur la structure syntaxique.

Enfin, une toute récente étude de Hartsuiker, Schriefers, Bock & Kikstra (in prep.) a mis en évidence l'impact d'une autre variable morphophonologique sur le traitement syntaxique : la transparence du cas. Les auteurs ont utilisé le paradigme de complètement de phrases ; les préambules contenaient un nom sujet et un nom local. Les auteurs ont manipulé le cas du nom local (datif vs. accusatif) en allemand²³ ; dans la condition datif, le cas est marqué sur le déterminant précédant le nom local tandis que dans la condition accusatif, le déterminant peut à la fois indiquer l'accusatif et le nominatif. La condition accusatif est donc ambiguë quant au rôle syntaxique du nom local, celui-ci pouvant donc potentiellement assurer la fonction sujet (au nominatif). Les auteurs rapportent un effet de la transparence du

²² En ce qui concerne l'effet d'amorçage, un effet d'inhibition plutôt que d'activation a en effet été rapporté lorsque l'amorce est présentée avant l'image à dénommer, tandis qu'un effet de facilitation apparaît quand l'amorce suit l'image (Schriefers, Meyer & Levelt, 1990).

²³ La détermination du cas en allemand dépend de la préposition qui l'introduit ; les auteurs ont donc manipulé la préposition.

cas sur le déterminant, un plus grand nombre d'erreurs d'accord s'étant produites lorsque le cas du nom local était ambigu. Les auteurs concluent que l'information morphophonologique relative au cas, portée par le déterminant, a influencé le mécanisme syntaxique d'accord.

En résumé, les dernières études décrites ci-dessus laissent penser que l'accessibilité de la forme phonologique du mot exerce une certaine influence sur le choix d'une structure syntaxique. Dans les études de Levelt & Maasen et de Bock, la difficulté d'accès à la forme du mot retarde sa position dans la phrase, quitte à forcer le choix d'une structure syntaxique au détriment de celle qui serait privilégiée si la forme du mot était accessible. Ces résultats ne permettent toutefois pas de spécifier exactement à quel niveau s'est jouée l'influence morphophonologique. Elle peut avoir eu un impact sur l'assignation des rôles syntaxiques d'agent et d'objet (chez Bock), ou sur le choix d'une coordonnée phrastique ou syntagmatique (chez Levelt & Maasen), autrement dit sur les processus responsables d'élaborer la structure fonctionnelle de la phrase. Le positionnement des mots dans la phrase en aurait subi les conséquences indirectement. Il se peut aussi que l'accessibilité de la forme du mot ait eu un impact direct sur l'ordre des mots dans la phrase, i.e. sur l'élaboration de sa structure positionnelle. Ces études ne permettent donc pas de distinguer entre un effet phonologique sur l'élaboration de la représentation fonctionnelle de la phrase, au stade de l'encodage grammatical, et un effet phonologique sur la représentation positionnelle de la phrase, au stade de l'encodage phonologique. L'étude de Hartsuiker et al., quant à elle, ne précise pas non plus quel est le processus syntaxique influencé par la transparence du cas ; il peut, comme les auteurs en ont conclu, s'agir du mécanisme d'accord lui-même. Mais il peut tout aussi bien s'agir du processus d'assignation des fonctions syntaxiques puisque dans la condition où le cas est ambigu, le nom local est un sujet syntaxique potentiel de la phrase. Dans ce dernier cas, les erreurs d'accord ne seraient qu'un témoin indirect de cet impact, étant donné qu'une assignation incorrecte de la fonction sujet, entraîne inévitablement une erreur d'accord.

L'Expérience 3 proposée par Hartsuiker et al., menée en néerlandais, semble toutefois permettre de rejeter cette dernière hypothèse. Ils ont en effet obtenu un taux d'erreurs d'accord plus élevé lorsque le déterminant du nom sujet était non marqué pour le genre que lorsqu'il était marqué. L'ambiguïté liée au cas dans la recherche précédente n'étant plus présente, ce résultat semble bien témoigner d'un effet de l'information morphophonologique du déterminant sur le mécanisme d'accord puisqu'aucune raison apparente ne pourrait plus expliquer une erreur d'assignation des rôles syntaxiques.

Ce résultat, accompagné de ceux issus des deux autres études récentes sur l'accord grammatical rapportées dans notre revue de littérature (en italien : Vigliocco, Butterworth & Semenza, 1995, et en néerlandais : Meyer & Bock,

1998) apportent les données les plus convaincantes en faveur d'une prise en compte de l'information morphophonologique par les processus syntaxiques. L'effet observé dans ces trois études est non seulement fort net, il procure également des indications plus précises quant à l'exacte localisation de l'influence morphophonologique dans le modèle. Dans ces études, il apparaît que l'effet porte bien sur le mécanisme d'accord, et donc sur une opération syntaxique ayant lieu au stade de l'encodage grammatical.

Nous présentons dans ce chapitre la première expérience menée en français et traitant de cette question. Elle vise à déterminer dans quelle mesure deux variables morphophonologiques de genre présentes dans le syntagme nominal sujet peuvent influencer la réalisation de l'accord de l'adjectif prédicatif. La même expérience a également été menée en italien, ce qui nous permettra de discuter l'effet de ces deux variables dans une perspective interlinguistique (Franck, Vigliocco & Collina, 1999).

5.3 La morphophonologie du genre en français

Nous l'avons déjà mentionné à plusieurs reprises, le français comporte un système flexionnel complexe, et particulièrement en ce qui concerne le genre. Les manipulations que nous avons effectuées dans cette expérience exploitent deux propriétés de ce système flexionnel, à savoir l'élision de l'article et la corrélation entre la syllabe finale du nom et son genre. Voyons comment le genre est marqué sur le plan morphophonologique sur ces deux segments du syntagme nominal que sont le déterminant et le nom.

Les déterminants les plus courants en français présentent l'opposition de genre ; l'article *le/la*, le numéral *un/une*, ou encore les possessifs *mon/ma*, *ton/ta*, *son/sa*. Mais lorsque l'article et le possessif sont devant des mots commençant par une voyelle, les mêmes formes *l'*, *mon*, *ton*, *son*, précèdent les mots masculins et féminins ; elles ne sont donc plus porteuses d'information distinctive de genre. Notons également que la marque de genre disparaît dans les formes plurielles (*les*, *mes*, *tes*, *ses*).

La morphophonologie du nom est bien plus complexe étant donné la grande variabilité des marques phonologiques de genre. Les trois types de genre que l'on trouve dans la langue française (i.e. le genre conceptuel, le genre grammatical et le genre de surface) peuvent se résumer en deux groupes : le groupe des noms dont le genre est déterminé par le sexe de l'entité à laquelle le nom réfère (genre conceptuel), et le groupe des noms dont le genre est déterminé de façon arbitraire par la langue (genres grammatical et de surface). En ce qui concerne la morphologie du premier groupe, le masculin se présente comme le cas non marqué tandis que le féminin est le cas marqué ; cette différence se manifeste par la présence d'un trait supplémentaire dans les paradigmes féminins. Nous avons

discuté le rôle potentiel de ce trait supplémentaire sur l'adjectif prédicatif pour rendre compte des différences observées entre sujets féminins et masculins dans nos études précédentes (Chapitres 2 et 3). Dans ce premier groupe, le suffixe du nom suffit à en déduire le genre, autrement dit, genre et suffixe entretiennent une corrélation de 1. On distingue toutefois un grand nombre d'oppositions suffixales différentes. La plupart sont de type zéro/consonne, en ce sens que le féminin consiste en l'ajout d'une consonne par rapport à la forme au masculin (ex. époux/épouse), mais il y a aussi des variations de type voyelle nasale/voyelle orale + nasale (ex. paysan/paysanne), des variations consonantiques (ex. veuf/veuve), des variations zéro/ [ɛs] (ex. prince/princesse), des variations vocaliques + variation zéro/consonne (ex. jumeau/jumelle) ainsi que des variations consonantiques + variation de timbre (ex. fermier/fermière). Notons que ces variations constituent des groupes, et qu'au sein de chaque groupe, des consonnes et voyelles différentes peuvent être engagées.

Les noms du second groupe ne présentent pas d'opposition de genre, il n'y a donc pas de variation de genre à proprement parler. On peut néanmoins dégager certaines régularités entre le suffixe des noms et le genre grammatical. Ainsi, Dubois (1965) relève que les suffixes en *-age*, *-ment* et *-isme* sont masculins tandis que les suffixes en *-tion*, *-té* et *-ie* sont féminins. Il semble que de façon plus générale, certaines terminaisons constituent un assez bon indicateur du genre du nom qui les porte. C'est ce que révèle une analyse effectuée sur l'ensemble des noms du *Petit Larousse* par Bernard Quemada (rapportée dans Tucker, Lambert & Rigault, 1997). Par ailleurs, Tucker et al. (1997) ont testé expérimentalement la connaissance implicite qu'ont des locuteurs francophones adultes de ces régularités morphophonologiques de genre. Ils ont présenté aux participants de nombreuses listes de non-mots portant des terminaisons différentes en leur demandant d'estimer le genre de ces non-mots. Bien que les locuteurs étaient évidemment incapables d'explicitier les règles sur lesquelles ils basaient leurs jugements, ces derniers étaient dans la majorité des cas relativement proches des résultats obtenus par l'étude du *Petit Larousse*.

Ils ont ainsi montré que bon nombre de corrélations objectives entre les terminaisons et le genre des mots étaient maîtrisées par ces locuteurs. Le Tableau 15 contient les terminaisons reconnues (sur les plans objectif et subjectif) comme prédictives et non prédictives du genre. Ce tableau contient les terminaisons ayant cette valeur de prédiction à la fois à l'oral et à l'écrit.

5.4 Expérience

Deux variables morphophonologiques de genre ont été manipulées en vue d'étudier leur impact sur la réalisation de l'accord. La première concerne l'information morphophonologique présente sur l'article, selon que celui-ci est

élide ou non, la seconde concerne l'information morphophonologique présente sur le nom, selon que ce dernier porte une terminaison prédictive ou non de son genre.

Tableau 15 : Valeur prédictive du genre des terminaisons, d'après Tucker et al. (1977).

Terminaisons prédictives du genre (plus de 75%)		Terminaisons non prédictives du genre (entre 50% et 70%)
Masculin	Féminin	Masculin et Féminin
-illon	-ade	-uge
-ment	-ose	-é
-min (son in)	-ssion	-is /is/
-leau	-ation	-ique
-ton	-aison	-ège
-isme	-ite	-oie
-oir	-enne	
-eur		
-être		

5.4.1. Méthode

Participants. Quarante étudiants de l'Université catholique de Louvain ont participé à l'expérience dans le cadre des travaux pratiques de psychologie expérimentale. Ils étaient âgés de 20 à 30 ans et le français était leur langue maternelle.

Matériel. Le matériel de l'expérience se composait de 32 items expérimentaux constitués d'un préambule et d'un adjectif. Les préambules contenaient un nom sujet et un nom local. Les variables manipulées dans les préambules étaient : (1) le genre du nom sujet (Masculin vs. Féminin), (2) le genre du nom local (Masculin vs. Féminin), (3) l'information morphophonologique sur l'article du nom sujet (marqué vs. non marqué), et (4) l'information morphophonologique sur la terminaison du nom sujet (terminaison prédictive ou non prédictive).

Tous les noms possédaient un genre grammatical. La sélection des terminaisons prédictives du genre a été faite de sorte que toutes les terminaisons de cette condition étaient associées à un genre dans plus de 75% des cas, à la fois dans le code oral et le code graphique de la langue française, d'après l'étude de Tucker et al. (1977). Les terminaisons non prédictives ont été choisies de telle manière qu'elles étaient associées à un genre particulier dans moins de 70% des cas (entre 50% et 70%). La condition de non prédiction peut donc être considérée comme une condition où il n'y a pas d'association systématique entre la terminaison et le genre du mot. Les terminaisons utilisées pour construire le matériel sont celles illustrées dans le Tableau 15.

Les terminaisons des noms locaux ont été contrôlées en ce sens qu'elles présentaient toutes une forte valeur prédictive du genre, bien que nous ayons quelque peu relâché les critères de définition de cette classe en ne considérant que les résultats de Tucker et al. sur le code oral de la langue. Les items expérimentaux étaient répartis dans huit conditions expérimentales, chaque condition comprenant quatre items. Des exemples de préambules dans différentes conditions expérimentales sont présentés dans le Tableau 16. Les adjectifs ont été choisis de telle sorte que la marque de genre qu'ils portaient s'entende à l'oral. Les deux formes (masculin et féminin) étaient présentées à l'écran l'une en-dessous de l'autre, aléatoirement²⁴.

Tableau 16 : Exemples d'items dans différentes conditions expérimentales.

Condition	Item
MM, article non marqué, terminaison prédictive	L'ancêtre du violon - PRECIEUX
MF, article non marqué, terminaison prédictive	L'ancêtre de la trompette - PRECIEUX
MM, article marqué, terminaison non prédictive	Le cortège du carnaval - AMUSANT
MF, article marqué, terminaison non prédictive	Le cortège de la procession - AMUSANT
FF, article non marqué, terminaison non prédictive	L'applique de la cuisine - LUMINEUSE
FM, article non marqué, terminaison non prédictive	L'applique du couloir- LUMINEUSE
FM, article marqué, terminaison prédictive	La raison du conflit - EVIDENTE
FF, article marqué, terminaison prédictive	La raison de la bataille - EVIDENTE

Une série de 36 items de remplissage ont été construits de manière à voiler certaines régularités du matériel. Deux tiers de ces items contenaient des noms sujets animés, et des structures syntaxiques différentes ont été introduites (un tiers de sujets isolés, un tiers de relatives et un tiers du type expérimental).

Les items expérimentaux et de remplissage ont été répartis dans deux listes différentes de sorte qu'un même item (défini par son contenu lexical) n'apparaissait jamais qu'une fois par participant. Chaque participant ne recevait qu'une seule des listes, selon une répartition aléatoire. Une session commençait par une série d'exercices au cours desquels les participants se familiarisaient avec la tâche.

Procédure. Les participants ont été testés séparément en présence de l'expérimentateur. Le matériel était présenté visuellement sur écran d'ordinateur. Les deux formes de l'adjectif apparaissaient pendant 800 ms, suivies d'un écran blanc d'une durée de 500 ms puis du préambule pendant 900 ms. Il était demandé aux participants de lire silencieusement les adjectifs puis les préambules et de

²⁴ Il a été montré (Vigliocco & Franck, 1999) qu'une telle présentation de l'adjectif n'avait d'impact ni sur le résultat, ni sur la prise de conscience de l'objectif de l'expérience.

répéter ces derniers en les complétant avec l'adjectif. Les consignes insistaient pour que les participants répondent et articulent rapidement.

Catégories d'erreurs. Les erreurs ont été classées en deux catégories : (1) les erreurs d'accord en genre de l'adjectif avec le nom sujet ; seules les réponses ne comportant qu'une erreur d'accord tout en étant parfaitement correctes par ailleurs ont été répertoriées dans cette catégorie, et (2) les erreurs miscellanées ; quand une partie de la phrase manquait, quand un nom était remplacé par un autre, ou simplement quand aucune réponse n'était fournie.

Design et analyse des données. Deux analyses de variance ont été faites sur chaque type d'erreur ; la première en prenant les sujets comme facteur aléatoire (F1), la seconde en prenant les items comme facteur aléatoire (F2). L'ensemble des variables étaient réparties selon un design intra-sujet. La variable 'genre du nom local' a été considérée comme une variable intra-item puisqu'elle ne faisait varier l'item que sur le nom local en maintenant le nom sujet et l'adjectif. Les autres variables ont été traitées comme des variations inter-items.

5.4.2. Résultats

Un total de 105 (8.2%) erreurs d'accord ont été récoltées ainsi que 183 (14.3%) erreurs miscellanées. La distribution des erreurs d'accord se trouve dans le Tableau 17.

Tableau 17 : Distribution des erreurs d'accord en genre dans les différentes conditions expérimentales. Dans la colonne de gauche, la première lettre réfère au genre du nom sujet, la deuxième au genre du nom local.

	Article non marqué		Article marqué		Total
	Terminaison non prédictive	Terminaison prédictive	Terminaison non prédictive	Terminaison prédictive	
MM	6	3	1	0	10
MF	19	6	6	1	32
FM	14	14	10	5	43
FF	12	7	0	1	20
Sous-total	51	30	17	7	105
Total	81		24		105

Distribution des erreurs d'accord. L'analyse de variance pratiquée sur les données révèle un effet principal de la marque sur l'article ($F(1,39)=3.64$, $p=0.001$; $F(1,24)=14.11$, $p=0.001$) ainsi qu'un effet principal de la prédictibilité du genre du nom sujet ($F(1,39)=12.84$, $p=0.001$; $F(1,24)=4.17$, $p=0.05$). L'effet de ces deux variables sur les erreurs d'accord est illustré dans la Figure 9.

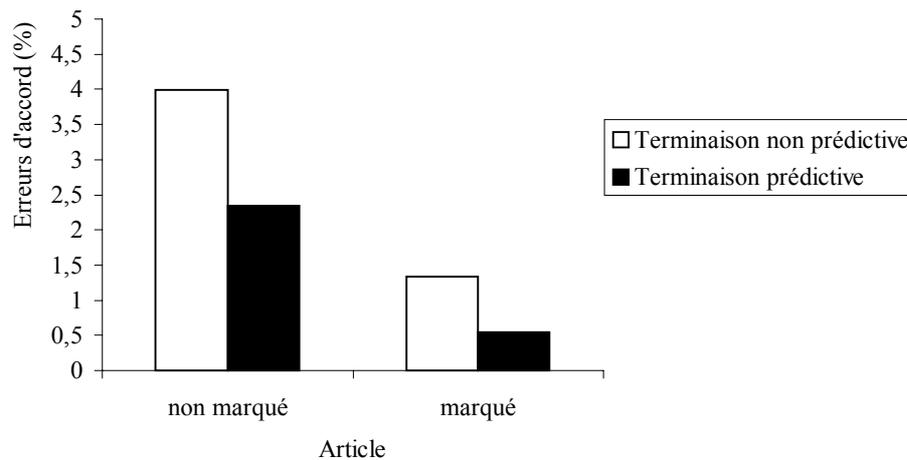


Figure 9. Distribution des erreurs d'accord en fonction de l'information morpho-phonologique sur l'article et sur le nom du syntagme nominal sujet (français).

L'interaction entre le genre du nom sujet et celui du nom local est significative, témoignant d'un effet d'attraction ($F(1,39)=12.47$, $p=0.001$; $F(1,24)=20.59$, $p<0.001$). Aucun autre effet ni interaction ne sont apparus dans l'analyse.

Distribution des erreurs miscellanées. Le seul effet significatif à rapporter est un effet principal du genre de N1, avec plus d'erreurs lorsque N1 est masculin que lorsqu'il est féminin ($F(1,39)=18.14$, $p<0.000$; $F(1,24)=9.11$, $p=0.006$).

5.4.3. Discussion

Les résultats obtenus indiquent un effet des deux variables morphophonologiques manipulées : les participants ont produit plus d'erreurs d'accord de l'adjectif prédicatif avec le sujet de la phrase lorsque le déterminant était élide que lorsqu'il était complet, et lorsque la terminaison était peu prédictive du genre que lorsqu'elle était fortement prédictive. Par ailleurs, nous avons également répliqué l'effet d'attraction classique ; un plus grand nombre d'erreurs sont apparues lorsque le nom sujet et le nom local étaient de genres différents. Les noms sujets féminins ont engendré un plus grand nombre d'erreurs d'accord que les noms masculins, comme dans les études précédentes, bien que cette différence ne soit pas significative dans la présente recherche.

Par cette expérience, nous avons montré que deux informations morphophonologiques de genre, l'une présente sur l'article, l'autre sur le nom, sont utilisées par les locuteurs francophones pour mener à bien l'accord de l'adjectif prédicatif. Avant d'aborder l'interprétation théorique de ces résultats, la recherche menée en italien est brièvement présentée.

5.5 Aperçu de la recherche en italien

5.5.1. La morphophonologie du genre en italien

La morphologie du genre en italien est considérablement plus simple que celle du français. L'opposition de genre se marque sur l'article ; on distingue deux articles masculins, *il* et *lo*, et un article féminin, *la*. Comme en français, l'article est élide (*l'*) lorsqu'il précède un nom commençant par une voyelle.

En ce qui concerne le genre des noms, la morphophonologie est la même selon que le nom a un genre conceptuel ou un genre grammatical. Il existe trois flexions nominales de genre : *-a*, *-o* et *-e*. La terminaison suffixale du nom est, contrairement au français, porteuse d'information morphophonologique transparente dans la plupart des cas. En effet, 87 % des noms possèdent soit le suffixe *-a* soit le suffixe *-o*, tous deux étant strictement réservés à un seul genre, i.e. respectivement le féminin et le masculin. Autrement dit, la toute grande majorité des noms possèdent un genre prédictible à partir de leur suffixe. La terminaison en *-e* touche 10 % des noms et est, quant à elle, ambiguë puisqu'elle peut être associée aux deux genres. Enfin, les 3 % restants concernent des mots issus de langues étrangères et peuvent ainsi posséder des suffixes divers, non prédictifs pour l'italien.

5.5.2. Expérience

Les mêmes variables que celles manipulées en français ont guidé la construction des 32 items expérimentaux en italien. Deux exemples d'items sont présentés dans le Tableau 18.

Un total de 48 participants ont été testés sur ce matériel au moyen de la même procédure expérimentale que celle utilisée en français.

Tableau 18 : Exemples d'items dans deux conditions expérimentales dans la recherche en italien.

Condition	Item
MM, article marqué, terminaison prédictive	Il premio del torneo - AMBITO/A <i>Le prix du tournoi-DESIRE/E</i>
FM, article non marqué, terminaison non prédictive	L'indagine del governo - INGIUSTO/A <i>L'enquête du gouvernement - MALHONNETE</i>

Les erreurs d'accord en genre produites par les participants constituent 5.2% de la totalité des réponses. Les résultats indiquent un effet d'attraction, plus d'erreurs ayant été produites lorsque le nom sujet et le nom local étaient de genres différents. Une seule des variables morphophonologiques est apparue comme facteur déterminant des erreurs d'accord en genre : la prédiction du suffixe. L'information véhiculée par l'article n'a entraîné aucun effet. L'impact de ces deux

variables morphophonologiques sur les erreurs d'accord est illustré par la Figure 10.

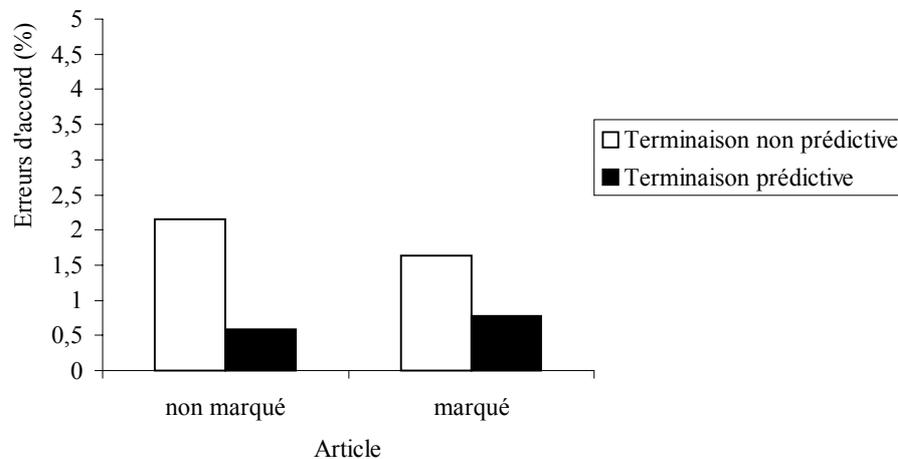


Figure 10. Distribution des erreurs d'accord en fonction de l'information morphophonologique sur l'article et sur le nom du syntagme nominal sujet en italien.

5.6 Discussion générale

L'étude réalisée en français a révélé l'importance de l'information morphophonologique présente sur le syntagme nominal sujet dans la construction de l'accord en genre de l'adjectif prédicatif. A la fois l'information sur l'article et celle sur le nom sujet semblent jouer un rôle dans la réalisation de ce mécanisme syntaxique. L'absence de ces informations s'est avérée être une source significative d'erreurs d'accord, les deux informations ayant un effet indépendant sur le processus. Les locuteurs italiens se sont, eux aussi, montrés sensibles à l'information morphophonologique. Toutefois, seule celle présente sur le nom sujet semble être utilisée pour construire l'accord. En l'absence de cette information, un plus grand nombre d'erreurs ont été commises, tandis que l'absence d'information sur l'article n'a pas modifié le pattern d'erreurs.

L'observation d'un effet morphophonologique sur le mécanisme d'accord en genre en français et en italien confirme les résultats déjà obtenus en italien par Vigliocco et al. (1995) pour l'accord en nombre, ainsi que ceux en néerlandais récoltés par Meyer & Bock (1998) en ce qui concerne l'accord en genre du pronom relatif sujet, et par Hartsuiker et al. (in prep.) pour l'accord en nombre entre le sujet et le verbe.

Dans les paragraphes qui suivent, nous examinons trois cadres théoriques (voir Franck, Collina & Vigliocco, 1999) avancés pour rendre compte de cet effet, ainsi que les hypothèses relatives aux différences interlinguistiques rapportées.

5.6.1. Interprétation théorique de l'effet morphophonologique

Les résultats apportés par ces recherches sont incompatibles avec le modèle strictement séquentiel développé par Garrett (1993) selon lequel les mécanismes syntaxiques tels que l'accord, ayant lieu au stade de l'encodage grammatical, sont aveugles aux représentations morphophonologiques. Dans ce modèle, l'encodage grammatical opère sur la base d'un minimum d'informations en vertu du principe de cloisonnement informationnel. Il semble toutefois que l'effet morphophonologique soit conciliable avec une telle position théorique, à condition que l'on ajoute au modèle une composante de *contrôle*. Une première proposition dans ce sens a été avancée par Butterworth (1980). Selon lui, chaque composante du système de production est soumise à un module de contrôle responsable de vérifier l'adéquation de ses représentations finales. Il existerait donc un module responsable du contrôle lexical, un autre responsable du contrôle syntaxique, etc. Cette conception implique une forte redondance de l'information puisque, pour être capables de contrôler l'output d'une composante, les modules doivent eux aussi posséder les représentations de ces composantes.

D'autres auteurs ont ainsi postulé l'existence d'une composante de contrôle unique, dite d'édition préarticulatoire, dont la fonction est de détecter les erreurs de toutes natures, lors de la phase préarticulatoire de la production, i.e. lorsque tous les aspects proprement linguistiques ont été spécifiés (Motley & Baars, 1979 ; Motley, Baars & Camden, 1981). En utilisant une technique destinée à susciter la production de lapsus, Motley et al. (1981) ont montré que des permutations de phonèmes sur des non-mots étaient plus courantes lorsqu'elles donnaient lieu à des séquences syntaxiques correctes que lorsqu'elles généraient des séquences incorrectes. Par exemple, la présentation des non-mots *ker hane* entraîne plus d'erreurs que le doublet *ker hame*. Une permutation dans le premier doublet produit une séquence syntaxiquement acceptable (*her cane*, sa canne), ce qui n'est pas le cas du second (*her came*, *sa vint). Les auteurs en déduisent qu'il existe un mécanisme préarticulatoire, mis en œuvre après l'encodage phonologique, capable de repérer une erreur syntaxique et de la renvoyer à son système de traitement.

Un résultat intéressant non discuté par Motley et al. est que la condition où la permutation donne lieu à une erreur syntaxique entraîne plus d'erreurs partiellement corrigées que celle ne menant pas à une erreur syntaxique. Autrement dit, il semble qu'en cours même d'articulation (et pas seulement en phase préarticulatoire) les locuteurs soient capables de détecter une anomalie syntaxique dans leur production. Cette observation peut par contre s'interpréter dans le cadre du modèle proposé par Levelt (1989) selon lequel une double boucle de contrôle assure la vérification des énoncés : au niveau préarticulatoire (boucle interne) et en cours d'articulation (boucle externe). L'originalité de cette approche est l'hypothèse que le contrôle ou *monitoring*, est assuré par le système de compréhension du langage d'autrui et ne nécessite donc pas d'ajouter une nouvelle

composante au modèle. Une expérience de Lackner & Tuller (1976), dans laquelle les participants devaient rapidement répéter des suites de syllabes tout en essayant de détecter leurs erreurs, plaide en faveur de ce double système de monitoring : lorsqu'il s'entendaient parler, ils détectaient non seulement plus d'erreurs, mais le temps de détection était également plus long que quand ils ne s'entendaient pas parler. Cela semble indiquer que le système de monitoring opère à la fois avant et pendant l'articulation, et qu'il est assuré par le système de compréhension du langage d'autrui.

Dans ce cadre, l'effet morphophonologique rapporté sur la réalisation de l'accord peut s'expliquer par le fait que, lors de la phase préarticulatoire, une erreur d'accord est détectée et renvoyée pour être corrigée au système syntaxique, *après* que la morphophonologie du mot ait été spécifiée. En arrivant au stade préarticulatoire, un même nombre d'erreurs d'accord auraient été commises, qu'elles soient porteuses ou non d'information morphophonologique. Toutefois, les erreurs dans les phrases dont le sujet est spécifié sur le plan morphophonologique seraient plus facilement détectées et renvoyées pour correction que celles dans les phrases ne comportant pas une telle spécification. L'effet rapporté *mime un effet de rétroaction* des représentations morphophonologiques sur le traitement syntaxique et donne ainsi l'illusion d'une interaction entre ces deux composantes, alors que cet effet est médiatisé par une composante de contrôle externe à ces deux traitements. L'effet morphophonologique dans ce modèle est illustré par la Figure 11. En gris clair, l'erreur n'est pas détectée par le système de monitoring parce que ni l'article, ni le nom, ne portent d'information morphophonologique. En noir, l'erreur est détectée par le moniteur et renvoyée pour correction à l'encodage grammatical, grâce à la présence d'information morphophonologique soit sur le nom, soit sur l'article.

Cette optique permet de concilier nos résultats et l'hypothèse de séquentialité unidirectionnelle du modèle de production modulaire : le traitement syntaxique initial a effectivement été opéré *avant* le traitement morphophonologique et sans être influencé par lui. Sont-ils toutefois compatibles avec l'hypothèse du cloisonnement informationnel ? La réponse à cette question repose sur la définition de l'output que le système de monitoring procure au système syntaxique. Si le monitoring, assuré par le système de compréhension, est lui-même composé de différents systèmes cloisonnés, comme le suggèrent certains modèles de la compréhension de phrases (ex. Frazier, 1990), c'est la composante de traitement syntaxique qui détecte l'erreur et en assure le renvoi à l'encodage grammatical. Dans ce cas, les représentations qu'il fournit sont de nature syntaxique et le cloisonnement informationnel est préservé. Si par contre, les différentes sources d'information, i.e. conceptuelle, syntaxique et morphophonologique, interagissent lors de la compréhension (ex. Bates & MacWhinney, 1989) alors l'output du système de contrôle utilisé par le système de production syntaxique sera de nature 'multi-informationnelle', et l'hypothèse de cloisonnement en production n'est plus

tenable. On le voit, l'hypothèse du cloisonnement informationnel de la production syntaxique repose sur celle du cloisonnement de la compréhension syntaxique, si l'on veut pouvoir rendre compte de l'effet morphophonologique rapporté dans ce cadre.

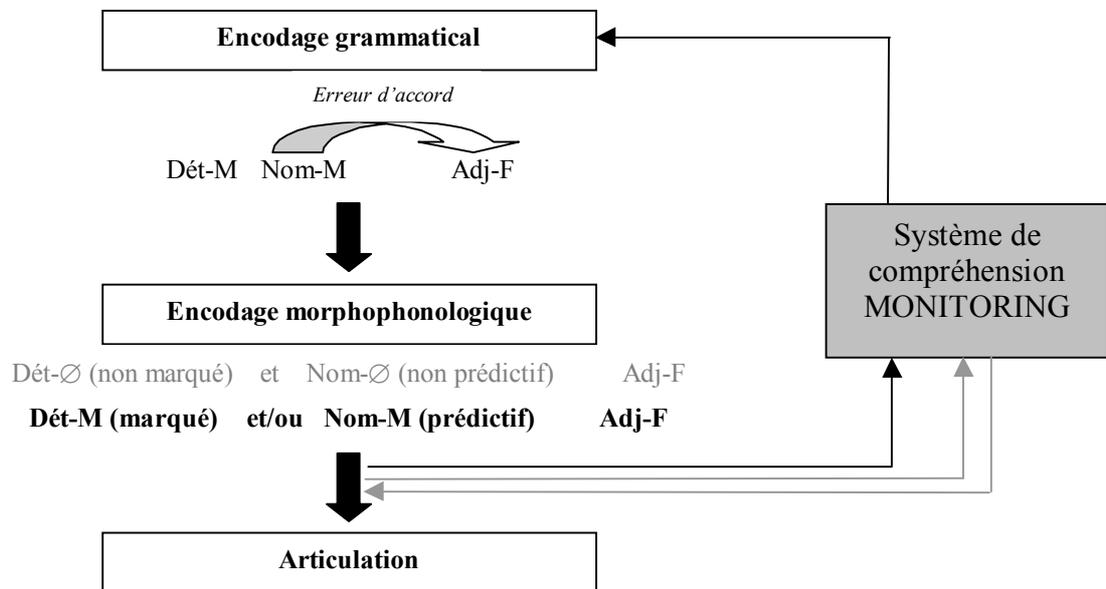


Figure 11. Le contrôle de l'accord en genre par le système de monitoring, selon que le nom et l'article ne portent pas d'information morphophonologique (gris clair), ou que l'un des deux en est porteur, inspiré de Levelt (1989).

Une conception interactive de la production (ex. Dell, 1986 ; Bates & MacWhinney, 1989 ; Stemberger, 1985) fait, elle aussi, l'hypothèse que l'output de chaque niveau de production est soumis à une forme de contrôle. Ce contrôle est toutefois assuré au sein même du système de production, sans qu'il ne soit nécessaire de postuler l'intervention de systèmes spécifiques de contrôle ou d'une composante de monitoring assurée par le système de compréhension. Le contrôle provient des *interactions entre les niveaux* et de la possibilité qui s'en suit de *rétroaction* d'un niveau sur celui qui le précède. Dans ce cadre, on peut avancer l'hypothèse que les traitements syntaxique et morphophonologique sont opérés en parallèle, chacun pouvant influencer l'autre par activation bidirectionnelle des connexions au sein du réseau. En ce qui concerne la réalisation de l'accord dans les différentes recherches rapportées, les éléments impliqués sont activés, chacun ayant des caractéristiques particulières concernant sa fonction dans la phrase, les relations qu'il entretient avec les autres, ses spécifications syntaxiques de genre et de nombre, et ses spécifications morphophonologiques de genre et de nombre.

La Figure 12 illustre un tel réseau et les activations entre les différentes unités impliquées pour la production de l'accord dans la phrase expérimentale *L'ancêtre (du violon est) précieux*. Les unités représentées dans la première couche du réseau

sont de nature syntaxique fonctionnelle (Déterminant, Nom, Adjectif), celles de la seconde couche sont de nature lexico-syntaxique (Féminin, Masculin), celles de la troisième couche sont de nature morphophonologique. Ces trois types de représentations sont reliées les unes aux autres et sont supposées être activées en parallèle. Bien que l'activation se propage initialement de haut en bas (du sens vers la forme), elle peut également se propager de bas en haut par rétroaction. Dans la Figure 12, le nom sujet possède une terminaison prédictive du genre masculin (*-être*) ; une fois la forme phonologique du mot *ancêtre* activée, cette information morphophonologique que contient la terminaison active par rétroaction le genre masculin qui augmente à son tour l'activation du masculin de l'adjectif, tout en inhibant le nœud féminin (pour rappel, les connexions entre les alternatives du même type sont inhibitrices).

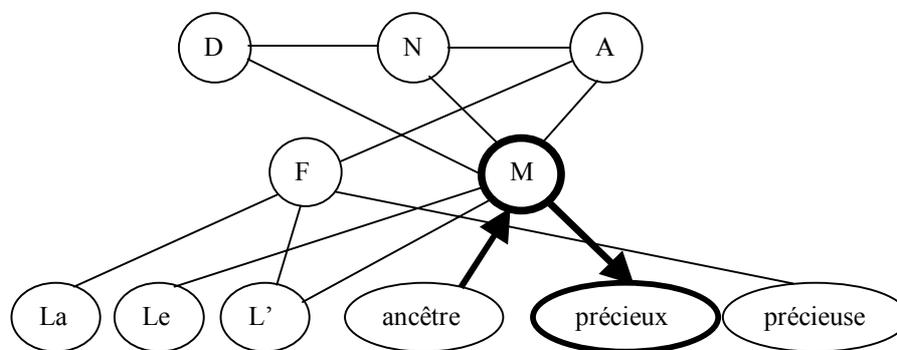


Figure 12. Rétroaction du niveau morphophonologique sur le niveau syntaxique dans le réseau activé pour la phrase *L'ancêtre (du violon est) précieux*, inspiré du modèle de Dell (1986).

La Figure 13 illustre la rétroaction provenant de la marque morphophonologique de genre présente sur l'article non élide du nom sujet sur l'activation du masculin de l'adjectif dans la phrase expérimentale *Le cortège (du carnaval est) amusant*.

Il est intéressant de noter que les études en néerlandais (Meyer & Bock, 1998) et en italien (Vigliocco et al., 1995), qui ont pourtant mis en évidence un effet de la présence d'une marque morphophonologique sur le sujet, n'ont rapporté aucun effet de cette marque sur le nom local. Par contre, un effet de la marque sur le nom local a été rapporté en néerlandais par Hartsuiker et al. (in prep.). Les propriétés morphophonologiques du nom sujet auraient un impact plus important que celles du nom local, étant donné la relation syntaxique particulière unissant le nom sujet à l'adjectif, ce qui n'est pas le cas du nom local. Même s'il est probable que les propriétés du nom local rétroagissent sur la sélection de la forme de l'adjectif parce qu'elles activent un nœud partagé, cette rétroaction est d'autant plus forte lorsqu'elle provient des propriétés du nom sujet. On s'attend par conséquent à ce

qu'un effet morphophonologique du nom local prédise l'observation d'un effet du nom sujet, mais non l'inverse.

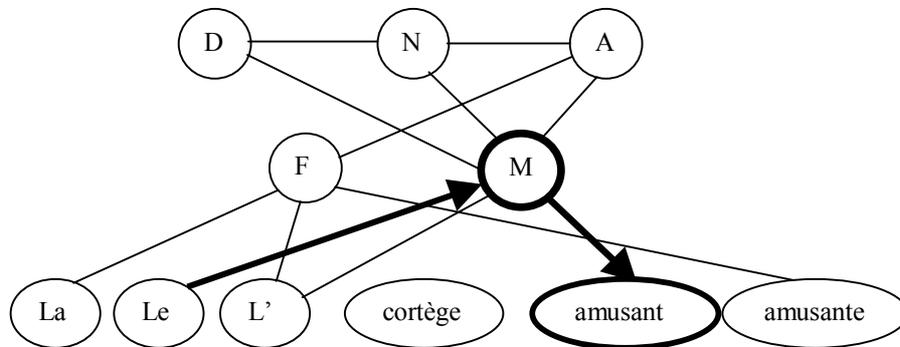


Figure 13. Rétroaction du niveau morphophonologique sur le niveau syntaxique dans le réseau activé pour la phrase *Le cortège (du carnaval est) amusant*, inspiré du modèle de Dell (1986).

Dans les Figure 12 et Figure 13 représentées ci-dessus, nous avons simplement illustré le mécanisme de l'impact morphophonologique sur le choix du genre de l'adjectif. Ce mécanisme repose sur l'hypothèse faite quant au *décours temporel* du traitement des différentes composantes de la production : l'effet morphophonologique est expliqué par le fait que les différents niveaux sont activés en même temps, et que l'activation entre les niveaux est bidirectionnelle. Cependant, ce schéma ne reflète pas l'aspect lié à la *quantification* de l'effet de cette variable. L'approche de Dell de la production au niveau lexical, mais aussi l'approche interactive de la compréhension de phrases basée sur le concept de satisfaction des contraintes (ex. MacDonald, Pearlmutter & Seidenberg, 1994, Trueswell, Tanenhaus & Garnsey, 1994), ou encore le modèle de compétition (Bates & MacWhinney, 1989) avancent un cadre théorique intéressant quant aux facteurs, aussi appelés *contraintes* (modèles de satisfaction des contraintes) ou encore *indices* (modèle de compétition) déterminant l'influence relative d'une information dans ce réseau interactif d'activations. Ces contraintes concernent essentiellement la fréquence : la fréquence d'apparition des différentes unités dans la langue, mais aussi la fréquence de l'association entre ces unités. Ces deux fréquences déterminent l'importance d'une information, et donc la probabilité qu'elle joue un rôle dans l'utilisation de la langue. Comme la validité des indices varie d'une langue à l'autre, les hypothèses les concernant fournissent un cadre d'interprétation fort pertinent des différences entre les langues. Elles sont discutées dans la section suivante.

Jusqu'ici, nos résultats ont été interprétés dans le cadre de la question des relations entre deux niveaux de traitement : le niveau syntaxique et le niveau morphophonologique. Nous avons montré que la perspective séquentielle de la production pouvait rendre compte de l'effet morphophonologique, pour autant

qu'elle postule l'existence d'un monitoring préarticulatoire, assuré par le système de compréhension. La perspective interactive, quant à elle, explique cet effet par un contrôle direct du niveau morphophonologique sur le niveau syntaxique, via les connexions bidirectionnelles entre les deux. Ces deux cadres interprétatifs reposent tous les deux sur l'hypothèse, originellement proposée par Garrett (1976), qu'il existe bel et bien un niveau de traitement morphophonologique, autrement dit, que le traitement syntaxique se décompose en deux niveaux : le premier dit fonctionnel, le second dit positionnel.

L'argument principal avancé par ce dernier provient de l'observation de deux types d'erreurs de substitution ayant chacune des caractéristiques propres. En ce qui concerne les erreurs du premier type, elles impliquent des unités de la classe ouverte, l'échange porte sur des unités de la même catégorie grammaticale, il n'y a pas de similarité phonologique entre les unités échangées, et les unités impliquées ont tendance à faire partie d'une entité syntaxique large (la proposition). Les erreurs du second type consistent essentiellement en erreurs dites de *stranding* : deux racines lexicales sont échangées tandis que leurs suffixes sont laissés en place. Ces erreurs ne respectent pas la catégorie grammaticale des unités échangées, elles impliquent souvent deux racines lexicales proches sur le plan phonologique, et elles se situent au sein d'une entité syntaxique plus petite (le syntagme). De ces observations, Garrett déduit que la syntaxe se divise en un niveau responsable de la syntaxe des mots de la classe ouverte (assignation des fonctions, élaboration d'un cadre hiérarchique spécifiant les relations syntaxiques entre ces unités), dit fonctionnel, et un niveau responsable de la syntaxe des mots de la classe fermée, la morphosyntaxe, dit positionnel. Or, comme ce dernier est apparemment influencé par des facteurs phonologiques, les interférences entre ces deux types de représentations (morphosyntaxiques et phonologiques) sont considérées comme un indicateur qu'elles sont traitées au même niveau, que l'on peut appeler l'encodage morfo-phonologique. Cette décomposition est illustrée dans la Figure 1 du premier chapitre.

En partant lui aussi de l'observation d'un corpus d'erreurs, Stemberger aboutit toutefois à des conclusions différentes. Il suggère tout d'abord qu'il n'y a pas de similarité phonologique entre les racines lexicales dans les erreurs morphosyntaxiques de *stranding* ou plus généralement dans les erreurs portant sur les mots de la classe fermée. Par ailleurs, il montre que les erreurs d'échange sur des mots de la classe ouverte ayant une forte fréquence d'occurrence dans la langue, comme c'est le cas des mots de la classe fermée, ne présentent pas non plus de similarité phonologique. Il explique cela par le fait qu'un mot de haute fréquence entretient de fortes connexions avec sa représentation phonologique, ce qui réduit la probabilité d'une erreur de ce type. L'analyse de l'occurrence des erreurs du second type menée par Stemberger montre, contrairement à celle de

Garrett, que ces erreurs sont fortement contraintes sur le plan syntaxique²⁵. Il en déduit que le traitement des morphèmes grammaticaux est assuré conjointement par le lexique, dans lequel ils sont stockés au même titre que les morphèmes lexicaux, et par la syntaxe, contenant les règles propres à leur utilisation.

Dans ce cadre théorique, l'effet que nous rapportons, que l'on ne doit plus qualifier de morphophonologique mais de morphologique, relève des propriétés internes de la syntaxe, ou plutôt des relations entre les composantes syntaxique et lexicale. Les différentes terminaisons nominales ainsi que les différentes formes de l'article sont représentées dans le lexique, au même titre que les unités lexicales. La Figure 14 illustre l'activation dans le réseau lexical impliqué lors de la production de l'item en italien *Il premio (del torneo è) ambito*. La flexion *-o* est représentée par un nœud propre, et son activation lors de l'activation de *premio* renforce l'activation de la forme au masculin de l'adjectif *ambito*. Comme l'indique le schéma, la flexion *-o* possède une représentation propre et partagée par l'ensemble des unités pouvant la porter.

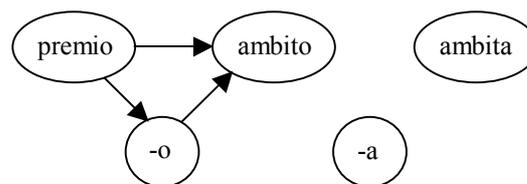


Figure 14. Activation de l'adjectif masculin au sein du réseau lexical, inspiré du modèle de Stemberger (1985).

Une question se pose toutefois quant à la représentation lexicale des terminaisons nominales que nous avons manipulées : alors qu'il s'agit de flexions en italien, leur statut en français n'est pas clair. Bien qu'il ne s'agisse pas de flexions au sens strict du terme, puisque ces terminaisons ne viennent pas modifier le mot mais en font partie intégrante, elles sont néanmoins porteuses d'une information syntaxique. Nous suggérons que ces terminaisons doivent donc, elles aussi, être représentées dans le lexique, de manière similaire aux flexions, comme représentations partagées entretenant des relations privilégiées avec les autres unités portant le même genre.

En ce qui nous concerne, nous pensons qu'il est actuellement impossible de trancher entre les deux conceptions du statut du traitement morphologique dans un modèle de production. Nous pensons toutefois que la distinction entre deux niveaux de traitement syntaxique est pertinente pour d'autres raisons, comme il sera abordé dans le Chapitre 6. Pour cette raison, nous avons choisi d'adopter la terminologie proposée par Garrett et largement partagée par les psycholinguistes ;

²⁵ Ce point est discuté lors de la présentation du modèle de Stemberger dans le Chapitre 1.

nous continuerons donc de nous référer aux processus étudiés sous le terme de processus morphophonologiques.

5.6.2. Hypothèses explicatives des différences interlinguistiques

Bien que les deux derniers cadres d'interprétation abordés postulent l'intervention de niveaux de représentation différents, i.e. syntaxique et morphophonologique pour le premier, syntaxique et lexical pour le second, tous les deux adoptent les mêmes principes de fonctionnement. Dans ces modèles, la production du langage se fait par un mécanisme de propagation de l'information, par activation au sein d'un réseau de connexions bidirectionnelles. La façon dont l'activation se propage dans le réseau, et par conséquent l'impact d'une information sur un traitement particulier, repose sur le poids relatif des unités impliquées et des connexions qu'elles entretiennent avec les autres unités du réseau. En nous basant sur les principes proposés dans le cadre des modèles interactifs de la compréhension de la phrase (développés par les équipes de MacDonald, Tanenhaus, Trueswell et autres), et en particulier du modèle de compétition de Bates & MacWhinney, nous envisageons ici les contraintes impliquées dans la réalisation de la tâche expérimentale de complètement de phrases, et dans l'apparition de l'effet morphophonologique rapporté, ainsi que leur potentiel explicatif pour rendre compte des différences entre le français et l'italien.

Une contrainte essentielle de ces modèles concerne la *fréquence* d'apparition des mots dans la langue. Dans les modèles interactifs, la fréquence d'utilisation d'une unité dans la langue détermine son seuil d'activation au repos. Lorsque deux unités entrent en compétition, la plus fréquente a plus de chances de remporter la place puisqu'elle est plus facilement activée. En ce qui concerne la réalisation de l'accord dans une phrase comme *L'ancêtre du violon est précieux*, les fréquences d'utilisation respectives de *précieux* et de *précieuse* jouent donc un rôle important dans la sélection finale de la forme de l'adjectif. Cette contrainte peut être représentée dans le modèle par un seuil d'activation plus bas de l'unité la plus fréquemment utilisée. Cet effet de fréquence peut parfaitement rendre compte de l'observation que nous avons faite dans l'ensemble de nos études sur l'accord en genre en français (voir Chapitres 2 et 3, et la présente étude) selon laquelle un plus grand nombre d'erreurs sont produites dans les phrases ayant un sujet féminin. Ces erreurs consistent à sélectionner l'adjectif masculin plutôt que l'adjectif féminin. La forme masculine est en effet la forme dite 'par défaut' ; c'est elle qui l'emporte dans les conjonctions (ex. *le mari et sa femme sont prétentieux-M*), lors de l'utilisation du pronom impersonnel 'on' (ex. *on est prétentieux d'en demander autant*) ou encore dans les constructions impersonnelles (ex. *il est prétentieux de penser que etc.*). Elle est donc plus fréquemment utilisée que la forme féminine.

Par ailleurs, cette contrainte de fréquence ne pèse pas uniquement sur les nœuds, elle pèse également sur les connexions entre les nœuds. Le poids de ces connexions est déterminé par la corrélation, dans la langue, entre les deux unités,

autrement dit, par la fréquence de leur co-occurrence. Par exemple, l'article *le* est systématiquement associé au masculin. L'article élide, par contre, peut être associé au masculin et au féminin. Il n'y a donc pas d'association privilégiée entre l'article élide et l'un des deux genres ; les deux connexions existent au sein du réseau. Toutefois, on peut imaginer que si la majorité des mots de la langue commençant par une voyelle sont masculins, l'article élide sera plus souvent utilisé en connexion avec le masculin qu'avec le féminin ; par conséquent, la connexion de l'article élide avec le masculin sera plus forte et aura donc un seuil d'activation inférieur à la connexion avec celle entre l'article élide et le féminin.

Le modèle de compétition a introduit la notion de *validité* d'une information (d'un indice), comme nous l'avons présenté dans l'Introduction générale. La validité d'un indice dépend de la force de la connexion, et donc de la fréquence de co-occurrence entre cet indice et ce qu'il est supposé indexer. Il faut cependant faire la distinction entre la validité de l'indice et son *utilisation* effective. Nous avons à plusieurs reprises insisté sur le fait qu'un input maximal de la composante syntaxique permet de limiter les erreurs liées à la perte d'information. Toutefois, cette prise en compte de différentes sources d'information représente un coût qu'il est également important de minimiser. Comme l'ont suggéré Bates et MacWhinney, l'utilisation d'un indice ne repose pas que sur sa validité, elle repose également sur son coût. En présence de plusieurs indices ayant une grande validité, il se peut donc que les locuteurs choisissent de ne tenir compte que de certains d'entre eux pour minimiser le coût de cette prise en compte.

Comment le cadre théorique du modèle de compétition peut-il rendre compte de ce 'trade-off' entre input maximal et coût ? Nous avons montré que les locuteurs italiens n'utilisent que l'information suffixale ; aucun effet de l'information sur l'article n'a été observé. Pourtant, ces indices morphophonologiques ont tous les deux une grande validité : la terminaison du nom est hautement prédictive de son genre, comme l'est l'article qui le précède. Les deux indices étant valides, les locuteurs italiens semblent faire le choix d'en favoriser un par rapport à l'autre, et ceci pour minimiser les coûts liés au traitement de ces informations. L'indice privilégié est le suffixe du nom. Pourquoi ? Sans doute parce que sa validité est meilleure que celle de l'article. La validité d'un indice ne repose en effet pas seulement sur la fréquence de son association avec cette fonction, i.e. le nombre de fois que cet indice est présent lorsque la fonction est activée, elle repose également sur sa *fiabilité*, i.e. le nombre de fois que cet indice est associé à cette fonction par rapport au nombre de fois où il est associé à une autre. Les terminaisons *-a* et *-o* ont une fiabilité maximale puisqu'elles sont systématiquement associées à un et un seul genre. L'article, quand il est complet, a lui aussi une fiabilité maximale, mais cela n'est pas toujours le cas puisqu'il peut parfois prendre la forme élide, et être ainsi associé aux deux genres.

La 'stratégie' d'utilisation de l'information morphophonologique par les locuteurs italiens semble donc optimale en termes d'efficacité : ils font usage d'une information extrêmement valide et fiable, i.e. le suffixe du nom, et peuvent par conséquent se permettre de négliger une information qui, certes, est valide mais n'est pas directement utile, qui est moins fiable, et qui donc ne ferait qu'alourdir le coût du traitement syntaxique. Notons qu'une mise à l'épreuve de cette hypothèse explicative passe par l'analyse des propriétés objectives de la langue. En particulier, il s'agira de démontrer que la fiabilité du suffixe est plus grande que celle de l'article. Pour ce faire, il faudra montrer que l'ambiguïté de genre véhiculée par le suffixe *-e*, touchant 13% des noms de la langue est, comme nous l'avons supposé, moindre que celle véhiculée par l'article élide, dont les proportions par rapport aux articles complets restent à déterminer.

En français par contre, nous avons rapporté que les deux indices morphophonologiques favorisent une réalisation correcte de l'accord. Un examen détaillé des erreurs dans les différentes conditions expérimentales semble toutefois indiquer que l'information sur l'article (81 erreurs dans la condition non marquée contre 24) joue un rôle plus important que celle sur la terminaison du nom (68 erreurs dans la condition non prédictive contre 37). Par ailleurs, dans une première recherche que nous avons menée en manipulant les mêmes variables, seul l'effet de la marque sur l'article était significatif. Le 'seuil de prédictibilité' des terminaisons dans cette recherche avait été fixé plus bas que celui de la seconde étude, à savoir à 70%, d'après les tables de Tucker et al. (1977) : les terminaisons prédictives avaient une fréquence de co-occurrence avec un genre de plus de 70%, tandis que cette fréquence était de moins de 70% pour les terminaisons non prédictives. Aussi, les terminaisons avaient été sélectionnées uniquement sur la base des tables pour le code écrit, tandis que nous avons tenu compte également du code oral dans la seconde recherche. Autrement dit, l'effet de la terminaison obtenu dans la présente recherche repose fortement sur le choix de critères suffisamment stricts de délimitation entre terminaisons prédictives et non prédictives.

Les locuteurs francophones semblent donc, eux aussi, suivre une stratégie visant à maximiser l'efficacité des opérations syntaxiques. En présence de deux indices morphophonologiques, ils privilégient le plus valide, à savoir l'information sur le déterminant. La validité de cet indice est en effet maximale quant il est complet puisque dans ce cas la corrélation entre l'article et le genre est de 1. Toutefois, bien qu'il soit valide, il n'est pas entièrement fiable puisque cette information peut être absente : lorsque l'article est élide, mais aussi lorsqu'il est au pluriel (*les*), ou bien lorsqu'il s'agit d'un déterminant quantificateur (ex. *chaque, deux, quelques, peu de, bon nombre*, etc.). Un autre indice est donc également pris en considération, i.e. la terminaison du nom. Cet indice est toutefois moins valide, puisque l'association entre la terminaison et le genre n'est pas maximale mais varie entre 75% et 100%. Elle est aussi moins fiable, puisque la langue contient assez peu de ces terminaisons prédictives, comme l'indique le Tableau 15.

Etant donné la faible prédiction que procurent ces terminaisons, on pourrait même se demander pourquoi les locuteurs francophones s’embarrassent d’une telle information au risque d’augmenter le coût du traitement syntaxique sans qu’il n’y ait de réel bénéfice. Outre les cas où le déterminant n’est pas porteur d’information morphophonologique, la prise en compte de l’information nominale est tout à fait pertinente lors de l’utilisation de noms ayant un genre conceptuel. Ces noms, contrairement à ceux ayant un genre grammatical, sont en effet porteurs de suffixes flexionnels constituant d’excellents indicateurs du genre (ex. *-eur/-euse, -ant/-ante, -ien/ienne*). Il semble donc tout à fait stratégique que les locuteurs francophones utilisent l’indice lié à la terminaison du nom pour construire l’accord, puisque dans un certain nombre de situations, il peut s’avérer extrêmement valide et fiable.

Enfin, Bates & MacWhinney ont suggéré que la *perceptibilité* d’un indice influence également son utilisation. Cette hypothèse est avancée pour rendre compte de la compréhension du langage : si un indice n’est pas audible, il va de soi que le compreneur peut difficilement l’utiliser. Mais la perceptibilité d’un indice se trouve elle aussi ‘encodée’ dans le réseau lexical, que les modèles interactifs supposent être commun aux systèmes de production et de compréhension (ex. MacKay, 1987). Le réseau lexical se forme progressivement au cours de l’apprentissage et est mis à jour tout au long de la vie. Un individu normal est à la fois confronté à du langage oral et écrit, ainsi qu’à des situations de production et de compréhension. On peut donc faire l’hypothèse qu’une information morphophonologique de genre qui n’est pas audible (comme c’est souvent le cas sur les adjectifs et les participes ; ex. gentil/gentille, parti/partie) est moins souvent associée à la propriété syntaxique de genre et que, par conséquent, la connexion entre les deux unités est plus faible. Autrement dit, la perceptibilité d’un indice pourrait influencer la force des connexions entre cet indice et les nœuds qu’il indexe au niveau lexical.

En faveur de cette hypothèse, plusieurs observations sur la production écrite de l’accord sujet-verbe en français suggèrent que les erreurs portent principalement sur des verbes dont le nombre n’est pas marqué à l’oral (ex. Hupet, Schelstraete, Demaeght & Fayol, 1996). Par ailleurs, Barbaud, Ducharme & Valdois (1982) rapportent une tendance à la féminisation des noms commençant par une voyelle au sein des locuteurs français du Canada. La mauvaise perceptibilité du genre de ces noms, dont le déterminant est souvent ambigu pour le genre (*l’, son, [set]* pour cet/cette), pourrait être à l’origine de la modification de la représentation lexicale de leur genre. Bien que l’article élide *l’* soit utilisé également pour les noms féminins et masculins, ce n’est pas le cas des deux autres déterminants. Le possessif *son* précède tous les mots masculins. Par contre, bien que les mots féminins commençant par une consonne soient précédés du possessif *sa*, ceux commençant par une voyelle sont, comme les mots masculins, précédés de *son*. Le possessif *son* est donc ambigu quant à la marque de genre qu’il porte puisqu’il peut

précéder à la fois un mot masculin et un mot féminin. En revanche, le démonstratif masculin *cet*, bien qu'il soit différencié du féminin *cette* à l'écrit, se prononce de la même manière que la forme féminine lorsqu'il précède un mot commençant par une voyelle (ex. [sɛt] ami). Enfin, l'article indéfini *un* constitue lui aussi un cas d'ambiguïté à l'oral, en particulier au Canada où la liaison avec le nom masculin commençant par une voyelle a tendance à être prononcée de la même façon que l'indéfini féminin *une* sur le modèle anglais de *an* (ex. [ɒn] ami). L'hypothèse de la perceptibilité prédit une tendance à la masculinisation liée à l'ambiguïté du possessif, mais une tendance à la féminisation consécutive à l'ambiguïté du démonstratif et de l'indéfini. Étant donné que le démonstratif et l'indéfini sont largement plus utilisés que le possessif dans la langue, cette hypothèse permet de rendre compte de l'observation de Barbaud et al. allant dans le sens d'une féminisation des mots en canadien français.

5.7 Conclusion

L'effet de variables morphophonologiques sur la réalisation de l'accord, que nous rapportons ici en français et en italien, confirme les récentes observations faites en néerlandais. Dans ces trois langues, un effet de certaines variables conceptuelles sur la construction de l'accord a également été rapporté (voir Introduction générale). Ces effets morphophonologique et conceptuel indiquent que le système responsable d'assurer l'accord est capable de prendre en considération des informations qui ne sont a priori pas strictement nécessaires. Pour reprendre la terminologie proposée par Vigliocco & Franck (1999), ces observations confortent l'hypothèse d'un input maximal selon laquelle les opérations syntaxiques puisent parmi l'ensemble des sources d'information potentiellement pertinentes pour effectuer leur travail correctement : lexicosyntaxique bien entendu, mais aussi conceptuelle et morphophonologique. Un tel fonctionnement entraîne sans doute un certain coût, puisque la composante syntaxique doit intégrer un plus grand nombre d'informations que si elle se limitait à ne traiter qu'une seule source. Néanmoins, les observations montrent que la prise en compte d'informations additionnelles assure une réalisation plus correcte des opérations syntaxiques, sans doute parce que la redondance permet de pallier à l'éventuelle perte d'information au cours de la production.

Les résultats obtenus jusqu'à présent ne permettent toutefois pas de conclure quant au mécanisme par lequel l'information morphophonologique est prise en compte par la composante syntaxique. Cela peut se faire par l'intervention d'un système externe de contrôle préarticulatoire assuré par le système de compréhension du langage, comme le suggère Levelt. Il se peut également que ce contrôle soit assuré par la composante syntaxique elle-même, grâce à un réseau de connexions bidirectionnelles qu'elle entretiendrait avec la composante morphophonologique, comme le suggère une perspective interactive forte. Une

dernière possibilité consiste à situer le traitement morphologique non pas au niveau de l'encodage phonologique mais bien aux niveaux lexical et syntaxique. Dans une telle perspective, la question soulevée par nos résultats n'est plus celle des relations entre la syntaxe et la morphophonologie, mais bien celle des relations entre la syntaxe et le lexique. Dans ce cadre, nos résultats confortent l'hypothèse d'un fonctionnement interactif entre les composantes lexicale et syntaxique et sont compatibles avec le modèle de Stemberger qui ne postule pas de relations bidirectionnelles entre la syntaxe et la phonologie.

En ce qui concerne l'estimation 'quantitative' des indices morphophonologiques manipulés, nous avançons que leur impact dépend à la fois de leur validité, de leur fiabilité et de leur perceptibilité, ces trois critères étant tous étroitement liés à la fréquence d'occurrence et de co-occurrence des unités lexicales dans le réseau interactif. Cette approche met en évidence l'optimisation de la régulation de l'input informationnel par le système syntaxique en même temps qu'elle rend compte des différences observées entre les locuteurs français et italiens.

Enfin, l'hypothèse que l'impact des différents indices sur la réalisation de l'accord varie en fonction de leur validité et de leur fiabilité relatives pose la question de l'interprétation de l'effet conceptuel rapporté dans le Chapitre 2. Dans cet article, nous montrons que des noms ayant un genre conceptuel sont plus facilement accordés que des noms comportant un genre grammatical. Les terminaisons de ces derniers constituant des indices de moindre validité et fiabilité que celles portées par les mots ayant un genre conceptuel ; les variables morphophonologique et conceptuelle apparaissent être confondues. Cette question est abordée dans la Synthèse théorique du Chapitre 8.

Chapitre 6

Attraction in sentence production: The role of syntactic structure and processing complexity

In this paper, we report four experiments conducted in French and in English in which we induced subject-verb agreement errors to explore the role of syntactic structure during sentence production. Previous studies have shown that attraction errors occur when a preverbal noun disagrees in number with the subject head noun. In this series of experiments, speakers were asked to complete complex sentential preambles. We found that agreement errors were more frequent following an intermediate modifier (e.g., *The threat-S to the presidents-P of the company-S ARE serious) than an immediately preverbal modifier (e.g., *The threat-S to the president-S of the companies-P ARE serious). It is suggested that attraction is determined by the syntactic distance between the interfering noun and the head noun. Furthermore, all experiments revealed that the number of plurals in the sentence preamble increases the probability of an agreement error. We conclude that agreement errors are determined by the structural factor of syntactic distance and by a factor of complexity related to plurality.

6.1 Introduction

For the past century, many researchers have believed that errors provide us with a particularly well suited window from which to observe many aspects of mental life. In the psycholinguistic literature, the primary data upon which models of language production are based are slips of the tongue (Fromkin, 1971; Garrett, 1976, 1980, 1982, 1990; Dell, 1986; Stemberger, 1985). In this paper, we deal with a particular type of speech error: failures in number agreement between the sentential subject and the verb. We focused on the organisation of syntactic segments during the realisation of agreement to understand how syntactic structure contributes to the occurrence of errors. To do so, complex sentence preambles were created in which multiple NPs varied in number-marking. This manipulation introduced variations in processing complexity. Hence, this study explores how syntactic structure and processing complexity affect the incidence of agreement errors.

Agreement is a very widespread phenomenon that can be found in most natural languages which makes it a key tool for comparing languages and elaborate general principles about how syntactic relationships are ensured in sentence production. Two languages were studied, namely English and French. The basic characteristics of syntactic encoding are expected to be the same in the two languages, but differences may be found as to the way they deal with syntactic number, as a consequence of potentially different conceptual representations (at a higher level) and morpho-phonological realisations (at a lower level). The parallel study of these two languages will allow us to draw general principles about mechanisms and units of grammatical encoding as well as to detect eventual language-specific aspects of sentence production.

In English as in French, subject-verb agreement in number is governed by a simple rule: a singular subject requires a singular verb while a plural subject requires a plural verb. Despite the simplicity of this rule, errors sometimes occur, as illustrated in (Ex. 43) and (Ex. 44).

Ex. 43 : *The influence of some contemporary writers and fashions are allowed to enter

Ex. 44 : *Je souhaite que la mention de ces quelques conférences vous seront utile(s)

I hope that the mention-S of these conferences-P will be-P useful for you

In (Ex. 43) the verb ('are') agrees with the plural nouns ('writers and fashions') which are closer to the verb, instead of the singular subject head noun ('influence'). Similarly, in (Ex. 44) there is agreement between the verb ('seront') and the plural noun that precedes it ('conférences'), rather than agreement between the verb and the singular subject head noun ('mention').

Many recent experimental studies carried out in different languages have attested to the role of an intervening local noun (i.e., a noun embedded in a PP or in a clause that modifies the subject NP) which disagrees in number with the subject head noun in inducing erroneous number agreement in oral production (e.g., Bock and Cutting, 1992; Bock and Eberhard, 1993; Bock and Miller, 1991; Vigliocco, Butterworth and Semenza, 1995; Vigliocco, Butterworth and Garrett, 1996a; Vigliocco, Hartsuiker, Jarema and Kolk, 1996b). In a typical experiment, participants were visually or acoustically presented with sentence preambles manipulating the number on the head and local nouns (e.g., The editor(s) of the book(s)) and their task was to repeat the preamble and provide a sentence completion. Written number agreement production was also studied using a dictation task; this allowed the examination of verbs that have the same plural and singular ending when presented auditorily in French (Fayol et Got, 1991; Fayol, Largy and Lemaire, 1994; Hupet, Fayol and Schelstraete, 1998). Errors in number agreement between the subject head noun and the verb were computed and analysed. All studies, conducted in different languages, reported a reliable increase in agreement errors when the local noun mismatched in number with the head noun.

Interestingly, the same influence of a local noun was found on the production of gender agreement in French and Italian. It was reported that gender agreement between a subject head noun and a predicative adjective was significantly disrupted when a local noun with a different gender from the head noun was present in the sentence (Vigliocco and Franck, 1999; Vigliocco and Franck, submitted).

Ex. 45 : *Le travail de la couturière est délicat
The handwork (M) of the dressmaker (F) is delicate (F)

In (Ex. 45), the predicative adjective ('délicat') should agree with the masculine subject head noun ('le travail') and not with the feminine local noun ('la couturière').

6.2 Interpreting the attraction effect

To describe agreement errors such as the ones reported above, grammarians such as Quirk et al. (1972) introduced the principle of proximity also termed attraction. The label "proximity" emphasizes the closeness of the verb to the mismatching local noun. Attraction denotes agreement with a closely preceding noun phrase in preference to agreement with the head of the noun phrase that functions as subject. Quirk et al. noticed that a conflict between grammatical concord and proximity concord tends to increase with the distance between the head noun and the verb. In 1924, Jespersen proposed the first processing account of erroneous agreement with a preverbal local noun, venturing the hypothesis that "if the verb comes long after the noun, there is no more mental energy left to remember what was the number of the subject" (Jespersen, 1924, p.345). Again, this suggests that the probability of finding proximity concord rather than grammatical concord increases with the distance between the subject head noun and the verb in the uttered sentence. This account supposes that in cases where the subject NP and verb are discontinuous, such as when the subject NP is followed by a modifying expression, mental energy is required to keep track of the information about the agreement source until the target becomes available.

Along these lines, Fayol and collaborators studied the relation between agreement and working memory in language production by adding a concurrent task to the sentence completion task described above. Fayol, Largy and Lemaire (1994) reported that agreement errors in written French were more common when working memory was overloaded by a concurrent task. In their experiments, erroneous agreement with a local noun was more likely when the participant was engaged in remembering a series of 3-4 unrelated words than when working memory was not involved in a concurrent task. According to Fayol et al., subject-verb agreement is computed automatically on the basis of spreading activation of the number feature from the closest preceding noun to the verb. In most cases, the subject noun immediately precedes the verb; this ensures correct and rapid

agreement (because it is automatic). However, when there is a preverbal local noun, activation will spread from this too. In order to guarantee correct agreement with the head noun, a non-automatic checking mechanism is assumed to be activated, which consumes working memory resources. A concurrent memory load task would reduce resources available for the checking procedure, and therefore increase the probability of agreement of the verb with a local noun.

This view that the agreement operation errs because the verb is directly preceded by a local noun that disagrees in number with the subject noun, will be referred to as the linear distance hypothesis. The crucial factor that influences the occurrence of errors in this account is the linear proximity between words in the sentence, specifically between the local noun and the verb. In this theoretical framework, the structure in which syntactic segments are framed when agreement takes place is assumed to be a linear chain that parallels the uttered word string. Relations between words are determined by the linear distance that separates them and working memory is assumed to keep track of the segments positioned first in the sentence.

Although this account seems to cover the majority of the available data, it is inconsistent with some experimental results of agreement error induction reported by Bock and Cutting (1992). In three experiments, they found that errors were significantly more common when the preamble contained a prepositional phrase (PP) modifier, as in the examples in (Ex. 46 and Ex. 47), than when the preamble contained a relative clause or a complement clause modifier, as in the examples in (Ex. 48 and Ex. 49).

Ex. 46 : The editor of the history books

Ex. 47 : The report of the destructive fires

Ex. 48 : The editor who rejected the books

Ex. 49 : The report that they controlled the fires

Note that crucially, the preambles in (Ex. 48 and Ex. 49) are syntactically more complex than those in (Ex. 46 and Ex. 47). The linear distance hypothesis would therefore predict even more errors for preambles such as (Ex. 48 and Ex. 49) than for preambles such as (Ex. 46 and Ex. 47) because the presence of a clausal modifier should involve greater working memory demands than a prepositional phrase modifier.

Bock and Cutting argued that the different error rates for prepositional phrases and clausal modifiers can be considered evidence in favour of a «clause packaging» organisation of production units during sentence planning. A mismatching local noun would be more likely to interfere with agreement computation if it is encoded simultaneously with the head noun, i.e. within the same encoding unit (the clause). The presence of a «clausal boundary» (as in Ex. 48 and Ex. 49) would reduce the probability of an agreement error by insulating the head noun from the number feature of the local noun. In contrast, in (Ex. 46

and Ex. 47) both nouns are within the same clause. We will refer to this view as the clause packaging hypothesis. The general framework of sentence production in which this hypothesis can be integrated assumes that the process of building a syntactic structure for the sentence, also called «grammatical encoding», occurs in two steps: a stage at which functional structures are built and a stage at which words are positioned in their left-to-right order in a string (Garrett, 1976). The processes at the functional level involve the integration of lexical units with their conceptual and syntactic properties (also called lemmas) into a syntactic frame that specifies their grammatical function, as well as syntactic relations between them. Agreement is supposed to take place at this stage of production. In contrast with the framework underlying the linear distance hypothesis, a sentence like (Ex. 50) is assumed to be encoded in two separate units, i.e. ‘the claim was rejected’ and ‘that wolves were stealing babies’.

Ex. 50 : The claim {that wolves were stealing babies} was rejected

Such a framework assumes that lexical units, before being positioned in the final left-to-right order, are inserted into higher-level syntactic constituents structured hierarchically, i.e. clauses, over which syntactic operations proceed in relative insulation from one clause to another (e.g., Bock, 1987; Dell, 1986; Garrett, 1982, 1988). As the different constituents are assembled, the information they contain becomes available for positional-level processing. This stage involves the retrieval of phonological representations for words and their ordering in a left-to-right order.

The hypothesis of a mode of operation constrained by higher-level constituents is widely shared by modern theories of sentence production (Bock, 1987; Dell, 1986; Fromkin, 1971; Garrett, 1988; Levelt, 1989; Stemberger, 1985). However, these theories divide over how selective the mechanisms are and therefore how strong the hierarchy of control is (Bock and Cutting, 1992). If one assumes that the functional level processes operate over a tree-like constituent structure in which smaller units are the phrases, as illustrated in Figure 1, this hierarchical structure may have finer grained control over a syntactic operation like agreement than what is assumed by the clause packaging hypothesis. Let us assume that agreement computation consists of a feature-matching operation (Kempen and Hoenkamp, 1987; Kempen and Vosse, 1989; Vigliocco et al., 1995) during the construction of such a hierarchical syntactic frame. The number feature of the subject head noun would percolate from a noun (N) node up the syntactic tree to the highest NP projection (NP') and then to the sentence (S) node, in order to be available for subject-verb agreement. If there is a NP modifier, it may be that the number feature of the noun contained in the modifier (the local noun) also percolates up the syntactic tree and is therefore sometimes erroneously used to compute agreement with the verb. However, this is an unlikely event because the syntactic path from the local NP to the highest NP projection is always longer than the path from the

subject head noun to this projection. In this theoretical framework, the syntactic distance between a mismatching local noun and the subject head noun is the crucial factor in agreement error production.

Consider the different structures associated with examples (Ex. 46 and Ex. 47) and (Ex. 48 and Ex. 49). For relative or complement clauses, the tree structure includes an extra node for the NP modifier, in comparison to the structure with a phrasal modifier, as illustrated in Figure 15.

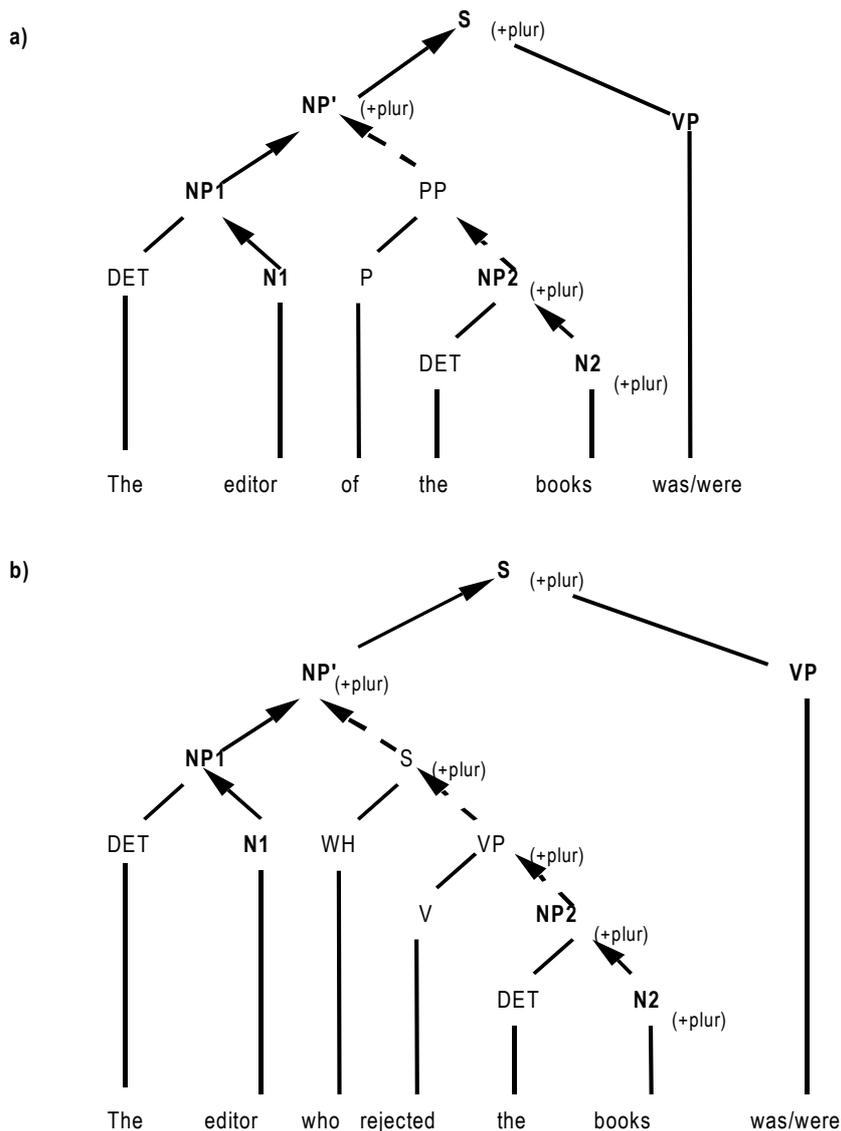


Figure 15. An account of the different error rates for prepositional phrases and clausal modifiers obtained by Bock and Cutting (1992) in terms of syntactic distance. The syntactic path that a mismatching local plural feature has to cover in order to interfere with agreement construction is shorter in (a), where from NP2 to NP' there is just a PP node, than in (b) where from NP2 to NP' there are two nodes (VP and S).

Therefore, the syntactic distance between the head noun and the local noun which carries a mismatching number feature is greater in (Ex. 48 and Ex. 49) than in (Ex. 46 and Ex. 47), and the longer the path, the less likely it is that the mismatching feature will influence agreement. On this account, the presence of a clause boundary does not insulate the matrix clause from the modifying relative or complement clause but it creates a longer syntactic path from the mismatching noun to the head noun. Hence, one would expect fewer errors with clausal than with prepositional phrase modifiers, but one would still expect a substantial number of errors. This is what Bock and Cutting found in their study (they report 37 errors with prepositional phrase modifiers and 25 errors with relative clause modifiers in the mismatching conditions in Experiment 1).

Bock and Cutting also reported that more agreement errors occurred when the subject pronoun embedded in the complement clause modifier mismatched in number with the head (e.g. The report that they controlled the fire(s)), than when it matched (e.g. The report that he controlled the fire(s)). This result is consistent with an account in which the syntactic distance between a plural nonhead NP and the head NP affects erroneous agreement.

To summarise, according to the syntactic distance hypothesis, agreement errors occur at a stage during grammatical encoding in which a syntactic tree structure is realised and features such as number are passed to different portions of this tree structure. The probability that a mismatching local feature interferes with agreement computation is dependent upon its syntactic distance from the subject head noun. On this account, the clause boundary has an effect because it makes the path substantially longer, not because it insulates the local noun from the treatment of the subject head noun. Experimental evidence in favour of the syntactic distance hypothesis and falsifying a linear distance view was reported in an experiment carried out by Vigliocco and Nicol (1998). They presented participants preambles containing a phrasal modifier (The helicopter for the flight) and asked them to make up questions with subject-auxiliary inversion (Is the helicopter for the flight safe?). The agreement error pattern was the same as the error pattern found when participants were asked to complete the preamble by making a declarative sentence. In other words, errors were equally common when the local noun was adjacent to the verb in the linear word chain and when it was separated from the verb by the subject head noun. This result was expected if one considers that agreement takes place when the sentence is structured hierarchically before it is linearised. Phenomena like word order inversion in the case of questions are assumed to occur after syntactic relations between the different segments have been established. The data indicate that agreement errors arise as a consequence of syntactic proximity of the local NP and head NP within the tree structure.

Nevertheless, although this result does not support the linear hypothesis, it is still compatible with a more restrictive view of how hierarchy can control syntactic

operations, like the one suggested by the clause packaging hypothesis. The clausal organisation is the same for declarative and interrogative structures and a similar attraction effect is expected because the local noun is part of the subject clause. Furthermore, the use of interrogative sentences does not allow a direct comparison with data available in the literature which were all obtained with declarative sentences. In this paper, we explore the effect of the local noun's position within the subject clause by introducing two phrasal subject modifiers in declarative sentences, as illustrated in (Ex. 51) and (Ex. 52).

Ex. 51 : The helicopter for the flight over the canyons

Ex. 52 : The helicopter for the flights over the canyon

Manipulating the number mismatches between the head noun and the local noun (as in Ex. 51) and between the head noun and the intermediate noun (as in Ex. 52) allows us to put to test the three hypotheses presented above because they make different predictions about which preamble type will elicit more attraction errors. The linear distance hypothesis, as defended by Fayol and collaborators, predicts a clear attraction effect with the local noun that directly precedes the verb, as it is assumed that an automatic procedure makes the verb agree with the closest preceding noun in the word string. This hypothesis therefore predicts a high error rate in (Ex. 51) whereas no (or fewer) attraction errors should occur in (Ex. 52). According to the clause packaging hypothesis, similar interference should be found with the local and the intermediate noun, as both nouns are situated in the same encoding unit as the subject noun: There should be no difference between (Ex. 51) and (Ex. 52). The syntactic distance hypothesis predicts that the intermediate noun, situated high in the hierarchical syntactic structure, will interfere to a greater extent than the final noun with the agreement process, because its number feature can occasionally percolate up the tree. The local noun, which is deeper in the tree, should influence the agreement process either not at all or to a lesser extent. In other words, more attraction errors should be found in (Ex. 52) than in (Ex. 51).

The use of the particular preamble structure containing two potentially interfering nouns allowed the emergence of a particularly counter-intuitive pattern of results. Indeed, data reported throughout all experiments presented in this paper clearly indicates that it is not the intuitively disturbing number feature on the local noun, next to the verb, that disrupts the agreement process, but rather the feature on the intermediate noun, that is further from the verb in the uttered sentence.

6.3 A preview of the experiments

Four experiments eliciting errors of subject-verb agreement in number are reported, which allowed us to test the differing predictions of the linear distance, the clause packaging and the syntactic distance hypotheses. Agreement is used as a

tool to study the way syntactic units are processed at the stage of grammatical encoding. Each experiment was conducted in parallel in French and English.

Some differences have been previously reported between the two languages, concerning how conceptual representations can influence the process of agreement (e.g., Bock and Miller, 1991; Vigliocco et al., 1996; Vigliocco and Franck, 1999). Nevertheless, both languages present a clear attraction effect and we hypothesise that the same cause underlies this phenomenon. The syntactic distance hypothesis we put forward assumes as a general principle the characterisation of grammatical encoding in both languages as follows: syntactic segments are organised hierarchically, before words are linearly ordered in a to-be-uttered string. It suggests that in both languages, subject-verb agreement is ensured by the percolation of number features from the subject noun and verb up to the hierarchical syntactic structure, where a unification process checks for matching. Comparative data will provide us with an empirical basis to test whether these two assumptions of the syntactic distance hypothesis indeed guide the operation of agreement in two different languages.

Experiments 1 (French) and 2 (English) were designed to test the hypotheses of linear distance, clause packaging and syntactic distance. Speakers were presented with complex sentence preambles containing a head noun and two modifiers: an intermediate noun, situated relatively high in the tree structure (close to the subject head noun in the word string) and a local noun situated low in the syntactic tree (close to the verb in the word string). To summarise the predictions (see above): the linear distance hypothesis predicts a high attraction effect with the local noun but no attraction with the intermediate noun, whereas the hypothesis of syntactic distance predicts the opposite pattern. The clause packaging predicts no difference in attraction between the two modifiers.

In Experiments 3 (French) and 4 (English), we presented participants with the same materials that we used in the previous experiments, however we manipulated a concomitant memory load in an attempt to better evaluate whether the plurality, which is assumed to be more complex, is sensitive to the reduction in working memory resources by a concurrent task. The prediction is that preamble conditions containing more plurals, being more complex, should be more sensitive to the presence of an additional memory load than less complex conditions containing fewer plurals.

6.4 Experiment 1 (French)

In Experiment 1 (French), as well as in the parallel English Experiment 2, we presented speakers with preambles (complex NPs) of the type illustrated in (Ex. 53).

Ex. 53 : L'ordinateur avec le programme de l'expérience
 The computer with the program of the experiment

These preambles contained three nouns: a head noun (computer), an intermediate noun (program) and a local noun (experiment). The number of each noun was manipulated. The intermediate and local NPs were both within modifying prepositional phrases. The intermediate NP modified the head NP, while the local noun NP modified the intermediate NP. Therefore, the two modifying nouns occupied two different positions in the hierarchical syntactic structure, with the local NP being the most embedded in the tree.

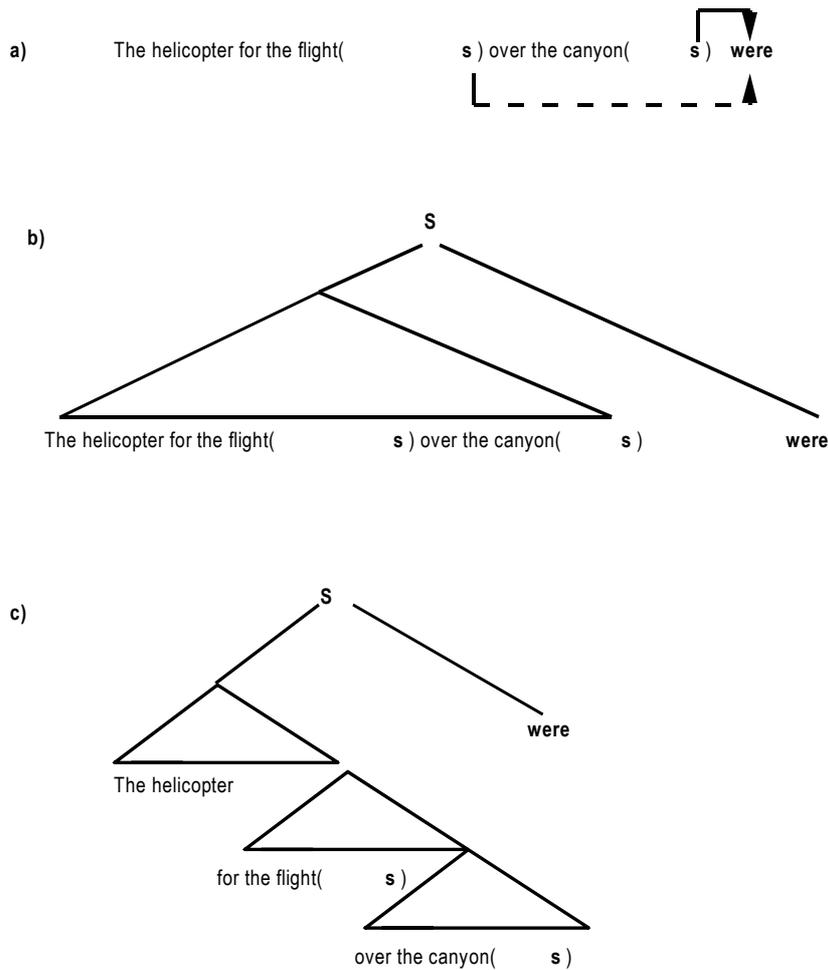


Figure 16. Predictions for double modifiers: In (a), the linear distance hypothesis predicts that the local noun (canyons) has a greater probability to influence agreement than the intermediate noun (flights) since it occupies a linear position closer to the verb. In (b), a clause packaging view predicts that a feature has the same probability to influence agreement computation if it is on the local or on the intermediate noun since they belong to the same clause and are encoded together. In (c), the syntactic distance hypothesis predicts more errors of agreement with the intermediate noun than the local noun since the syntactic distance its feature has to travel up the tree is shorter.

The use of these materials allowed us to put to the test the three hypotheses presented above. The syntactic distance hypothesis predicts that more verb agreement errors will occur when the intermediate noun disagrees in number with

the head (The computer with the programs of the experiment) than when the local noun disagrees with the head (The computer with the program of the experiments). The reason for this is that the intermediate noun is higher in the tree structure than the local noun and therefore has more chance to have its number feature percolate up to the maximal projection of the head NP and unify with the verb's feature. Note that this prediction is totally unexpected given the widespread view of agreement in terms of proximity concord. The hypothesis of clause packaging predicts no difference between the intermediate noun and the local noun in the way they will attract verb agreement. In this framework, both nouns are part of the same clause as the head noun and there is no reason to assume that they will have a different impact on the agreement process. These three hypotheses are illustrated in Figure 16.

6.4.1. Method

Participants. Fifty-six students, aged between 18 and 45, of the Université catholique de Louvain took part in the Experiment. They were paid 100 BF for their participation.

Materials. The experimental materials were sentence preambles composed of a head noun (N1), an intermediate noun (N2) and a local noun (N3), as illustrated in Table 1 (Tableau 19). Both N2 and N3 were embedded in prepositional phrases. The variables which were experimentally manipulated were: (1) the number of N1 (Singular vs. Plural), (2) the number of N2 (Singular vs. Plural), and (3) the number of N3 (Singular vs. Plural). For each item, eight different versions were created to represent the different combinations of singular and plural. Experimental items are listed in Appendix 1 (Annexe 8).

Tableau 19 (Table 1) : Different Versions of a Sentential Preamble in Experiments 1 and 3 in French. S = singular noun, P = plural noun.

Condition	Sentence preamble
SSS	L'inscription sur la porte de la toilette <i>The inscription on the door of the toilet</i>
SSP	L'inscription sur la porte des toilettes <i>The inscription on the door of the toilets</i>
SPS	L'inscription sur les portes de la toilette <i>The inscription on the doors of the toilet</i>
SPP	L'inscription sur les portes des toilettes <i>The inscription on the doors of the toilets</i>
PPP	Les inscriptions sur les portes des toilettes <i>The inscriptions on the doors of the toilets</i>
PPS	Les inscriptions sur les portes de la toilette <i>The inscriptions on the doors of the toilet</i>
PSP	Les inscriptions sur la porte des toilettes <i>The inscriptions on the door of the toilets</i>
PSS	Les inscriptions sur la porte de la toilette <i>The inscriptions on the door of the toilet</i>

Eight 96-item lists were created. In each list there were 32 experimental items and 64 fillers. All the experimental conditions were represented within each list. Each condition was represented by four preambles. In each list there was only one of the eight possible versions of the same item. The filler items were NPs-PPs; half had a singular head noun and half had a plural head noun. For all filler items, the modifying noun matched in number with the head noun. Preambles were arranged in a pseudo-random order: each list started with 5 fillers and there were no more than three consecutive experimental preambles.

Procedure. Each participant was tested individually in a session lasting approximately 30 minutes. Participants were instructed to read the sentence beginnings, then to repeat and complete them. They were asked to speak as quickly as possible and to use the verb 'to be' (être) (this maximised the number of inflected verb forms produced collected). The experiment was presented visually and programmed on Superlab for Apple Macintosh. Presentation times were 3000 ms for both experimental and filler items. The experimenter regulated the appearance of the items on the screen.

Scoring. The completions were transcribed and then assigned to one of the following scoring categories: (1) Correct Responses were scored when the participant correctly repeated the preamble and completed the sentence with a correctly inflected verb. (2) Agreement Errors were scored when the completion met the criteria above but the verb form failed to agree in number with the subject of the sentence. (3) Number repetition errors were scored when the participants changed the number marking of N1, N2 or N3. Note that agreement errors after a repetition error were considered separately from other agreement errors as the conditions in which they appeared have been modified by the number repetition error. (4) Miscellaneous Responses were scored when the participant failed to apprehend the preamble (or parts of it), when he/she failed to repeat some words in the preamble, and when he/she produced a completion lacking the main verb. If two different utterances were produced in succession, only the first was scored, including those cases in which an agreement error was produced and immediately corrected.

Design and data analyses. For the production experiment, the main statistical tests were carried out using the number of agreement errors, the number of errors in the repetition of number, and the number of miscellaneous responses as the dependent measures. For each dependent variable, we conducted: (a) a repeated measures ANOVA, where the three orthogonal factors were: number of N1, number of N2 and number of N3, both with subjects (F1) and items (F2) as random factors; (b) pairwise comparisons between conditions.

6.4.2. Results

Application of the scoring criteria yielded 1150 (64.2%) correct responses, 79 (4.4%) subject-verb agreement errors, 417 (23.3%) repetition errors and 146 (8.2%) miscellaneous responses. The proportion of errors in the different conditions is presented in Table 2 (Tableau 20).

Agreement errors. There was no main effect of the number of any of the nouns ($F_s < 1$). There were more agreement errors when the number of N1 was in mismatch with the number of N2 ($F(1,55) = 31.05, p < .001$; $F(1,31) = 40.48, p < .001$) and when the number of N2 was in mismatch with the number of N3 ($F(1,55) = 9.19, p = .004$; $F(1,31) = 11.48, p = .002$). There were more agreement errors when N1 and N3 matched in number than when they mismatched ($F(1,55) = 5.49, p = .023$; $F(1,31) = 8.80, p = .006$).

Relevant contrasts with regard to the linear distance hypothesis showed no difference between SSP and its baseline SSS and between PPS and its baseline PPP. Relevant contrasts with regard to the clause packaging hypothesis showed significantly more errors in SPS than SSP ($t(55) = 4.96, p < .001$) and in PSP than PPS ($t(55) = 2.8, p = .007$). Finally, relevant contrasts with regard to the syntactic distance hypothesis revealed significantly more errors in SPS than its baseline SSS ($t(55) = 4.15, p < .001$) and in PSP than its baseline PPP ($t(55) = 4.43, p < .001$). Conditions SPP and PSS also induced more errors than the control conditions SSS and PPP (respectively $t(55) = 1.99, p = .05$ and $t(55) = 2.01, p = .04$). PSP did not differ significantly from PSS but there were more agreement errors in SPS than SPP ($t(55) = 2.57, p = .013$).

Tableau 20 (Table 2) : Distribution of Responses in Experiment 1 (French) in percent.

Condition	Correct responses	Agreement errors	Repetition errors	Miscell. responses
SSS	10.27	0.17	1.17	0.89
SSP	8.87	0	2.34	1.28
SPS	6.47	1.23	4.07	0.73
SPP	6.75	0.56	4.07	1.12
PPP	8.98	0.17	2.29	1.06
PPS	7.76	0.39	3.35	1.00
PSP	6.58	1.17	3.68	1.06
PSS	8.48	0.73	2.29	1.00
Total	64.17	4.41	23.27	8.15

Repetition errors. ANOVAs revealed a significant effect of the number of N2, with more errors produced when N2 was plural than when it was singular ($F(1,55) = 13.32, p = .001$; $F(1,31) = 10.97, p = .002$). The interaction between the number of N1 and the number of N2 was significant, indicating that more repetition errors

occurred when both nouns mismatched in number compared to when they matched ($F(1,55) = 25.80, p < .001$; $F(1,31) = 11.39, p = .002$). A mismatch between the numbers of N2 and N3 was also a significant source of error, as indicated by the significant interaction between these two numbers ($F(1,55) = 10.49, p = .002$; $F(1,31) = 6.55, p = .016$).

Miscellaneous responses. There was no significant effect ($F_s < 1$).

6.4.3. Discussion

Experiment 1 shows that in contrast with the traditional view of attraction errors, when two potentially interfering nouns were present in the preamble, subject-verb agreement was not disturbed by the local noun (N3), preceding the verb, but rather by the intermediate noun (N2), which was further from the verb. No attraction at all was reported when the number feature on the local noun mismatched the feature on the head noun. In contrast, an important attraction effect was found with the feature of the intermediate noun when it mismatched in number with the head. This finding is in line with the prediction of the syntactic distance hypothesis, which assumes that the hierarchical structure in which syntactic segments are organised during grammatical encoding influences the incidence of agreement errors. In this framework, agreement takes place when the subject, intermediate and local nouns are placed in a tree structure in which the subject is situated in the highest position whereas the local noun appears lower in the tree structure. The higher the number feature in the hierarchy, the greater the probability that the verb will be affected. This is indeed what was found; although in most of the cases agreement is correctly computed between the subject and the verb, attraction errors only occurred when the intermediate noun mismatched in number, while local noun number did not influence error rate at all.

This result falsifies the account of attraction errors based on linear distance. The linear distance hypothesis predicted that speakers would produce more errors with the noun closer to the verb in the linear word string, in virtue of an automatic mechanism of copying of its feature onto the verb. Our results show that no attraction effect was found with the local noun that preceded the verb. Data also contradict the clause packaging hypothesis, which predicted that any number feature present in the same clause as the subject would yield interference with the agreement process, whatever its position in the clause. We found that although the intermediate noun and the local noun were situated in the same encoding unit (i.e. the same clause), they did not have the same impact on agreement attraction; only the intermediate noun yielded errors.

Before turning to general theoretical considerations, it would be useful to ensure that this is a widespread phenomenon, occurring in languages other than French. Therefore the same experiment was conducted in English.

6.5 Experiment 2 (English)

6.5.1. Method

Participants. Forty undergraduate students from the University of Arizona participated in this experiment. They received course credit or \$3 for their participation.

Materials. Materials were constructed as in Experiment 1. Conditions are illustrated in Table 3 (Tableau 21). Experimental items are listed in Appendix 2 (Annexe 9).

Tableau 21 (Table 3) : Different Versions of a Sentential Preamble in Experiments 2 and 4 in English.

Condition	Sentence preamble
SSS	The threat to the president of the company
SSP	The threat to the president of the companies
SPS	The threat to the presidents of the company
SPP	The threat to the presidents of the companies
PPP	The threats to the presidents of the companies
PPS	The threats to the presidents of the company
PSP	The threats to the president of the companies
PSS	The threats to the president of the company

Procedure. Lists were recorded on a digital recording system by a female speaker. Participants were instructed to listen to the sentence beginnings, then to repeat and complete them. They were asked to speak as quickly as possible, but no other constraint was put on the form or content of the completions. Eight practice preambles (of the filler type) were presented to the participants at the beginning of the session. If the participant failed to understand a preamble, the experimenter repeated it.

Scoring. Same as in Experiment 1, except for the addition of the category Uninflected Verb Responses. These responses were so categorised when the completion met the criteria for a correct response but the verb form produced was uninflected for number (i.e., a past tense of a regular verb).

Design and data analysis. Same as in Experiment 1.

6.5.2. Results

Application of the scoring criteria yielded 784 (60.1%) correctly inflected verbs, 80 (6.2%) agreement errors, 129 (10.1%) number repetition errors, 118 (9.2%) miscellaneous responses and 184 (14.4%) uninflected verb forms. Table 4 (Tableau 22) reports the distribution of responses in the experimental conditions.

Tableau 22 (Table 4) : Distribution of Responses in Experiment 2 (English) in percent

Condition	Correct responses	Agreement errors	Repetition errors	Miscell. responses	Uninflected verb forms
SSS	9.38	0.16	0.31	1.17	1.48
SSP	8.20	0.39	0.86	1.09	1.95
SPS	5.78	1.95	2.03	0.86	1.88
SPP	6.48	1.64	1.80	0.86	1.72
PPP	7.50	0.39	0.86	1.88	1.88
PPS	7.81	0.31	1.56	1.17	1.64
PSP	6.17	1.25	2.42	0.94	1.72
PSS	8.75	0.16	0.23	1.25	2.11
total	60.08	6.25	10.08	9.22	14.38

Agreement errors. Errors were more common when N1 was singular than when it was plural ($F(1,39) = 7.4, p = .009$; $F(1,31) = 8.1, p = .008$) and when N2 was plural than when it was singular ($F(1,39) = 11.04, p = .002$; $F(1,31) = 7.94, p = .008$). There were more errors when N1 and N2 mismatched in number, as indicated by the significant interaction between N1 and N2 ($F(1,39) = 34.89, p < .001$; $F(1,31) = 33.8, p < .001$). Finally, the interaction between N2 and N3 was also significant ($F(1,39) = 5.42, p = .02$; $F(1,31) = 4.3, p = .04$), indicating that more errors were produced when they were in number mismatch. The interaction between N1 and N3 was not significant ($F_s < 1$).

The relevant contrasts with regard to the linear distance hypothesis showed that errors in SSP were not significantly more common than in SSS and that there were no more errors in PPS than in the baseline PPP. With regard to the clause packaging hypothesis, SPS yielded significantly more errors than SSP ($z = 2.72, p = .006$) and PSP yielded more errors than PPS ($z = 2.7, p = .007$). Relevant comparisons for the syntactic distance hypothesis were the following: there were more errors in SPS and SPP than in the SSS baseline (respectively $z = 3.41, p < .001$ and $z = 2.95, p = .003$) and PSP yielded more errors than the baseline PPP ($z = 2.4, p = .02$), however PSS did not. There was no difference between SPS and SPP.

Repetition errors. The analysis of variance showed a significant main effect of the number of N2, with more errors for plurals than for singulars ($F(1,39) = 8.13; p = .007$; $F(1,31) = 7.15; p = .01$), a significant interaction between the numbers of N1 and N2 ($F(1,39) = 6.7, p = .01$; $F(1,31) = 12.1, p = .001$) and between the numbers of N2 and N3 ($F(1,39) = 13.3, p < .001$; $F(1,31) = 16.9, p < .001$), indicating that more errors occurred when these numbers differed.

Pairwise comparisons on singular NIs showed significantly more errors for SPS and SPP than the baseline SSS ($z = 2.9$, $p = .005$; $z = 3.2$, $p = .001$, respectively). Number repetition errors were equally likely in SPS and SPP ($z = .45$, $p = .64$), but they were more common for SPS than SSP ($z = 2.5$, $p = .01$). For plural NIs, errors were more common in PSP than in the base-line PPP condition ($z = 2.78$, $p = .005$). Finally, errors were equally common for SPS and PSP.

Miscellaneous responses. None of the variables had an effect on miscellaneous errors ($F_s < 1$).

6.5.3. Discussion

Experiment 2 in English confirms the results obtained in the first Experiment in French: the attraction effect depends entirely on the presence of a number mismatch between the head noun and the intermediate noun. The number feature on the local noun appears to have no control over the subject-verb agreement process. As we already discussed in regard to the first experiment, the hypotheses of linear distance and clause packaging cannot account for this result. Rather, it appears that the determining factor in the occurrence of attraction is the syntactic distance between the syntactic segments of the sentence, as is specified in the hierarchical functional structure, before words are positioned in their left-to-right order.

Nevertheless, an alternative explanation to the syntactic distance hypothesis could be that the items which induced high error rates were actually more difficult to interpret. Conditions may have different degrees of comprehension difficulty or difficulty in the assignment of concepts to the appropriate sentence roles (see for e.g., Bock and Miller, 1991; Hupet et al., 1996). Parallel to the on-line production experiment, two off-line rating tests were conducted on the experimental items to estimate their plausibility and their imageability. The aim of these tests was to make sure that the differences found between conditions in agreement errors do indeed reflect differences in agreement computation and not differences in the apprehension of the preamble. The results of the plausibility and imageability judgement tasks exclude this possibility. Scores for both the French and the English materials showed little variation between the eight conditions as shown in Table 5 (Tableau 23).

We found no correlation in the items between plausibility and the occurrence of agreement errors (Pearson's correlation in English: $r = .123$; in French: $r = .292$) nor between imageability and the occurrence of agreement errors (English: $r = .06$; French: $r = .139$). This is not to say that there was no difference at all between conditions. In both languages, it was found that items in the SSS condition were judged more imageable than items in any other condition. But conditions inducing a high number of agreement errors did not receive lower plausibility and imageability scores.

Tableau 23 (Table 5) : Mean scores and standard deviations (between brackets) for the different number conditions in the plausibility and imageability judgement tasks in French and English

Condition :	Plausibility		Imageability	
	French	English	French	English
SSS	5.9 (1.7)	6.0 (.75)	4.7 (2.1)	5.3 (1.0)
SSP	5.6 (1.9)	5.9 (.76)	4.3 (2.1)	4.9 (.90)
SPS	5.5 (2.0)	5.8 (.71)	4.3 (2.1)	4.9 (.96)
SPP	5.5 (1.9)	5.7 (.78)	4.4 (2.1)	4.9 (.77)
PPP	5.3 (2.2)	6.0 (.74)	4.3 (2.2)	5.0 (1.0)
PPS	5.4 (2.1)	5.8 (.91)	4.4 (2.2)	4.9 (.83)
PSP	5.4 (2.1)	5.8 (.69)	4.5 (2.1)	4.7 (.85)
PSS	5.5 (2.0)	6.1 (.52)	4.1 (2.2)	5.0 (.89)

It is also worth noting here that miscellaneous responses, which are also a measure of the difficulty participants had in apprehending the preambles, did not differ across conditions in either language.

One can therefore reasonably conclude that the agreement error pattern reported in Experiments 1 and 2 supports the role of syntactic distance in the grammatical encoding of the sentence. Furthermore, the same effect was reported in French and in English, indicating that the factor of syntactic distance is not limited to the structural particularities of one specific language.

Although the main predictions with regard to the syntactic distance hypothesis were supported, other aspects of the data require discussion. One unexpected finding is that the mismatch effect found between the head noun and the intermediate noun is increased when there is an additional mismatch between the intermediate noun and the local noun, i.e. SPS and PSP conditions. In the condition with plural heads, such a double mismatch (PSP) yielded significantly more errors than the single mismatch (PSS) in both languages. A similar pattern was found for singular heads in French where significantly more errors were produced in SPS than SPP, however, there was no difference between these two conditions in English. The syntactic distance hypothesis predicted no difference between PSP and PSS, nor between SPS and SPP; if anything, a difference should have been found in the other direction, since incorrect feature percolation from the local noun, although deep in the tree, might still have influenced agreement realisation.

The second interesting aspect of the data concerns the asymmetry between singular and plural heads. An interesting cross-linguistic pattern of nearly all studies on agreement is that, in conditions of number mismatch, errors with plural head nouns (e.g., *The babies on the blanket*) were far less common than errors with

singular head nouns (e.g., *The baby on the blankets*). Hence, in most studies the attraction effect only occurred in the conditions with singular head nouns and plural local nouns, (Bock and Cutting, 1992; Bock and Eberhard, 1993; Bock and Miller, 1991; Fayol and Got, 1991; Fayol et al., 1994; Hupet, Schelstraete, Demaeght et Fayol, 1996; Vigliocco, Butterworth and Semenza, 1995, Vigliocco et al., 1996). Bock and Eberhard (1993; see also Eberhard, 1997; Tiersma, 1982) suggested that this asymmetry reflects a difference in the markedness of the nouns: plural nouns possess a marked grammatical feature for number that singular nouns lack. The probability that a marked local noun interferes with an unmarked head noun is greater than the opposite: a feature with a specified value (i.e., the plural form), being more salient, is more likely to replace a feature with no value (i.e., the singular form) than vice versa. However, in Experiment 1 (French), we found no hint of such an asymmetry, as attested by the absence of an effect of the subject's number. In the English experiment, there was a main effect of head number, indicating that, overall, more errors were produced when N1 was singular than when it was plural. Nevertheless, in contrast to previous studies, a large number of errors were produced in the condition with a plural head and a double number mismatch, i.e. PSP, which contrasts with the usual absence of an attraction effect with plural heads.

Our interpretation of these two aspects of the data (i.e. the effect of a double mismatch and the large number of errors with plural heads) lies in the concept of processing complexity that we assume is involved in the treatment of plural features. In this framework, the traditional asymmetry between singular and plural heads would emerge from greater complexity in preambles with plural nouns than with singular nouns. Plural forms would be more complex to deal with because they are both semantically and morphologically more complex than singular forms. Evidence for greater complexity associated with plurality comes from different data on language acquisition, production and comprehension (for a detailed review, see Eberhard, 1997). Developmental studies have shown that in children's speech, nouns appear in their singular form before the plural form (Cazden, 1968, Mervis & Johnson, 1991). It has also been found that children understand singular items before their plural counterparts (Clark, 1973; Lapointe, 1986). In his study of spontaneous errors, Stemberger (1985) reports a strong asymmetry, with more errors in the production of a plural than a singular. It also appears that when people are asked to memorise nouns, plural forms tend to be remembered as singulars while no errors of number arise for singular forms (Molen and Morton, 1973). Furthermore, a closer look at the various data on subject-verb agreement reveals that in many experiments, when the subject head noun and the local noun have the same number, errors are more common for plural heads than for singular heads (e.g., in French, Fayol et al., 1994; Franck, 1998; in English, Bock and Cutting, 1992; in Italian, Vigliocco et al., 1995; in Spanish, Vigliocco et al., 1996a). Fayol et al. (1994) also found that verb agreement with plural head nouns was more

disrupted by a secondary task than agreement with singular head nouns (such as when a memory load task was added to the sentence dictation task). It has also been found that elderly people, who often show a diminution of working memory resources, seem to be more sensitive to the presence of plurals than young adults (Fayol, Hupet, & Largy, 1999).

The traditional interpretation in terms of markedness cannot account for the absence of asymmetry reported in French, nor for the high error rate in PSP in English. On the view we defend here, it is not the presence of marking per se that explains the pattern of errors, but rather, the fact that marking increases processing load. In contrast to previous studies, our experiment used longer preambles with an additional NP modifier. This manipulation created the conditions for an increased error rate with plural head nouns in the PSP condition, due to the presence of the additional plural feature on N3, a novel result never reported before. A second potential source of complexity would be related to an increased memory load in the conditions in which number alternates (SPS and PSP). Why would this occur? Presumably because it is harder to build and use a representation for the sentence in which the mappings are changing or possibly switching from one morpho-phonological form to another may be more demanding.

Therefore, two factors may be responsible for inducing errors in our experiment: a structural factor, i.e. syntactic distance, that determines the probability of interference with the agreement process as a function of the plural feature's position in the syntactic tree, and a factor related to processing complexity. If the results of our first experiment can be plausibly described in processing-complexity terms, then increasing the memory load by introducing an additional task should be particularly disruptive for those conditions that, we assume, already have a high processing load. In the second pair of experiments, we test this possibility by introducing greater load in working memory.

6.6 Experiment 3 (French)

Experiments 3 (French) and 4 (English), aimed at replicating the effect found in Experiments 1 and 2 while assessing the presumed role of processing load. A dual-task paradigm was used. The primary task consisted in the sentence completion task, using the same materials as in the previous experiments. The secondary task consisted in memorising and reporting a series of one (low memory load condition) or three (high memory load condition) digits while performing the sentence completion task.

The general idea underlying this manipulation is that performance should be particularly disrupted by the additional memory task in just those conditions that already present a high processing load. Furthermore, if the idea of two different sources of complexity is correct, we should find that: (a) The effect of a memory

load should be greater for those conditions in which there are two or three plural features (PPP, PPS, PSP, SPP) than in the conditions in which there is one or no plural feature (SSS, SSP, SPS, PSS), and (b) There should be an effect of the memory load in the two conditions where number alternated, i.e. SPS and PSP. This effect should even be stronger in PSP since complexity is the highest (due to the combination of two plural nouns and of number alternation).

Experiments 3 and 4 are discussed together following the Results section of Experiment 4.

6.6.1. Method

Participants. 48 students of the Université catholique de Louvain took part in the Experiment. They were all between 19 and 25 and received credit for their participation.

Materials. The primary materials for this experiment consisted of the same sentence preambles (experimental and filler items) used in Experiment 1. Each preamble was associated with a one-digit load in one condition and a three-digit load in another condition. Memory load was manipulated between subjects. In the three-digit load, one (different) number was associated with each sentence preamble to be presented to the participants. Sixteen lists were created (corresponding to the combination of the different experimental conditions). In lists 1-8, the experimental preambles were associated with one-digit load. In lists 9-16 the same preambles were associated with three-digit load.

Procedure. Same as in Experiment 1, except that participants were visually presented with one (or three) digit(s) before each preamble. Digits were presented visually for 1000 ms just after a 500 ms. fixation point. A blank screen of 100 ms separated the digit from the preamble. Triplets of digits were presented simultaneously and vertically on the screen to prevent participants from chunking the digits. Participants were asked to repeat the digits back as soon as they had performed the completion task.

Scoring. Same as Experiment 1. A category for Memory errors was added that scored when one (or more) of the digits was not reported or incorrectly reported or when they were not reported in the right order.

Design and data analyses. The effect of the four variables (number of N1, number of N2, number of N3 and memory load) on the different error types (agreement, repetition, memory, miscellaneous) was assessed by 2X2X2X2 ANOVAs with participants (F1) and items (F2) as random factors. Balanced ANOVAs were performed on the subjects analysis, because memory load was part of a between-subject design, whereas other variables were between-subjects. Items were part of a repeated measures design because all variables were within-items. Subsequent pairwise comparisons between conditions were conducted.

6.6.2. Results

Application of the scoring criteria yielded 808 (52.6%) correct responses, 95 (6.2%) subject-verb agreement errors, 401 (26.1%) repetition errors, 398 memory errors (25.9%) and 53 (3.5%) miscellaneous responses. The data are presented in Table 6 (Tableau 24).

Agreement errors. Memory load had no effect on the overall production of agreement errors.

Tableau 24 (Table 6) : Distribution of errors in Experiment 3 (French). The first number of each cell corresponds to the responses in the one-digit condition while the number between brackets corresponds to the responses in the three-digit condition.

Condition	Correct responses	Agreement errors	Repetition errors	Memory errors	Miscell. responses
SSS	10.29 (8.2)	0	1.17 (1.17)	0.78 (3.26)	0.52 (0.65)
SSP	8.70 (6.9)	0.26 (0.13)	3.39 (2.86)	0.52 (3.78)	0.13 (0.13)
SPS	5.47 (3.26)	1.69 (1.69)	4.43 (3.8)	2.60 (6.38)	0.13 (0.39)
SPP	7.29 (5.21)	0.65 (1.04)	4.82 (2.86)	1.17 (5.6)	0.26 (0.65)
PPP	9.24 (5.73)	0.26 (0.39)	1.69 (2.86)	1.43 (4.56)	0.26 (0.65)
PPS	5.99 (5.08)	0.26 (0.91)	4.56 (3.78)	2.21 (5.47)	0.65 (0.26)
PSP	5.21 (3.5)	2.21 (2.21)	4.69 (3.78)	1.56 (5.99)	0.26 (0.78)
PSS	8.98 (6.12)	0.52 (0.13)	2.99 (2.21)	0.91 (5.6)	0.65 (0.52)
Total	61.20 (44.0)	5.86 (6.51)	27.73 (24.5)	11.20 (40.6)	2.86 (4.04)

More errors occurred when there was a number mismatch between N1 and N2 and between N2 and N3 (respectively $F(1,46) = 48.88$, $p < .001$; $F(1,31) = 28.15$, $p < .001$; $F(1,46) = 20.49$, $p < .001$; $F(1,31) = 10.79$, $p = .003$). Again, there were significantly more errors when N1 and N3 had the same number than when they were in number mismatch ($F(1,46) = 40.52$, $p < .001$; $F(1,31) = 24.03$, $p < .001$). The three-way interaction was marginally significant ($F(1,46) = 4.17$, $p = .047$; $F(1,31) = 4.22$, $p = .049$). SPS yielded more errors than SPP ($t(47) = 2.55$; $p = .014$) and PSP yielded more errors than PSS ($t(47) = 5.92$; $p < .001$). There was no difference between SPS and PSP. SSP and PPS were not significantly different from their relevant baselines (SSS and PPP).

We analysed the data separately for two groups of conditions: the first one with conditions containing less than two plurals (SSS, SSP, SPS, PSS) and the second with conditions containing two or more plurals (PPP, PPS, PSP, SPP). The group with more plurals yielded more errors than the group with fewer plurals (61 vs. 34), a difference which was significant ($F(1,46) = 8.29$, $p = .006$; $F(1,31) = 7.66$, $p = .009$). The group with more plurals was slightly more sensitive to the additional memory load (26 errors in the low condition vs. 35 in the high

condition) than the group with less plurals which even showed a tendency in the opposite direction (19 errors in the low condition against 15 in the high condition). Nevertheless, the interaction between the group and the memory load was not significant.

There was no effect of the load in the two conditions where number alternated, i.e. SPS and PSP.

Repetition errors. The memory load did not have any effect on the occurrence of repetition errors.

The number of N2 was significant with more errors occurring when it was plural than when it was singular ($F(1,46) = 10.87, p = .002$; $F(1,31) = 6.86, p = .014$). There was a significant mismatch effect between number marking of N1 and N2 ($F(1,46) = 12.05, p = .005$; $F(1,31) = 10.34, p = .003$) and between number marking of N2 and N3 ($F(1,46) = 26.14, p < .001$; $F(1,31) = 17.61, p < .001$). The three-way interaction number of N1 X number of N3 X memory load was marginally significant by subjects but not by items ($F(1,46) = 4.11, p = 0.048, F(2 < 1)$).

There were more repetition errors in the group containing two or more plurals than in the group containing less than two plurals (223 against 178), however this difference was only significant in the item analysis ($F(1,46) = 2.5, p = .12$; $F(1,31) = 5.51, p = .025$). The two groups did not differ as to their sensitivity to the memory load ($F_s < 1$).

Memory errors. There was a strong effect of the memory load, with more errors in the three-digit condition than in the one-digit condition ($F(1,46) = 45.50, p < .001$; $F(1,31) = 94.68, p < .001$).

More memory errors occurred when N2 was plural than when it was singular ($F(1,46) = 12.75, p < .001$; $F(1,31) = 15.12, p < .001$) and when N1 and N2 mismatched in number ($F(1,46) = 12.34, p = .001$; $F(1,31) = 14.97, p = .001$). The three way interaction of number of N1xnumber of N2xmemory load was significant ($F(1,46) = 5.48, p = .024$; $F(1,31) = 6.65, p = .015$), indicating that the interaction between the numbers of N1 and N2 reported just above was only significant under the high memory load condition. The mismatch between the number of N2 and the number of N3 was also a significant source of memory errors, although the analysis by items was only mildly significant ($F(1,46) = 8.78, p = .005$; $F(1,31) = 3.69, p = .064$).

Miscellaneous responses. None of the variables had a significant effect on miscellaneous errors ($F_s < 1$).

6.7 Experiment 4 (English)

6.7.1. Method

Participants. 80 undergraduate students of the University of Wisconsin-Madison participated in the present study. They received course credit for their participation. They were all native English speakers.

Materials. Same as in Experiment 3. Each sentence preamble from Experiment 2 was digitised.

Procedure. One (or three) digits were displayed for 500 ms at the centre of a computer screen 400 ms before the auditory presentation of a sentence preamble via headphones connected to the computer. Participants were instructed to memorise the digits, repeat and complete the sentence beginning, and then report the digits. Instructions emphasised accuracy in remembering the numbers as well as speed in completing the sentences.

Scoring criteria. Same as for Experiment 2.

Design and data analysis. Same as for Experiment 3.

6.7.2. Results

Application of the scoring criteria yielded 1426 (55.7%) correct responses, 134 (5.2%) agreement errors, 452 (17.7%) memory errors, 395 (15.4%) number repetition errors and 392 (15.3%) miscellaneous responses. Table 7 (Tableau 25) reports the distribution of responses in the different scoring categories for the experimental conditions.

Agreement errors. Memory load did not have a significant effect on the occurrence of agreement errors ($F(1,78) = 1.32, p = .255$; $F(1,31) = 3.52, p = .07$).

Errors were more common when N2 was plural than when it was singular ($F(1,78) = 9.56, p = .003, F(1,31) = 7.11, p = 0.012$). The interaction between the numbers of N1 and N2 was significant ($F(1,78) = 12.7, p < .001, p < .001$; $F(1,31) = 26.09, p < .001$), showing that errors were more common when there was a mismatch in number between N1 and N2, than when they matched. A mismatch effect was also found between the numbers of N2 and N3 ($F(1,78) = 13.95, p < .001$; $F(1,31) = 17.22, p < .001$). The interaction between N1 and N3 was also significant ($F(1,78) = 21.81, p < .001$; $F(1,31) = 20.67, p < .001$), but in contrast with the two previous interactions, more errors occurred when the two nouns had the same number than when they were in number mismatch.

We conducted an analysis of variance in which we contrasted memory load (low vs. high) and number of plurals in the preambles (less than two vs. two or more, see Experiment 3). The group with less plurals yielded 60 agreement errors

whereas the group with more plurals yielded 74 errors. The difference between the two groups was not significant ($F(1,78) = 1.58, p = .213$; $F(1,31) = 1.46, p = 2.236$) and there was no interaction between group and memory load ($F_s < 1$).

Memory load did not have any effect on the two conditions where number alternated, SPS and PSP (respectively $t(39) = 0.63, p = .53$ and $t(39) = 0.46, p = .65$).

Tableau 25 (Table 7) : Distribution of Responses in Experiment 4 (English).

Condition	Correct Responses	Agreement errors	Repetition errors	Miscell. responses	Memory errors
SSS	9.45 (8.36)	0.16 (0.23)	0.47 (0.23)	1.95 (1.95)	0.78 (2.58)
SSP	9.53 (6.17)	0.23 (0.16)	1.09 (0.94)	1.56 (2.11)	0.55 (4.22)
SPS	6.72 (5.94)	1.56 (1.88)	2.81 (1.95)	2.03 (1.56)	0.94 (2.89)
SPP	6.25 (4.45)	0.70 (1.25)	3.51 (4.14)	2.11 (1.88)	1.25 (4.20)
PPP	8.59 (7.03)	0.39 (0.47)	1.25 (1.17)	1.72 (2.11)	0.94 (3.13)
PPS	7.42 (6.56)	0.31 (0.23)	1.48 (1.33)	2.34 (2.03)	1.80 (3.52)
PSP	5.47 (3.44)	1.09 (1.33)	3.75 (4.84)	2.03 (1.64)	1.17 (4.40)
PSS	8.67 (7.34)	0.23 (0.23)	0.70 (1.17)	2.34 (1.25)	1.33 (3.52)
Total	62.11(49.3)	4.69 (5.78)	15.08 (15.78)	16.09 (14.53)	8.75 (26.56)

Repetition errors. The memory load did not have an effect on the repetition of number ($F_s < 1$).

Significantly more errors were produced when N2 was plural ($F(1,78) = 7.06, p = .01$; $F(1,31) = 4.86, p = .035$) as well as when N3 was plural ($F(1,78) = 34.13, p < .001$; $F(1,31) = 31.56, p < .001$). The numbers of N1 and N2 interacted, in that more errors occurred when they were different than when they were the same ($F(1,78) = 95.83, p < .001$; $F(1,31) = 93.75, p < .001$). There were also more errors when N2 and N3 were in number mismatch, as attested by the significant interaction ($F(1,78) = 11.42, p = .001$; $F(1,31) = 9.33, p = .005$). The three-way interaction between numbers of N1, N2 and N3 was also significant ($F(1,78) = 31.85, p < .001$; $F(1,31) = 16.58, p < .001$).

There were more errors in the group with more plurals than in the group with less plurals (275 against 120), a difference which was highly significant ($F(1,78) = 53.22, p < .001$; $F(1,31) = 29.60, p < .001$). However, the two groups were equally sensitive to the memory load as attested by the absence of an interaction between the group and the memory load ($F(1,78) = 1.86, p = .176$; $F(1,31) = 2, p = .167$).

Memory errors. Participants had more difficulty memorising three digits than one digit ($F(1,78) = 40.73, p < .001$; $F(1,31) = 198.66, p < .001$). More memory

errors occurred when N1 and N3 had the same number than when their numbers differed ($F(1,78) = 7.33, p = .008$; $F(1,31) = 5.08, p = .031$). There was no other significant effect ($F_s < 1$).

Miscellaneous responses. The three-way interaction between the numbers of N1 and N2 and the memory load was significant ($F(1,78) = 4.55, p = .036$; $F(1,31) = 5.47, p = .026$). There was no other effect ($F_s < 1$).

6.7.3. Discussion of Experiments 3 and 4

Experiments 3 and 4 replicated the main finding of Experiments 1 and 2: no attraction in subject-verb agreement was found with the number feature of the local noun whereas a strong attraction effect occurred with the number feature of the intermediate noun. These results come to confirm that the linear distance view is inappropriate to account for the attraction phenomena. Again, our results suggest that the view in terms of clausal packaging of encoding units is insufficient, since we found a significant difference between the attraction effect when the mismatching nonhead was in local noun position and when it was in intermediate position. This indicates that these two potentially interfering nouns, both situated in the same clause as the subject head noun, do not have a similar status for attraction as entailed by the clause packaging account. Rather, Experiments 3 and 4 further support the syntactic distance hypothesis in demonstrating once again, in both languages, that only features situated high in the tree structure can interfere with the process of verb agreement.

From results of Experiments 1 and 2, we hypothesised that the presence of two prepositional phrases in our materials introduced some additional complexity that was not present in the classical preambles used in previous studies, containing only one subject modifier. Two sources of complexity were assumed to have influenced our results. The first source was related to plurality (to account for the high amount of errors in PSP), the second was related to number alternation (to account for the higher error rate in SPS and PSP vs. SPP and PSS respectively). We further assumed that dealing with these two sources of complexity would require working memory resources and therefore predicted an interaction with the addition of a secondary task of digit memorisation. Experiments 3 and 4 showed that grouping together conditions with two or three plurals (PPP, PPS, PSP, SPP) and conditions with one or no plural (SSS, SSP, SPS, PSS) revealed that the group with more plurals yielded considerably more agreement errors than the group with less plurals in French. In English, although there were more agreement errors in the group with more plurals, the difference between the two groups was not significant. Considering errors in the repetition of number, their occurrence was considerably increased in the group with more plural features in both languages. Plurality therefore appears to be a factor influencing the occurrence of language errors. Nevertheless, the addition of the digit memorisation task did not influence at all the pattern of language errors found in Experiments 3 and 4, nor their rate of

occurrence. As a consequence, conditions with more plurals did not appear to be more sensitive to the memory load, nor were the conditions with number alternation (SPS and PSP). However, performance on the memory task was very poor, as it was found that in the one-digit condition participants produced around ten percent errors whereas more than one fourth (in French and even forty percent in English) of their responses were incorrect in the three-digit condition. This result suggests that attention given to the language task was highly predominant over the memory task, although it is not clear whether it falls within a conscious or unconscious trade-off strategy. It also indicates that the realisation of agreement is a highly automatised process which resources are not shared with other operations, no matter if they are linguistic or non linguistic.

6.8 General discussion

In this paper we reported four experiments, two in French and two in English. Our main goal was to establish the role of syntactic structure in inducing subject-verb attraction errors. We found a pattern of errors that had never been reported before: an important attraction phenomena occurred with the modifier that was not close to the verb in the uttered sentence, whereas the modifier directly preceding the verb had no interfering impact on the agreement process. The manipulation of double modifiers allowed another novel result to appear: a very high amount of attraction errors in the PSP condition. This result reduced, in English, the traditional asymmetry reported between singular and plural heads, and even suppressed it in French. The roles of syntactic structure and processing complexity are discussed in the following two sections. An alternative, still hypothetical, interpretation of our results is then put forward which suggests that the central factor determining attraction is the conceptual salience of the nouns bearing number features in the speakers' mental models.

6.8.1. The role of syntactic structure

Three hypotheses were considered that account for the attraction effect widely reported in the literature. In the framework of the linear distance hypothesis, based on linguistic accounts of agreement, the local noun effect is interpreted in terms of proximity concord. In this view, speakers would « rely spontaneously on proximity criteria, making the verb agree with the closest noun or pronoun » (Fayol et al., 1994). In most cases, the verb immediately follows the subject and agreement is correct. When another noun separates the verb from the head noun, it has been suggested that a checking procedure would ensure the relevance of the agreement relation. Different properties of the preverbal noun may influence the triggering of the editing-check process, like, for example, the presence of a plural feature or the presence of a preverbal noun that is a semantically implausible subject of the verb (Hupet et al., 1996). The four experiments showed that when two potentially

interfering number features were manipulated in the preamble, the preverbal noun (local noun) that immediately preceded the verb in the word string had no impact on subject-verb agreement. In contrast, the intermediate noun situated further from the verb but proximal to the head noun considerably disturbed agreement computation. The 'intermediate noun effect' we report closely parallels the 'local noun effect' found in studies that manipulated only one interfering, preverbal noun. This result clearly falsifies the linear distance hypothesis, as we found no effect of the local noun in any of the four experiments.

The second hypothesis we reviewed was clause packaging, as suggested by Bock and Cutting (1992); this hypothesis also fails to explain this effect. In this framework, elements that share similar structural properties within a clause are potentially interfering in their competition for the same mechanisms. The prediction of this view was that number features on the intermediate noun and on the local noun would have the same impact on verb agreement, as both nouns are part of the same clause. Results show that, within the clause, syntactic segments with similar properties do not have the same status for agreement computation. The syntactic position of the segments within the clause plays a role in that only features on segments situated high enough in the tree may interfere with the agreement process. This is not to say that clausal structure does not constitute a functional unit of sentence production (Bock and Levelt, 1994), however, it appears that it is relevant to distinguish between finer-grained units on which syntactic mechanisms operate, like the position in the hierarchical phrasal structure. Bock and Cutting (1992) admitted that some aspects of the clausal structure may influence the process of agreement: «As the size or complexity of a clause increases, the efficiency of grammatical processes that operate within that clause may decrease». Our data show that indeed the syntactic position of a number feature within the clause influences the grammatical process of agreement.

Our results are compatible with the idea that agreement is computed at the level of grammatical encoding when lemmas are retrieved and organised hierarchically before they are put in left-to-right order for articulation. The relevant factor responsible for agreement errors would therefore not be the linear distance between a local noun and the verb, nor the clausal organisation of lemmas during agreement computation, but the syntactic distance between the head noun and any potential interfering number feature in the sentence. In this framework, the mechanism for building agreement was described in terms of feature percolation from the head noun to the S node where the features would unify with features from the verb (Vigliocco et al., 1995). In this view, errors would occur when features from a nonhead noun situated high in the tree structure (i.e. close to the head noun) erroneously percolate through the tree to unify with verb number. The number feature of the local noun is too deep in the syntactic structure to percolate up the tree and influence verb agreement, as shown by the absence of an effect for N3.

It is important to note that the syntactic distance hypothesis can also account for the results reported in the literature, and is consistent with observations of speech errors which suggest that agreement errors originate during grammatical encoding and, more specifically, at a stage in which the syntactic structure of the sentence is computed, after grammatical functions have been assigned, but before words are placed in their left-to-right order in the to-be-uttered string.

6.8.2. The role of processing complexity

In contrast to previous studies, we used complex materials in which the subject head noun was separated from the verb by two modifying noun phrases. This manipulation of processing complexity revealed some interesting new information about agreement that the structural factor of syntactic distance alone cannot account for. The first observation was that PSP yielded a high amount of errors, even similar to the amount of errors in SPS in French. This finding contrasts with the small attraction effect for plural head nouns compared to singular heads reported throughout the literature. Second, although there was a high error rate in PSP, no or little attraction was reported when the plural head noun was followed by two singular nouns, i.e., PSS. The third point is the finding of a tendency for a smaller attraction effect when the singular head noun was followed by two plural nouns (SPP) compared to when the intermediate noun only was singular (SPS). These three aspects of the agreement pattern were reported consistently in all four experiments described in this paper.

To account for these findings, we suggested that a more general factor of processing complexity influenced the error pattern, in combination with the structural factor of syntactic distance. Different arguments in favour of a higher complexity of plurals were put forward (see Discussion of Experiment 2). A local noun would be more disruptive when it is plural, not because it is marked per se but because it demands more processing resources. In this framework, the finding of a high error rate in PSP, combined with a rather low error rate in PSS, and in PS as reported in the literature, could be the result of the presence of the additional plural mark on N3. Grouping results of the four experiments reported in this paper, we found indeed that language error rate (agreement and repetition) was a function of the amount of plural features in the preamble. This finding supports the idea that plurality is vector of some complexity. It is illustrated in Figure 17.

In virtue of the principles of syntactic distance and plural complexity, we expected to find a high amount of errors in SPS, as the potentially interfering intermediate noun is plural, but an even higher amount of agreement errors was expected in SPP, as a consequence of the additional plural feature. We suggested that number alternation may also constitute a variable of complexity that would have an adverse effect on the representation of the preamble in memory. The presence of two number changes would increase the difficulty of keeping track of the right number of the subject. In this framework, SPS and PSP would be

particularly prone to errors because of three factors: syntactic distance, processing complexity and number alternation.

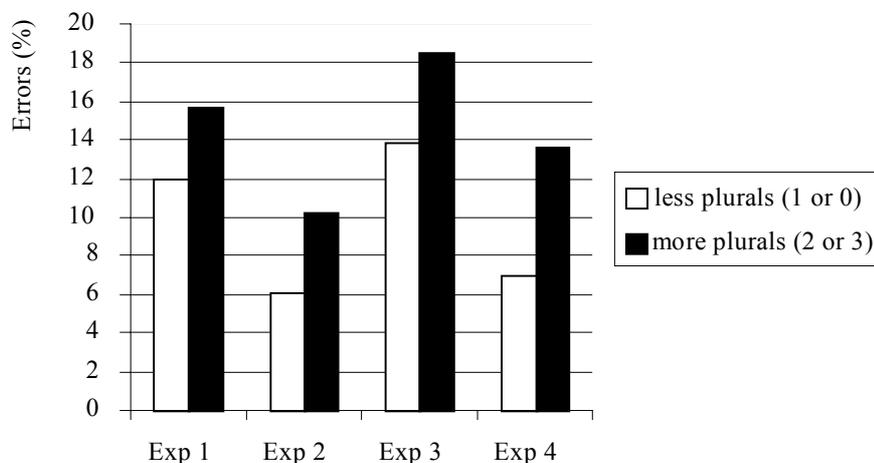


Figure 17. Errors (agreement and repetition) as a function of the amount of plural features in the preamble. In all Experiments, significantly more errors were produced in the conditions with more plurals (SPP, PSP, PPS, PPP) than in conditions with less plurals (SSS, SSP, SPS, PSS).

In Experiments 3 and 4, we asked participants to memorise digits at the same time as they were accomplishing the sentence completion task. The rationale for this manipulation was the assumption that the processing complexity related to plurality and to number alternation would consume the same cognitive resources as the digit memorisation task. Under this assumption, we predicted that the additional memory load would hamper the sentence completion task specifically in the conditions where complexity was high. We did not find this expected interaction: neither of the two tasks decreased performance in the conditions assumed to be more complex. It may be that the amount of errors produced was too small to find a significant interaction between the group (less vs. more plurals) and the memory load. There was indeed a trend in the right direction that was even significant in the subject analysis in English. Nevertheless, the very high amount of memory errors combined with the absence of an effect of the additional task on agreement errors seems rather to attest to the fact that participants focused on the sentence completion task rather than on the memorisation task. In a trade-off proceeding between the linguistic task and the memory task, most resources would have been allocated to the linguistic task. A closer look at memory errors shows that there was indeed a consistent tendency for more memory errors in conditions with more plurals than in conditions with less plurals, whatever the load condition (in English: 237 vs. 215; in French: 215 vs. 183). This may indicate that plurals do actually consume some memory resources, but that our speakers allocated all of them to the language task. In support of this, some experiments are reported in the literature that involved a dual-task paradigm (Fayol et al., 1994, Hupet et al., 1996,

1998). They all report an increase of agreement errors in the presence of an additional task, whether it was linguistic (memorisation of a list of words) or non-linguistic (click counting, number memorisation). Interestingly, error increase was higher in the condition with two plurals, PP, than in the SS condition (Fayol et al., 1994: Experiments 1, 2 and 3). The main difference between these studies and the present one is that they were all concerned with written agreement whereas we studied oral production. The highly automatised procedure of agreement in oral language is most probably responsible for the absence of an impact of a secondary task on its production, and it seems therefore sensible to conclude that a dual-task paradigm may not be appropriate to study levels of complexity in the oral format. Further studies will need to be conducted in order to establish the nature of the complexity related to plurality, as well as the apparent difficulty related number alternation.

6.8.3. Cross-linguistic considerations

The main aim of the present study was to put to test, in a cross-linguistic perspective, the syntactic distance hypothesis. This hypothesis comprises two important assumptions: (1) agreement takes place at a stage of sentence production where syntactic segments are structured hierarchically before linearisation, and (2) features from the subject percolate through the syntactic tree to reach the S-node, where agreement is ensured by a mechanism of unification between subject and verb features. Very similar error patterns were found in both languages, supporting the syntactic distance account of agreement errors. For the most part, the general pattern of results was the same in both languages. Nevertheless, three differences did emerge. We first review these three points and then provide a possible explanation for the differences.

The first difference concerns the overall number of language errors (agreement and repetition). More errors were produced in French than in English: between 27.7% (Experiment 1) and 32.3% (Experiment 3) in French and between 16.3% (Experiment 2) and 20.7% (Experiment 4) in English.

Second, for agreement errors, we found a different balance between singular and plural heads in English and in French. An asymmetry was reported for Experiment 2 in English and a strong tendency was found in the English Experiment 4, whereas there was no asymmetry in any of the French experiments. Looking at particular conditions, both languages showed a strong mismatch effect in PSP; however the symmetrical SPS condition tended to yield more errors than PSP in English, whereas the opposite tendency was found in French. In other words, proportionally more errors were produced on PSP in French than in English. Similarly, in Experiment 1 (French), there was a rather high amount of errors in PSS which was not the case in the English experiments. This pattern concords with scattered results in the literature on French agreement which often reports an attraction effect with plural heads, even if it is most of the time smaller

than the effect with singular heads (e.g., Fayol et Got, 1991; Fayol et al., 1994; Franck, 1998; Hupet et al., 1996; Hupet et al. 1998).

Along these lines, the third variation concerns the global tendency to produce erroneous singular forms versus plural forms, a tendency which was found consistent in the two experiments in each language. This tendency concerns both verbal forms (via the observation of agreement errors) and nominal forms (via the observation of repetition errors). Considering verbs, we found that English speakers produced more erroneous plural verbs than singular verbs (82 singulars vs. 132 plurals), which yielded the reported asymmetry on agreement errors. A different pattern was found for French speakers who even showed a small preference for singular verbal forms compared to plural ones (97 singulars vs. 87 plurals). Interestingly, production of nominal forms followed exactly the same tendencies as verbal forms: more plural nouns in English (219 singulars vs. 306 plurals) but more singular nouns in French (458 singulars vs. 360 plurals).

We claim here that a single factor can account for the three first cross-linguistic differences mentioned above. This factor has to do with the complexity of French plural morpho-phonology compared to English. In French, there exists many different ways of marking morpho-phonologically plurality on verbs. Most of them are audible in oral French in the present and future tenses. Singular/plural oppositions on verbs consist in vocalic alternations (e.g., [a] / [ɔ], [ɛ] / [ø]) but most commonly in a multitude of zero/consonant oppositions (e.g., zero / [t], zero / [s], zero / [v], etc.) (Dubois, 1965). In contrast, in most English verbs, number is signalled by variations in the third-person present forms and is very simple morphologically: it consists in the single [s] / zero opposition. On the other hand, the way number is marked over nouns is rather similar in both languages when one considers the written format. In English, the plural form is obtained by adding an 's' to the noun ending. In most cases, the same rule governs formation of plural nouns in French. Nevertheless, this ending is not audible in French, whereas it is in English. However, although the noun number is rarely audible in French, the determiner that precedes it is always marked for number (le / la-singular vs. les-plural), and this marking is present in the oral format. Two major differences between English and French appear from this brief description of number morpho-phonology in the two languages which concern the complexity of the verb's number. First, the morpho-phonology of complex verbal forms is by far richer in French than in English, in that there is a multitude of number oppositions in French whereas there is only one in English. A simple principle guides our interpretation here, which is that the probability of making an error is a function of the possibility of making an error. In other words, one expects to find more errors of any type but relative to number in a language that presents a greater variability of number morpho-phonological marking, like French, than in a language with a smaller variability, like English. This is what we found.

Second, plural verbal forms are morpho-phonologically simpler than singular ones in English whereas it is the opposite in French. We suggested in some previous work (see Vigliocco and Franck, 1999) that adding an inflection may be computationally more expensive than changing it. This hypothesis was put forward to account for the finding that French speakers were particularly prone to produce gender agreement errors with feminine nouns compared to Italian speakers. In this study, participants had to make the predicative adjective agree with the subject head noun. In French, the feminine form of the adjective is obtained by adding a morpheme to the masculine form, whereas it consists in a vocalic variation in Italian ([o] / [a]). Data reported in the present study support this hypothesis: we found that French participants tended to produce more errors in PSP and PSS preambles than English ones and that globally, the French tend to produce singular verbs erroneously while the English tend to produce plural verbs erroneously.

Nevertheless, one could therefore wonder why English speakers tended to produce more plural nouns than singular ones, as revealed by the analysis of repetition errors. Following our reasoning, we would expect English speakers to produce more singular nouns than plural ones, like French speakers, as in both languages plural nouns are more complex than singular ones. Our interpretation is more hypothetical here. We suggested that agreement is computed through a mechanism of feature unification, and not, as it is traditionally assumed, by a directional process of copying the subject feature onto the verb. In this framework, one could imagine that in some cases a simpler verbal morphological ending (i.e., plural) would win the unification process and therefore ‘force’ the production of a plural noun to ensure agreement. Although this way of thinking in terms of ‘reverse agreement’ seems rather counter-intuitive, it is a necessary consequence of the view we defend according to which both the subject and the verb would get their number specification from the conceptual level and unify at the stage of grammatical encoding. Repetition errors, at least errors arising on the head noun, could reflect cases where subject and verb features mismatched at the level of the S-node, a situation that the speaker resolved by changing the subject head noun’s number rather than by changing the verb’s number or producing an agreement error if unification did not succeed. English data support this view, as it was found that nearly 80% of repetition errors on the head nouns consisted in changing the singular into a plural. In other words, a repetition error would be nothing else but a way to avoid producing an agreement error. We do not claim that all repetition errors occur as a consequence of ‘reverse agreement’, as they obviously can also arise during the mapping of the conceptual representation onto the syntactic structure or even earlier during preamble comprehension as mentioned above. More research is necessary to put to test the necessary consequence of the unification hypothesis that is reverse agreement.

6.8.4. Alternative explanation: the role of conceptual salience

Although our results perfectly fit an account of agreement errors in terms of syntactic distance between the subject head noun and a potentially interfering noun, the spreading activation framework could also describe the error pattern reported. To compute verb agreement, number features of the subject have to be accessible to the system. It seems sensible to assume that the accessibility of a feature like number directly depends on the accessibility of the lexical concept that bears it (Levelt, Roelofs and Meyer, 1999). Hence it is possible that the salience of an element in the speaker's mental model determines the probability of its feature to interfere with the agreement process. In most of the cases, the subject is the most salient noun in the sentence and agreement is therefore computed correctly. Items in the experiments reported here (e.g., *The publisher of the book about the king*) consisted in a head noun and two prepositional phrases: the first one, that contained the intermediate noun, was attached to the head noun, the second one, that contained the local noun, modified the first prepositional phrase. The prepositional phrase attached to the intermediate noun (*the book*) modifies it and in that it brings some information about it; it is not an anonymous book that one is referring to but that precise book which is about the king. In contrast, nothing modifies the local noun (*the king*) which occupies the position of 'comment' about the intermediate noun. The presence of a prepositional phrase modifying the subject's modifier would come to reinforce its salience in the speaker's mental model and at the same time would keep it activated as long as something is being said about it. In this framework, the intermediate modifier will interfere with the agreement process because its activation level is high as a consequence of its salience in the mental model of the speaker.

An experiment reported by Nicol (1995) brings some support to such an account of agreement errors. The author manipulated the attachment site of the clausal modifier: it was either attached to the subject head noun (e.g., *The owner of the house who charmed the realtor(s)*) or to the local noun (e.g., *The owner of the house which charmed the realtor(s)*). Very few errors of attraction with the local noun were produced, however the analysis of another type of error consisting of the incorrect selection of the sentential head appears quite informative. Nicol found that most (87%) of these errors, in which the completion could not be semantically related to the head noun, occurred in the low-attachment condition. She concludes that attachment site seems to determine the salience of the nouns, in that attaching a modifier to the local noun increased its probability to be mistaken as the subject head. In this view, although attachment site should not have any effect on agreement errors between the verb and the noun in the clausal modifier (i.e. 'the realtors') in Nicol's experiment, it is expected that the attachment site would affect the occurrence of verb agreement errors with the noun in the prepositional phrase (i.e. 'the house').

This explanatory framework does not provide with any detail about the mechanism through which subject-verb agreement is ensured, as it is limited to an explanation of how interference can occur during the agreement process. Nevertheless, it is entirely compatible with the mechanism of feature unification sketched by the syntactic distance hypothesis and developed in Kempen and Vosse (1989), although it assumes a different cause for interference.

6.9 Conclusion

In this paper, we provided new empirical findings and a theoretical account of the attraction phenomena in both French and English. We put to question the widespread conception of subject-verb agreement errors in terms of ‘proximity concord’ according to which the number feature on a local noun in the linear proximity of the verb is attracted to construct verb agreement. We also tested the notion that agreement operates on concurrent syntactic segments within a clause, without a finer account of their syntactic position within the clause. Our results are compatible with the hypothesis that syntactic relations between the different segments can be represented through a tree-structure and that syntactic operations at that stage are sensitive to the position of each segment in this relational hierarchy. In this framework, number features of segments situated high in the tree will interfere more with the agreement process than features situated low because the path in the tree they need to travel to influence the verb is shorter. The results are also compatible with the idea that agreement is not a directional process of feature copying but rather a process of unification of features from the subject noun and from the verb.

Another factor appeared to play a role in agreement computation: processing complexity. The four experiments showed that the presence of plural features influences the occurrence of agreement errors. This finding was interpreted as indicating that plurality is a significant source of complexity that influences language processing.

Finally, it was suggested that a view in terms of feature accessibility could also account for the data observed: A segment would be more salient in the sentence if it occupies a conceptual position about which some information is brought. This salience would render its number feature more available and therefore more susceptible to disrupt the agreement process. Nevertheless, it may well be that syntactic distance and feature accessibility are just two different ways of talking about the same underlying phenomena, as important elements tend to occupy higher positions in the sentence hierarchical structure. The crucial difference is that the syntactic distance hypothesis assumes that a syntactic factor determines the occurrence of agreement errors, while the hypothesis of features accessibility focuses on a more general factor related to the accessibility of elements in the mental model speakers construct. This latter may prove an interesting heuristic for

modelling: it is expected that the attraction of a segment's feature during agreement production would be a function of its activation weight in the model.

Chapitre 7

L'hypothèse de la saillance conceptuelle

Le Chapitre 6 a mis en évidence un taux plus important d'erreurs d'attraction avec un nom se situant loin du verbe qu'avec un nom adjacent au verbe dans la chaîne phrastique. L'hypothèse de la distance syntaxique a été avancée pour expliquer ce pattern, selon laquelle le facteur déterminant dans l'effet d'attraction d'un nom interférent est sa position dans la structure hiérarchique de la phrase. Nous présentons dans ce chapitre une expérience visant à départager cette hypothèse de celle que nous avons appelée l'hypothèse de la saillance conceptuelle. Cette dernière postule qu'un nom local aura un effet interférent d'autant plus fort sur l'accord qu'il occupe une position importante au niveau conceptuel, par exemple, s'il est l'objet d'un commentaire dans la structure du message. Une proposition relative a été ajoutée aux préambules classiques et son site d'attachement a été manipulé. Bien que, globalement, l'attachement au nom local n'ait pas entraîné plus d'erreurs que l'attachement au nom sujet, une différence significative est rapportée en ce sens lorsque le nom sujet était animé et le nom local inanimé. Cette interaction semble indiquer que le site d'attachement joue effectivement un rôle sur l'assignation des fonctions syntaxiques, et par conséquent influence indirectement le phénomène d'attraction.

7.1 Introduction

L'hypothèse de la saillance conceptuelle repose sur le postulat que l'information de nombre (ou de genre) portée par un nom aura d'autant plus de poids que ce nom joue un rôle conceptuel important dans le message. Ce rôle conceptuel serait à l'origine d'un degré d'activation plus élevé de ce nom et par conséquent, de sa marque de nombre, ce qui le rendrait plus susceptible d'interférer avec l'accord du verbe. Elle est ancrée dans la perspective des modèles interactifs selon lesquels la sélection d'un élément est opérée sur la base de son degré d'activation dans le réseau activé par le message à exprimer. Cette conception permet d'intégrer l'influence directe de facteurs de nature conceptuelle, comme l'importance d'un concept dans le message, sur la construction syntaxique. Dans le chapitre précédent, nous avons avancé que cette hypothèse pouvait rendre compte de l'effet rapporté en français et en anglais selon lequel un grand nombre d'erreurs d'attraction sont apparues dans des phrases dont le nom *intermédiaire* était de nombre différent du nom sujet (Ex. 54), tandis qu'aucun effet d'attraction n'a été

observé lorsque le nom *local* était en opposition de nombre avec le nom sujet (Ex. 55).

Ex. 54 : L'étude sur les effets de la drogue

Ex. 55 : L'étude sur l'effet des drogues

L'interprétation en termes de distance syntaxique que nous avons défendue dans le chapitre précédent repose sur l'intervention d'un facteur de nature syntaxique, à savoir la position du nom interférent dans la structure hiérarchique élaborée au niveau fonctionnel de la construction syntaxique. L'influence sur l'accord de la marque portée par un nom dépendrait, selon cette hypothèse, de la distance syntaxique entre ce nom et le verbe. Cette hypothèse situe le phénomène d'attraction au niveau de l'encodage grammatical. L'hypothèse de la saillance conceptuelle suggère quant à elle que l'explication de l'effet d'attraction est à chercher au niveau de la structure conceptuelle du message. Le nom ayant entraîné un effet d'attraction dans les préambules manipulés précédemment était toujours modifié par un syntagme prépositionnel. Autrement dit, le concept correspondant à ce nom était toujours l'objet d'un commentaire, au sens large du terme, au niveau de la structure conceptuelle. Différents types de commentaires étaient introduits, répondant aux questions telles que : 'de qui ?' (ex. *l'exposition avec les peintures de l'enfant*), 'où ?' (ex. *le récit des souffrances dans l'hôpital*), 'de quoi ?' (ex. *l'entraîneur des vainqueurs du match*), 'à propos de quoi ?' (ex. *la menace des réformes dans l'école*), etc. La présence de ce commentaire à propos du nom renforcerait son rôle dans la structure conceptuelle, ce qui se traduirait par un renforcement de son activation au niveau lexical, et par conséquent de sa marque lexico-syntaxique de nombre.

Notons que Bock & Cutting (1992) ont également rapporté un nombre significativement plus élevé d'erreurs d'attraction lorsque le nom local était précédé d'un adjectif (ex. *The report of the destructive FOREST fires*, le rapport sur les feux forestiers destructeurs) que lorsqu'il n'en était pas précédé (*The report of the destructive fires*, le rapport sur les feux destructeurs).

Sur la base des données récoltées dans le Chapitre 6, nous avons également effectué une analyse fine des erreurs de répétition commises par les participants francophones. Nous ouvrons ici une parenthèse sur cette analyse parce qu'elle suggère que le même facteur de saillance conceptuelle pourrait être à l'origine à la fois des erreurs d'accord et des erreurs de répétition. Par ailleurs, les erreurs de répétition concernent près d'un quart des productions des participants francophones, elles constituent donc une source d'information non négligeable quant aux processus en jeu dans la résolution de la tâche de complètement de phrases. Pour rappel, ces erreurs consistent à changer le nombre d'un des noms dans la phrase par rapport à ce qui a été présenté (ex. *L'ordinateur avec les programmes de l'expérience-Singulier* est répété : *L'ordinateur avec les programmes des expériences-Pluriel*).

L'analyse des erreurs de répétition indique qu'elles sont loin de survenir au hasard, et qu'elles suivent au contraire un pattern bien défini. Cette analyse n'est toutefois pertinente en ce qui nous concerne que si ces erreurs reflètent une difficulté relative au 'mapping' des niveaux conceptuel et syntaxique, et non s'il s'agit simplement d'erreurs relevant de l'encodage phonologique des noms. Une telle hypothèse n'est a priori pas à exclure puisque le changement de nombre en français se marque essentiellement sur le déterminant et que l'opposition de nombre sur ce dernier repose entièrement sur le phonème final (*le* vs. *les*, *la* vs. *les*). L'analyse statistique des erreurs d'accord combinées à une erreur de répétition (ex. *L'ordinateur avec les programmes de l'expérience*-Singulier donne lieu à une réponse du type : *L'ordinateur avec les programmes des expériences*-Pluriel *SONT en panne*) permet toutefois d'écarter l'hypothèse d'une origine phonologique des erreurs de répétition. Elle montre que ces erreurs d'accord suivent exactement le même pattern que les erreurs d'accord survenant sans erreur de répétition, à condition que l'on prenne comme référence l'item tel qu'il a été produit, i.e. en tenant compte de l'erreur de répétition. Dans ce cas, on retrouve l'effet du mismatch avec le nom intermédiaire et l'absence d'un tel effet avec le nom local. Si par contre, les erreurs d'accord combinées à une erreur de répétition sont analysées sur la base des items tels qu'ils ont été présentés, sans tenir compte du changement lié à l'erreur de répétition, le pattern qui s'en dégage est radicalement différent. Ces patterns sont illustrés dans le Tableau 26 pour les deux expériences menées en français dans le Chapitre 6 (Expériences 1 et 3). Notons que les mêmes observations tiennent pour les expériences menées en anglais, bien que moins d'erreurs de répétition aient été produites.

Tableau 26 : Répartition des erreurs d'accord combinées aux erreurs de répétition, selon que la condition de référence est celle présentée (sans tenir compte de l'erreur de répétition) ou celle effectivement produite (en tenant compte de l'erreur de répétition), dans les deux recherches françaises du Chapitre 6.

	Expérience 1		Expérience 2	
	Condition présentée	Condition produite	Condition présentée	Condition produite
SSS	1	1	2	0
SSP	7	1	6	0
SPS	1	9	5	10
SPP	9	9	8	8
PPP	7	3	5	2
PPS	8	3	8	2
PSP	3	11	3	12
PSS	7	6	3	6
Total	43	43	40	40

Cette analyse indique que l'erreur de répétition a bien lieu *avant* l'erreur d'accord, puisque ce dernier est construit sur la base de l'erreur de répétition. Il

semble donc que nous puissions affirmer que les erreurs de répétition surviennent avant l'encodage phonologique de la phrase. Comme l'indiquent le Tableau 20 (p. 159) et le Tableau 24 (p. 168), ces erreurs se concentrent dans les mêmes conditions que les erreurs d'accord, i.e. les conditions SPS, SPP et PSP. Les analyses statistiques effectuées confirment la forte ressemblance des patterns de ces deux types d'erreurs : interactions entre N1 et N2 et entre N2 et N3, ainsi qu'un effet principal du nombre de N2 (plus d'erreurs lorsque celui-ci est pluriel).

Une analyse plus détaillée des noms (N1, N2 ou N3) sur lesquels portent ces erreurs s'avère éclairante quant à l'hypothèse de la saillance conceptuelle avancée pour rendre compte des erreurs d'accord. Le Tableau 27 illustre les erreurs de répétition sur les trois noms dans les différentes conditions expérimentales. Les erreurs de répétition surviendraient lorsque l'activation de la propriété de nombre du nom qui la porte est trop faible. Un nom ayant une faible saillance conceptuelle est donc un candidat susceptible de subir un changement de nombre. Dans ce cas, l'unité lexicale serait encore suffisamment active pour être rappelée, mais sa marque de nombre serait facilement soumise à des modifications²⁶. Autrement dit, la même cause serait à l'origine des erreurs d'accord et des erreurs de répétition.

Tableau 27 : Répartition des erreurs de répétition sur les trois noms des préambules. Les données des expériences en français 1 et 3 du Chapitre 6 ont été groupées étant donné la forte ressemblance de leurs distributions.

Condition	Répétition de N1	Répétition de N2	Répétition de N3	Total
SSS	3	3	12	18
SSP	3	24	21	48
SPS	18	22	32	72
SPP	11	15	33	59
PPP	4	14	17	35
PPS	6	30	28	64
PSP	10	23	32	65
PSS	10	18	12	40
Total	65	149	187	401

L'analyse globale des erreurs de répétition a révélé deux fortes interactions : entre le nombre de N1 et celui de N2 et entre le nombre de N2 et celui de N3 (voir Chapitre 6). Autrement dit, de façon générale, ces erreurs apparaissent lorsque les noms contigus portent des nombres différents. Il est intéressant d'analyser le rôle de l'entourage des noms sur lesquels surviennent les erreurs de répétition. Nous avons pour cela effectué trois analyses de variance séparées pour les erreurs de répétition sur N1, N2 et N3. L'ANOVA sur les erreurs de N1 révèle une interaction entre N1 et N2, autrement dit, un plus grand nombre d'erreurs sur N1 surviennent

²⁶ La dissociation entre un rappel correct de la racine lexicale et une erreur de répétition du nombre est un phénomène qui a été rapporté au niveau expérimental par Van der Molen & Morton (1973). Ces auteurs ont montré que dans une tâche de mémorisation, la marque de nombre était régulièrement incorrectement rappelée tandis que la racine était correctement mémorisée.

lorsque N2 est de nombre différent. Les erreurs sur N3 sont, elles aussi, influencées par la présence d'un N2 de nombre différent de N3, comme en témoigne l'interaction entre ces deux nombres. Les erreurs les plus informatives sont néanmoins celles sur N2 puisque sur ce nom-là, les locuteurs pouvaient soit anticiper la marque du N3 soit persévérer sur celle de N1 ; elles procurent donc une information supplémentaire quant à la 'direction' de l'effet de contamination par le contexte. Cela n'est pas le cas des deux autres puisque les erreurs sur N1 ne peuvent qu'anticiper tandis que celles sur N3 ne peuvent que persévérer, étant donné leur position dans la phrase. En ce qui concerne les erreurs sur N2, nous avons observé une forte interaction avec N3, mais pas d'interaction significative avec N1. Les erreurs sur N2 sont donc anticipatives, comme l'indique l'interaction avec N3. Ce comportement des locuteurs correspond à une propriété du système de production déjà largement mise en évidence tant par l'étude des lapsus (Nooteboom, 1980) que par l'étude des erreurs expérimentales (Dell, 1986 ; Dell, Burger & Svec, 1997) : en temps normal, le système tend à anticiper.

The system must activate the present, deactivate the past and prepare to activate the future. (...) when the language production system is working well, it looks to the future and does not dwell on the past. (Dell et al., 1997, p. 123)

L'hypothèse de la saillance prédit que le degré d'activation du nombre de N2 sera plus fort que celui de N3, mais moins fort que celui de N1. Trois observations vont dans ce sens. Premièrement, l'analyse indique un continuum dans les pertes de l'information de nombre : 46.6 % surviennent sur N3, 37.2% sur N2 et seulement 16.2% sur N1. Le nombre du nom local semble être plus vulnérable que celui des deux autres, tandis que le nombre du nom intermédiaire semble plus vulnérable que celui du nom sujet. Deuxièmement, bien que les erreurs sur N2 soient globalement anticipatives, certaines sont persévératives. Par ailleurs, les erreurs d'anticipation sur N1 sont presque toutes (94%) suivies d'une erreur de persévération sur N2 lorsque N1 et N2 étaient de nombre différent. Les erreurs d'anticipation sur N1 sont donc en fait des échanges avec le nombre de N2. Or, comme l'indique Dell (1986), une anticipation devient un échange parce que le degré d'activation du premier élément (ici, N1) est encore suffisamment élevé pour rebondir sur le suivant (N2). Cela indique que le nombre de N1, bien qu'il arrive qu'il soit incorrectement produit, est tout de même suffisamment activé pour entraîner un effet de rebond sur N2 et ainsi pousser le système à persévérer, malgré que cela soit contraire à son mode de fonctionnement anticipatif normal. Enfin, le même raisonnement peut être suivi pour les erreurs sur N2. Celles-ci sont majoritairement anticipatives. Toutefois, si le degré d'activation de N2 est suffisamment fort, on s'attend à ce qu'une anticipation du nombre de N3 s'accompagne d'un rebond du nombre de N2 sur N3. Autrement dit, on s'attend à ce que lorsqu'une erreur d'anticipation survient sur N2, elle soit suivie par une persévération du nombre de N2 sur N3, qu'il y ait donc échange. Nous observons

en effet que 73% des erreurs sur N2 d'anticipation du nombre de N3 sont suivies d'une erreur de persévération sur N3 lorsque N2 et N3 étaient de nombre différent.

En résumé, l'observation des erreurs de répétition indique que : (1) elles ont lieu avant l'encodage phonologique de la phrase ; (2) le taux d'erreurs sur N1 est plus faible que celui sur N2 qui est lui-même plus faible que celui sur N3 ; (3) quand bien même une erreur apparaît sur N1, elle est suivie d'une erreur de persévération du nombre de N1 sur N2 alors que le système de production est normalement anticipatif et non persévératif ; cet effet de rebond témoigne d'une forte activation de N1 ; (4) les erreurs sur N2 sont majoritairement anticipatives. Cela pourrait être pris comme un indicateur d'un haut degré d'activation de N3. Cependant, cela ne fait que confirmer les études montrant qu'un système de production normal anticipe. En faveur de cette interprétation, nous rapportons que dans la plupart des cas, une erreur d'anticipation sur N2 est suivie d'une erreur de persévération sur N3. Cet effet de rebond indique un haut degré d'activation du nombre de N2. Ces quatre arguments soutiennent l'hypothèse qu'une cause commune est à l'origine des patterns d'erreurs d'accord et de répétition²⁷ : un haut degré d'activation du nom intermédiaire, comparativement à celui du nom local. L'hypothèse de la saillance conceptuelle suggère que ces différences dans les degrés d'activation de ces deux noms proviennent de différences dans leur saillance conceptuelle au niveau de la structure du message.

Une nouvelle expérience a été conduite pour départager les deux interprétations possibles des résultats présentés dans le chapitre précédent, i.e. l'hypothèse de la distance syntaxique et l'hypothèse de la saillance conceptuelle. Une proposition relative a été ajoutée aux préambules classiques composés d'un nom sujet et d'un nom local. Le site d'attachement de la relative était manipulé, de sorte qu'elle pouvait être rattachée au nom sujet (Ex. 56) ou au nom local²⁸ (Ex. 57).

Ex. 56 : Le potage des gamins qui *coule* sur le tapis

Ex. 57 : Le potage des gamins qui *mangent* sur le tapis

L'attachement de la relative est déterminé par les contraintes sémantiques de sélection du verbe : dans l'exemple (Ex. 56), le verbe *couler* requiert que son sujet soit un élément fluide, tandis que dans l'exemple (Ex. 57), le verbe *manger* requiert un sujet vivant. La contrainte d'attachement est exclusive en ce qu'elle impose soit le nom sujet, soit le nom local comme site d'attachement, les verbes étant choisis de telle sorte qu'il ne puisse pas y avoir d'attachement aux deux sites.

²⁷ Nos résultats n'excluent pas que les erreurs de répétition soient commises au niveau du processus de compréhension des préambules, et non lors de la production de la phrase. Ce problème d'interprétation, bien qu'intéressant, est toutefois secondaire à nos préoccupations puisqu'on s'attend à ce que les effets liés au degré d'activation soient les mêmes en production et en compréhension.

²⁸ Nous utiliserons le terme de *nom local* pour désigner le nom dont nous étudions l'effet d'attraction sur l'accord, à savoir celui présent dans le syntagme prépositionnel attaché au sujet (le nom local n'est donc plus le dernier nom dans la phrase, contrairement au chapitre précédent).

L'hypothèse de la saillance conceptuelle prédit un taux d'erreurs plus élevé dans la condition d'attachement au nom local puisque la relative est supposée renforcer le poids du nom local dans la structure conceptuelle de la phrase. L'hypothèse de la distance syntaxique, par contre, ne prédit aucune différence entre les conditions puisque dans les deux cas le nom local occupe la même position dans la structure syntaxique de la phrase.

7.2 Expérience

7.2.1. Méthode

Participants. Cinquante-deux étudiants de l'Université catholique de Louvain ont participé à l'expérience dans le cadre d'un travail pratique de psychologie expérimentale. Le français était leur langue maternelle.

Matériel. Le matériel se composait de 32 items expérimentaux formés d'un préambule contenant un nom sujet (N1), suivi d'un syntagme prépositionnel contenant le nom local (N2) et d'une proposition relative. Cette dernière pouvait soit être attachée au nom sujet, soit au nom local, en fonction des contraintes sémantiques de sélection du verbe de la relative. Les variables manipulées étaient : (1) l'animation du nom sujet et du nom local (nom sujet animé, nom local inanimé vs. nom sujet inanimé, nom local animé) et (2) le site d'attachement de la relative (au nom sujet vs. au nom local).

Le nom sujet et le nom dans la relative étaient toujours au singulier, tandis que le nom local était toujours au pluriel. Nous n'avons pas introduit comme variable expérimentale la discordance entre le nombre du nom et sujet et celui du nom local étant donné que l'effet du mismatch a déjà été largement rapporté antérieurement. Le nom sujet et le nom local étaient donc toujours de nombres différents, afin de maximiser la quantité d'erreurs produites. Des exemples d'items dans les différentes conditions expérimentales sont présentés dans le Tableau 28. La totalité du matériel se trouve en Annexe 10.

Tableau 28 : Exemples d'items dans les différentes conditions expérimentales.

Condition	Item
Attachement à N1, N1 Animé, N2 Inanimé	L'architecte des bâtiments qui mesure le clocher
Attachement à N2, N1 Animé, N2 Inanimé	L'architecte des bâtiments qui dépassent le clocher
Attachement à N1, N1 Inanimé, N2 Animé	La tente des vacanciers qui perce sous l'averse
Attachement à N2, N1 Inanimé, N2 Animé	La tente des vacanciers qui courent sous l'averse

Nous avons manipulé le caractère animé/inanimé des noms sujet et local en guise de test de la difficulté d'assignation des rôles syntaxiques dans la phrase. Bock & Miller (1991) ont montré que cette variable n'avait pas d'impact sur

l'accord dans des préambules classiques (de type nom sujet + préposition + nom local), tandis que des phrases complexes étaient sensibles à ce facteur. Etant donné la présence de trois noms au sein de nos préambules, nous voulions disposer d'une mesure de la difficulté d'assignation des fonctions syntaxiques que cela engendrait.

Quarante-huit items de remplissage ont été construits en vue d'introduire des sujets pluriels ainsi qu'une certaine variabilité syntaxique dans le matériel. L'ensemble des items ont été répartis dans deux listes expérimentales.

Procédure. Les préambules étaient présentés oralement via un casque relié à un magnétophone. Les participants entendaient les préambules et devaient immédiatement les répéter et les compléter au moyen d'un complément de leur choix. Il leur était suggéré d'utiliser le verbe *être* pour la facilité. Les verbes présents dans les relatives étaient tous du premier groupe (terminaison en -ER) de telle sorte que leur nombre n'était pas audible. L'expérimentateur réglait lui-même l'apparition de l'item suivant dès que le participant avait terminé sa phrase. Les réponses ont été enregistrées.

Catégories d'erreurs. Les réponses ont été réparties en trois catégories : (1) réponses correctes, lorsque le participant produisait une phrase complète et correcte, (2) erreurs d'accord du verbe, (3) erreurs miscellanées, lorsque le nombre d'un des trois noms était changé, une partie de la phrase manquait ou rien n'était produit.

Design et analyse des données. Deux analyses de variance ont été conduites, l'une prenant les sujets comme facteur aléatoire (F1), l'autre les items (F2). Des comparaisons paires ont également été effectuées par le test *t* de Student. Les variables manipulées faisaient partie d'un design intra-sujet. Le caractère animé/inanimé des noms constituait une variable inter-items tandis que le site d'attachement était manipulé intra-items.

7.2.2. Résultats

Un total de 1420 réponses correctes (85.3%) ont été récoltées, ainsi que 101 erreurs d'accord (6.07%) et 121 erreurs miscellanées (7.27%). La répartition des erreurs d'accord dans les conditions expérimentales se trouve dans le Tableau 29.

Erreurs d'accord. L'analyse statistique des erreurs d'accord révèle que ni le site d'attachement, ni le caractère animé/inanimé n'ont influencé leur occurrence (respectivement, $F_s < 1$ et $F_1(1,51)=2.91$, $p=0.094$; $F_2(1,30)=1.56$, $p=0.115$). Une interaction entre ces deux facteurs est cependant apparue, bien qu'elle n'atteigne pas tout à fait le seuil de signification dans l'analyse par sujets ($F_1(1,51)=3.53$, $p=0.066$; $F_2(1,30)=4.77$, $p=0.037$). Les comparaisons deux à deux indiquent que dans la condition AI, un taux significativement plus élevé d'erreurs d'accord ont été produites sous l'attachement à N2 que sous l'attachement à N1 ($t_{15}=2.92$,

$p=0.01$). Elles indiquent par ailleurs un effet significatif de l'animation dans la condition d'attachement à N2 ($t_{15}=3.47$, $p<0.01$).

Tableau 29 : Répartition des erreurs d'accord du verbe dans les conditions expérimentales.

	N1 Animé, N2 Inanimé	N1 Inanimé, N2 Animé	Total
Attachement à N1	22	24	46
Attachement à N2	37	18	55
Total	59	42	101

Erreurs miscellanées. Aucun effet significatif n'est ressorti de l'analyse de ces erreurs ($F_s < 1$).

7.3 Discussion

Le site d'attachement n'a pas globalement influencé l'occurrence des erreurs d'attraction, contrairement à la prédiction avancée par l'hypothèse de la saillance conceptuelle. Ce résultat est compatible avec l'hypothèse selon laquelle l'important taux d'erreurs d'attraction du nom intermédiaire observées dans le Chapitre 6 est lié à la position syntaxique de ce nom dans la structure fonctionnelle de la phrase, et non à son statut dans la structure conceptuelle de celle-ci. Par ailleurs, l'animation du nom sujet n'a pas non plus influencé l'occurrence des erreurs d'accord. Un effet de ce facteur a été rapporté en anglais uniquement lorsque l'assignation des rôles syntaxiques était rendue difficile, par exemple par une structure syntaxique complexe, comme l'ont montré Bock & Miller (1991). L'absence d'un effet de l'animation du nom sujet pourrait donc indiquer que l'assignation des rôles syntaxiques n'a pas posé de problème malgré la présence de trois noms dans la phrase, ce qui constituerait un argument supplémentaire en faveur de l'interprétation syntaxique des erreurs d'attraction survenues dans cette recherche.

Bien que nous n'ayons pas observé d'effet principal de l'attachement, une interaction est apparue entre le site d'attachement et l'animation des noms. Cette interaction signale : (1) que le site d'attachement a influencé les erreurs d'accord lorsque le nom sujet était animé et le nom local inanimé, et cela dans le sens prédit par l'hypothèse de la saillance, (2) que sous la condition d'attachement à N2, un effet de l'animation est apparu dans la direction opposée à celle couramment rapportée, puisque plus d'erreurs d'accord ont été produites lorsque le nom sujet était animé et le nom local inanimé. Ces deux caractéristiques de la distribution des erreurs d'accord sont a priori difficiles à expliquer et pourraient sembler ne refléter que les aléas du hasard. Charvillat (manuscrit non publié) rapporte toutefois un pattern tout à fait similaire dans une étude menée sur les processus de compréhension auprès de locuteurs francophones. L'auteur a manipulé des phrases

du type Sujet+verbe+N1+N2+relative. L'attachement de la relative à N1 ou N2 était déterminé par l'accord en genre de l'adjectif dans la relative, soit avec N1, comme en (Ex. 58) soit avec N2, comme en (Ex. 59).

Ex. 58 : Les vigiles ont intercepté le colis de la passagère qui paraissait SUSPECT

Ex. 59 : Les vigiles ont intercepté le colis de la passagère qui paraissait SUSPECTE

L'animation de N1 et N2 était également manipulée. Les participants devaient estimer, dans les plus brefs délais, si l'adjectif constituait ou non une suite grammaticalement correcte de la phrase en cours. Les jugements de grammaticalité portés sur l'adjectif dans la relative sont rapportés dans le Tableau 30. Un taux particulièrement élevé d'erreurs de jugement ont été commises dans la condition AI d'attachement à N2, où seulement 45% de ces phrases étaient considérées comme correctes (alors qu'elles l'étaient toutes). Ce pattern d'erreurs en compréhension est très similaire à celui que nous avons rapporté pour les erreurs d'accord en production. En compréhension comme en production, les phrases contenant un N1 animé, un N2 inanimé et un attachement à N2 se distinguent fortement des trois autres types en ce qu'elles semblent nettement plus difficiles à traiter.

Tableau 30 : Effets de l'attachement et de l'animation sur les jugements de grammaticalité portés sur l'adjectif cible. Les données sont les pourcentages de jugements positifs (Charvillat, manuscrit non publié).

	N1 Animé, N2 Inanimé	N1 Inanimé, N2 Animé	Moyenne
Attachement à N1	67	66	66.5
Attachement à N2	45	63	54
Moyenne	56	64.5	60.25

Les études sur l'attachement de la relative à des syntagmes nominaux complexes en français ont révélé une nette préférence pour l'attachement haut, i.e. au premier nom (Zagar, Pynte & Rativeau, 1994). Cette préférence pour la clôture précoce (*early closure*), également rapportée en espagnol (Cuetos & Mitchell, 1988), a été avancée comme argument à l'encontre de la conception selon laquelle la compréhension de la phrase repose sur l'application automatique de règles de traitement syntaxique (principes d'attachement minimal et de clôture tardive, ex. Frazier, 1990), indépendamment de facteurs conceptuels et pragmatiques²⁹. Par

²⁹ Gilboy, Sopena, Clifton & Frazier (1995) ont avancé l'hypothèse que l'application de principes structuraux 'modulaires', i.e. l'attachement minimal et la clôture tardive, n'était active que dans les phrases présentant des relations de type primaire. Les phrases primaires incluent un sujet, un prédicat et leurs constituants obligatoires. Les phrases non primaires, contenant des relatives, conjonctions, et autres compléments non obligatoires seraient quant à elles analysées sur la base de principes non grammaticaux (*construal principle* : attacher le complément au domaine thématique en cours de traitement). Cette hypothèse permet de rendre compte de la variabilité au sein de la langue entre des structures de phrases différentes et de la variabilité inter-langues dans les préférences d'attachement.

ailleurs, Charvillat a montré dans une tâche off-line de production de phrases que cette préférence d'attachement était elle-même modulée par l'animation des noms impliqués : la préférence d'attachement pour N1 est significativement moindre dans des syntagmes de type IA (82%) que dans des syntagmes de type AI (95.5%).

Une particularité de la condition AI comprenant une relative attachée à N2 est qu'elle combine deux indices consécutifs concurrents : le premier indice est l'apparition d'un N1 animé et d'un N2 inanimé, celui-ci entraîne un important engagement envers l'attachement à N1 (à 95.5%). Le second indice impose toutefois que l'attachement se fasse sur N2 (par une contrainte sémantique dans notre étude en production, par une contrainte syntaxique dans l'étude de Charvillat), nécessitant une révision de l'attachement initial. Cette discordance entre les indices entraîne un sentiment d'erreur grammaticale (comme le montre l'étude en compréhension) ou la production d'une erreur d'accord. Lorsque le sujet est inanimé, l'engagement envers l'attachement à N1 est moindre (82%) ; l'attachement à N2 survenant ensuite n'entraîne donc pas de difficulté puisqu'il ne nécessite pas de révision des relations syntaxiques établies. En résumé, tandis que l'apparition d'une structure AI s'accompagne d'un engagement envers l'attachement à N1, et donc d'un grand nombre d'erreurs lorsque celui-ci n'est pas respecté, une structure IA entraîne un engagement moindre quant à l'attachement, ce qui limite considérablement le risque d'erreur puisque le système reste 'ouvert' à différentes alternatives et donc à un attachement à N2.

7.4 Conclusion

L'absence d'effet principal de l'attachement que nous rapportons suggère que le facteur principal impliqué dans l'effet d'attraction rapporté dans le Chapitre 6 relève bien du niveau syntaxique, comme nous l'avions avancé dans le cadre de l'hypothèse de la distance syntaxique. Toutefois, il apparaît que des facteurs de nature non syntaxique peuvent agir, dans une certaine mesure, sur l'effet d'attraction. Ces contraintes, relatives à l'animation des noms, à la saillance conceptuelle et à la préférence d'attachement dans des phrases non primaires (voir Note 29 page 192), interagissent en fonction de leur influence relative, pour moduler l'ampleur de l'effet d'attraction. L'hypothèse de la saillance conceptuelle suggère que les erreurs d'attraction peuvent survenir à cause de la présence d'un nom local ayant une forte saillance au niveau du message. Elle a été mise à l'épreuve en testant l'effet de l'attachement d'une relative au nom local sur l'occurrence d'erreurs d'attraction. Comme le prédisait l'hypothèse de la saillance, la présence d'une relative attachée au nom local s'est avérée être un facteur d'augmentation de l'effet d'attraction de ce nom. Toutefois, cette augmentation n'est apparue que dans les structures contenant un nom sujet animé et un nom local inanimé. Cette interaction entre la saillance du nom local (opérationalisée par l'attachement d'une relative à ce nom) et son animation, i.e. une propriété

typiquement associée à la subjectivité du nom (Bates & MacWhinney, 1982 ; Keenan, 1976), suggère que l'effet de la saillance relève de la procédure d'assignation des fonctions syntaxiques dans la phrase. Dans ce cadre, une erreur d'accord reflète indirectement une erreur dans l'assignation de la fonction sujet au nom local.

Bien que cette explication ne porte pas sur le mécanisme d'accord en lui-même mais bien sur une procédure grammaticale précédant sa réalisation, i.e. l'assignation des fonctions syntaxiques, elle semble toutefois constituer un cadre interprétatif non négligeable des erreurs rapportées dans la littérature sur l'accord en nombre. La grande majorité des recherches se sont basées sur l'observation des erreurs d'accord du verbe *être*, or, comme nous l'avons souligné, celui-ci ne comporte pas de réelles contraintes sémantiques de sélection (voir Hupet, Fayol & Schelstraete, 1998, pour une critique allant dans ce sens). On ne peut donc pas exclure que ces erreurs, ou en tout cas certaines d'entre elles, aient consisté en l'assignation incorrecte de la fonction sujet au nom local. Un examen plus attentif des compléments de phrase suivant le verbe dans les productions des participants, rapporté par Nicol (1995), indique par exemple qu'une proportion non négligeable de ces compléments ne pouvaient être rattachés qu'au nom local (3.45% de la totalité des productions, alors que les erreurs d'accord ne représentent que 0.006%).

Une meilleure estimation des rôles respectifs des différentes contraintes discutées dans ce chapitre (animation, saillance, statistique des préférences d'attachement) permettra d'avancer une explication plus précise de l'implication de ces facteurs conceptuels sur l'occurrence des erreurs d'attraction. Le nombre d'erreurs récoltées étant toujours relativement faible, une mesure plus fine des réponses des participants, i.e. des temps de réaction pour initier la production de la phrase, s'avérerait sans doute être un indicateur plus sensible de l'influence respective de ces facteurs sur le phénomène d'attraction.

Chapitre 8

Synthèse théorique

Les années nonante ont vu fleurir de nombreuses recherches exploitant l'accord grammatical comme un outil pour étudier le traitement syntaxique lors de la production d'une phrase. Nous avons contribué à ces développements par une série d'études en français, certaines ayant également été menées en parallèle en italien ou en anglais par le Dr. Gabriella Vigliocco. L'ensemble des données récoltées et discutées dans la thèse reposent sur un même effet expérimental : l'effet d'attraction sur l'accord en genre ou en nombre. Cet effet a été utilisé tantôt comme simple moyen d'augmenter le taux d'erreurs d'accord, tout en manipulant certaines propriétés conceptuelles ou morphophonologiques du nom sujet (Chapitres 2, 3 et 5), tantôt comme moyen de déterminer l'impact des propriétés lexico-sémantiques, syntaxiques ou conceptuelles du nom local sur l'interférence avec le processus d'accord (Chapitres 4, 6 et 7). A première vue, il apparaît de l'ensemble de ces observations sur l'accord qu'elles conduisent à des conclusions parfois contradictoires. Certaines d'entre elles plaident en faveur de ce que nous avons appelé l'hypothèse d'un input minimal aux traitements syntaxiques, selon laquelle le processeur syntaxique fonctionne de manière relativement isolée des autres composantes, sur la base d'un minimum d'informations. D'autres mettent au contraire en évidence l'existence d'effets de variables conceptuelles ou morphophonologiques sur les opérations syntaxiques, étayant ainsi l'hypothèse que ces dernières sont contrôlées par un maximum d'informations disponibles.

Dans cette synthèse, nous tentons de dégager un fil conducteur capable de rendre compte de l'apparente hétérogénéité des résultats. Cette hétérogénéité est envisagée sous deux angles. Premièrement, nous examinons les diverses contraintes pesant sur la réalisation de l'accord qui ont été mises en évidence par les différentes recherches : les contraintes relatives au sens, à la morphophonologie, à la syntaxe, et les contraintes extra-linguistiques. Les interprétations avancées pour rendre compte des mécanismes qui sous-tendent les effets de ces contraintes sont confrontées et situées par rapport à la question plus générale de la modularité des opérations syntaxiques. Deuxièmement, des hypothèses explicatives de la variabilité de l'effet des différentes contraintes au sein d'une même langue et entre les langues sont avancées, ainsi que leurs prédictions en vue des recherches futures.

8.1 Les contraintes impliquées dans la réalisation de l'accord

8.1.1. Les contraintes relatives au sens

Les contraintes conceptuelles

Deux types de contraintes conceptuelles ont été étudiées : celles relatives au nom sujet et celles relatives au nom local. Les propriétés conceptuelles du nom sujet ont été manipulées dans les Chapitres 2 et 3. Dans le Chapitre 2, nous avons montré que l'accord en genre de l'adjectif prédicatif est facilité par la présence d'un corrélat conceptuel de genre, comme en témoigne le moindre taux d'erreurs d'accord avec des noms ayant un genre conceptuel qu'avec des noms ayant un genre grammatical. Par ailleurs, les recherches présentées dans le Chapitre 3 indiquent que la présence d'information conceptuelle ne coïncidant pas avec l'information grammaticale constitue un facteur d'erreur considérable : un nom féminin référant à une personne de sexe masculin (et inversement) est une source d'accord particulièrement difficile à traiter. Autrement dit, l'information conceptuelle, si elle coïncide avec l'information grammaticale, favorise la réalisation correcte de l'accord, tandis qu'elle la pénalise si elle discord avec l'information grammaticale. Ces résultats sont compatibles avec l'hypothèse de l'input maximal qui postule la perméabilité du traitement syntaxique aux informations conceptuelles (voir Figure 5). Elle s'avance toutefois peu quant aux mécanismes explicatifs de la façon dont ces informations influencent le fonctionnement de la syntaxe. Nous récapitulons ici les deux mécanismes avancés dans les chapitres précédents compatibles avec l'hypothèse maximaliste, i.e. l'unification et l'interaction entre les niveaux, et la façon dont ils rendent compte des effets conceptuels rapportés.

Le mécanisme que nous avons proposé à l'issue des Chapitres 2 et 3 pour rendre compte de l'influence des propriétés conceptuelles du nom sujet repose sur la conception de l'accord comme processus d'*unification* des marques se situant sur les deux éléments à accorder. Dans un tel cadre théorique, l'encodeur grammatical récupère les marques d'accord de la représentation conceptuelle du message pour l'ensemble des éléments à accorder. Dans le cas de l'accord en nombre entre le sujet et le verbe, la pluralité notionnelle de l'agent se propage à l'ensemble des éléments conceptuels qui lui sont reliés, dont l'action. L'action elle-même est donc conçue comme plurielle et cette pluralité notionnelle au niveau extra-linguistique détermine la pluralité syntaxique du verbe au niveau linguistique. Ce mécanisme a originellement été avancé pour rendre compte de l'effet conceptuel lié à la distributivité du sujet sur l'accord du verbe (ex. Vigliocco et al., 1995). Cette variable conceptuelle, de loin la plus largement étudiée, s'est avérée influencer l'accord en italien (Vigliocco et al., 1995), en espagnol (Vigliocco et al., 1996a), en français et en néerlandais (Vigliocco et al., 1996b). Un

tel effet a également été observé en anglais, dans une moindre mesure, à condition de faire précéder le nom sujet d'un adjectif (Vigliocco, Verès & Franck, 1997) ou d'illustrer les préambules par des images (Eberhard, 1999).

Dans le cas de l'accord en genre entre le sujet et l'adjectif prédicatif, l'information concernant le sexe de l'expérimenteur se 'propage' elle aussi au sein de la représentation conceptuelle à l'ensemble des entités conceptuelles liées à l'expérimenteur. Le qualificatif est donc spécifié pour le genre déjà au niveau conceptuel, et cette spécification notionnelle détermine le choix de la marque syntaxique que l'adjectif portera. Lors de l'encodage grammatical, le mécanisme d'unification assure que les deux segments syntaxiques, i.e. le segment du nom sujet et celui de l'adjectif prédicatif, portent bien la même marque. Dans cette perspective, l'accord avec un sujet ayant un corrélat conceptuel consiste en une mise en commun des informations présentes sur le sujet et l'adjectif. Si le corrélat conceptuel de genre concorde avec la marque syntaxique, il ne fera que renforcer la probabilité d'un accord correct, comme nous l'avons montré dans le Chapitre 2. Si par contre les informations conceptuelle et syntaxique sur le nom sujet discordent, le risque d'erreur d'accord augmente, comme il a été montré dans le Chapitre 3. Lorsque le nom sujet ne réfère pas à une entité sexuée, le qualificatif n'est pas spécifié notionnellement pour le genre. L'unification consiste alors à copier la marque grammaticale du nom sujet sur celle de l'adjectif prédicatif.

Sur le plan computationnel, la détermination conceptuelle de tous les éléments soumis à un accord permet de limiter les dégâts liés à la perte d'information. Les deux éléments étant spécifiés, l'information sur l'un des deux peut être perdue sans pour autant que cela ne compromette la réalisation de l'accord. Toutefois, pour qu'un tel modèle puisse rendre compte des données observées, il doit réserver une place particulière à l'information grammaticale : en cas de désaccord entre l'information grammaticale et l'information conceptuelle, la première est, dans la majorité des cas, privilégiée par les locuteurs. Les observations indiquent que même s'il arrive que l'information conceptuelle l'emporte, cela ne concerne qu'une minorité de situations (Chapitre 3). Comme l'ont suggéré Trueswell, Tanenhaus & Garnsey (1994) en ce qui concerne le traitement syntaxique en compréhension, les contraintes conceptuelles n'ont d'effet que si elles sont suffisamment fortes et si elles concordent avec l'information syntaxique :

Contextual constraints will have weak and/or delayed effects when they are not particularly strong or when they support a structure that is not highly available because of local frequency factors. Likewise, pragmatic knowledge will not override clear syntactic constraints. And contextual dimensions that are not correlated with syntactic alternatives will rarely provide much constraint. (Trueswell et al., 1994, p. 307)

Le poids accordé à l'information conceptuelle doit donc obligatoirement différer de celui accordé à l'information syntaxique. Cette question n'est pas abordée dans le modèle de l'unification. Une perspective *interactive* de la

production peut, quant à elle, rendre compte de l'effet conceptuel tout en accordant des poids différents aux contraintes conceptuelle et syntaxique. Dans une telle perspective, la présence d'un corrélat conceptuel de nombre ou de genre congruent avec l'information lexico-syntaxique du nom sujet vient renforcer l'activation de cette dernière et ainsi faciliter la réalisation correcte de l'accord. Une information conceptuelle incongruente active par contre la représentation lexico-syntaxique de genre opposée à la propriété lexico-syntaxique du nom, il y a donc conflit entre les deux. Le fait que la résolution du conflit se fasse majoritairement en faveur de l'information lexico-syntaxique peut facilement être implémenté dans le réseau par l'augmentation du poids donné à cette information, par rapport à celui octroyé à l'information conceptuelle³⁰.

Dans le Chapitre 4, nous avons discuté la pertinence d'étudier les propriétés conceptuelles du nom local pour tester l'hypothèse notionnelle. Nous avons avancé dans ce chapitre que l'observation d'effets de ces propriétés ne permettait pas de conclure quant à une influence directe de la représentation conceptuelle sur la réalisation de l'accord, mais plutôt que ces effets témoignaient de l'influence indirecte des propriétés conceptuelles sur l'assignation des fonctions syntaxiques dans la phrase (ou d'un effet lexico-sémantique direct sur l'accord, voir section suivante). La principale propriété conceptuelle manipulée concerne la possibilité du nom local d'occuper, sur le plan conceptuel, la fonction sujet de la phrase, autrement dit d'être considéré comme l'agent du verbe (ou l'expérienceur de la qualité, selon la relation étudiée). Bock & Miller (1991) ont montré que, dans des structures syntaxiques suffisamment complexes, un nom local animé entraînait plus d'erreurs d'accord qu'un nom local inanimé, ayant un degré de subjectivité (*subjecthood*) plus élevé que ce dernier. Hupet, Fayol & Schelstraete (1998) ont quant à eux montré qu'un nom local ne pouvant pas être agent de l'action sur le plan sémantique (ex. *Sous les étoiles rêve la mariée*) entraînait significativement moins d'erreurs qu'un nom local agent potentiel du verbe (ex. **Dans les trains passent-P le contrôleur*). Enfin, l'ajout d'information conceptuelle sur le nom local semble, lui aussi, augmenter sa plausibilité en tant que sujet du verbe. Bock & Cutting (1992) ont ainsi rapporté une augmentation du taux d'erreurs lors de l'ajout d'un qualificatif au nom local, et nous avons montré dans le Chapitre 7 que la présence d'une relative attachée au nom local avait un effet similaire. Toutefois, ce dernier effet n'a été observé que lorsque le nom local était inanimé et le nom sujet animé. Une interaction est donc apparue entre la variable relative à l'animation du nom local et celle relative à sa saillance conceptuelle (manipulée par l'attachement d'une relative). Cette interaction suggère que l'effet de saillance conceptuelle relève, lui aussi, d'une assignation incorrecte de la fonction sujet au nom local.

³⁰ Notons que le mécanisme d'unification ayant lieu au stade de l'encodage syntaxique est compatible avec la perspective interactive de l'effet conceptuel puisque tous deux postulent une détermination conceptuelle du processus syntaxique. Toutefois, alors que le mécanisme d'unification est un mécanisme proprement syntaxique, la nature du mécanisme d'accord n'est pas claire dans la perspective interactive.

Ces résultats indiquent que la propriété conceptuelle de subjectivité du nom local constitue un facteur important d'interférence avec le processus d'accord. L'impact du nom local sur l'accord n'est dans ce cadre pas directement lié à une interférence avec la construction de l'accord mais bien à une interférence dans l'assignation des fonctions syntaxiques précédant cette construction. Le nom local est erronément considéré comme source de l'accord parce qu'il est incorrectement identifié comme sujet du verbe sur le plan sémantique.

Les contraintes conceptuelles étudiées sont-elles dans le lexique ?

Dans le Chapitre 4, nous avons montré que les conclusions tirées du Chapitre 2, en termes de détermination conceptuelle de la cible de l'accord, étaient peut-être hâtives, et qu'un autre cadre théorique pouvait être avancé étant donné le statut de l'information conceptuelle manipulée. La recherche menée en français et présentée dans ce chapitre indique qu'un nom local de genre conceptuel entraîne plus d'erreurs qu'un nom local de genre grammatical (de surface). Ce résultat est largement discuté dans le Chapitre 4 ; il suggère que l'effet conceptuel rapporté sur le nom sujet pourrait non pas ressortir de la structure conceptuelle du message mais bien de la présence d'une propriété sémantique lexicale de ce nom. Les noms ayant un genre conceptuel sont en effet associés à l'un ou à l'autre sexe et il est donc probable que cette association soit stockée sous forme de propriété sémantique des unités lexicales nominales. Nous pouvons donc affirmer avec certitude que les résultats rapportés dans le Chapitre 2 soutiennent l'hypothèse que le système responsable d'élaborer les relations syntaxiques entre les mots est sensible à une certaine 'sémantique du genre'. Néanmoins, si cette information est encodée dans le lexique sous forme de propriété sémantique du lemme, comme le suggère la recherche présentée dans le Chapitre 4, l'effet conceptuel sur le nom sujet ne démontre pas que de l'information conceptuelle variable, externe au système linguistique, a pénétré le système syntaxique. Il peut simplement indiquer que le système responsable de la construction phrastique est sensible aux informations sémantiques lexicales, internes au système langagier. Notons que même s'il s'agit d'un effet lexical, il va à l'encontre de la perspective modulaire stricte (minimaliste) qui ne postule pas d'effet des propriétés lexico-sémantiques sur le traitement syntaxique, ce dernier étant supposé opérer uniquement sur la base des informations lexico-syntaxiques.

Un problème d'interprétation similaire concernant les processus de compréhension est discuté par Fodor (1983). David Swinney (1979) a montré que le temps de réaction dans une tâche de décision lexicale (le participant doit décider s'il a entendu un mot ou un non-mot) est plus bref lorsque le mot est précédé d'un contexte dans lequel il est prédictible (ex. *Parce qu'il avait faim, il est entré dans la boulangerie pour acheter une baguette*, où *baguette* est le mot cible). Le même résultat est obtenu si le mot *pâtisserie* remplace le mot *baguette*. Pourtant, il a également été montré que si l'on présente au sujet le mot *tige* au lieu du mot

pâtisserie, l'effet de facilitation est également observé (pour autant que le mot amorceur soit présenté dans un court délai après la cible). Autrement dit, dans un premier temps de traitement, l'ensemble des mots associés à *baguette* (i.e. *tige* et *pâtisserie*) entraînent un effet de facilitation, indépendamment du contexte phrastique. Fodor analyse ce résultat comme un indicateur du fait que le système de compréhension ne fait pas une utilisation intelligente de l'information contextuelle ou générale, mais qu'il se base simplement sur les associations superficielles entre les mots *au niveau lexical*, sans recourir à leur pertinence pragmatique. L'effet contextuel observé n'est rien d'autre qu'une *imitation* de la pénétrabilité cognitive du système, provenant de l'existence d'associations au niveau des représentations lexicales. Pour préserver un fonctionnement cloisonné, et ainsi assurer la rapidité de son fonctionnement, le module de compréhension a trouvé comme solution d'imiter les rapports entre les choses dans le monde par des rapports entre les entrées dans le lexique mental.

Les associations permettent à des systèmes stupides de se comporter comme s'ils étaient intelligents. (...) Ce type de lien n'est pas du savoir. Ce n'est même pas du jugement. Ce n'est que le mécanisme grâce auquel on peut reconnaître plus rapidement un mot dans un contexte approprié. (Fodor, 1983, p. 108)

Bien entendu, ce réseau lexical est formé à partir de notre expérience du monde, c'est elle qui en détermine la structure. L'élaboration de la structure du système lexical, i.e. du réseau de représentations, est influencée par le savoir (au niveau de son développement) mais cette influence ne joue pas 'on-line' lors du décodage de la phrase, quand le système doit traiter ces représentations. Selon les termes de Fodor, ces connexions entre unités lexicales préservent le *fonctionnement cloisonné synchronique* du système de compréhension en le rendant *pénétrable de manière diachronique*. De même, l'effet conceptuel rapporté dans le Chapitre 2, ou encore l'effet conceptuel des noms collectifs (Bock & Eberhard, 1993 ; Bock et al., 1999) pourraient très bien mimer la pénétrabilité du traitement de l'accord à de l'information extra-linguistique alors qu'ils ne font que refléter l'influence d'une information lexico-sémantique, interne au système linguistique. Ces réflexions illustrent combien il est important de déterminer exactement ce qui est propre à un système et ce qui ne l'est pas avant de s'interroger sur sa modularité.

Pour prouver que le système de production syntaxique est non cloisonné, il faut donc montrer qu'il est capable de fonctionner de façon 'intelligente'. Il doit pouvoir ajuster son fonctionnement en fonction d'une variable contextuelle, indépendamment des associations qu'il a construites avec l'expérience. Ou encore, il doit pouvoir faire une utilisation synchronique d'une information conceptuelle qu'il n'a pas stockée de manière diachronique dans son réseau associatif. Les études sur la distributivité du sujet nous semblent éclairantes à ce propos puisqu'il ne peut y avoir d'association *permanente* entre le nombre d'un syntagme distribué

et une propriété lexicale de nombre, le sens distribué provenant de la présence simultanée d'un nom sujet singulier et d'un nom local pluriel. Etant donné l'ampleur de ces études et de leur impact sur la conception maximaliste des opérations syntaxiques, nous ouvrons ici une parenthèse de réflexion concernant le concept de distributivité. Vigliocco, Hartsuiker, Jarema and Kolk (1996) ont manipulé en français la distributivité du sujet de la phrase, à savoir la référence conceptuelle plurielle (ex. *L'uniforme des soldats*) ou singulière (ex. *Le lecteur des romans*) des préambules. L'interprétation plurielle est dite distribuée parce que la seule façon de se représenter le préambule est d'imaginer la présence d'un uniforme sur chacun des soldats, ce qui n'est pas le cas des sujets non distribués (un même lecteur peut lire plusieurs romans). Vigliocco et al. rapportent un effet de la distributivité : les verbes sont plus souvent erronément accordés au pluriel lorsque le sujet de la phrase est distribué que lorsqu'il ne l'est pas. Toutefois, une analyse plus détaillée des résultats impose que l'on nuance les conclusions tirées par Vigliocco et al. Une condition a été ajoutée dans leur design expérimental, liée à la présentation visuelle du matériel : l'adjectif prédicatif présenté aux côtés de chaque préambule pouvait soit porter le même nombre que le nom sujet (condition de congruence syntaxique, ex. *L'uniforme des soldats - NEUF*), soit porter le nombre opposé à celui du nom sujet (condition d'incongruence syntaxique, ex. *L'uniforme des soldats - NEUFS*). Les auteurs rapportent un effet principal de la distributivité, mais mentionnent également la présence d'une interaction significative entre cette variable et la congruence syntaxique. L'effet de distributivité repose en fait sur une différence de 2 erreurs entre les deux conditions lorsque le nombre du nom sujet et celui de l'adjectif sont congruents, tandis qu'il repose sur une différence de 21 erreurs lorsque ces deux nombres sont incongruents (voir Tableau 31).

Tableau 31 : Distribution des erreurs d'accord sujet-verbe en fonction de la distributivité du sujet et de la congruence syntaxique dans les études de Vigliocco et al. (1995) en italien et de Vigliocco et al. (1996b) en français.

	Vigliocco et al., 1995 (Italien)		Vigliocco et al., 1996b (Français)	
	Sujet non distribué	Sujet distribué	Sujet non distribué	Sujet distribué
Congruence syntaxique	6	11	8	10
Incongruence syntaxique	21	79	13	34
Total	27	90	21	44

L'étude de la distributivité menée en italien par Vigliocco, Butterworth & Semenza (1995) et présentée dans le Chapitre 1 comportait, elle aussi, la présentation d'un adjectif devant le préambule. A nouveau, l'adjectif pouvait être congruent ou incongruent sur le plan syntaxique avec le nom sujet. Une forte interaction est rapportée entre la congruence syntaxique et la distributivité, allant

dans le même sens que dans l'expérience en français (voir Tableau 31). Autrement dit, dans ces deux études, l'effet de distributivité semble reposer essentiellement, si pas entièrement, sur la présence d'une discordance syntaxique au sein de l'énoncé. L'effet conceptuel observé n'apparaîtrait qu'à condition que la phrase contienne une anomalie syntaxique.

Dans les deux études, la condition de cohérence syntaxique n'a entraîné qu'une très faible proportion d'erreurs. Bien que ce faible nombre d'erreurs soit un obstacle à toute interprétation théorique quant à l'effet de la distributivité dans cette condition de congruence, il constitue en soi une donnée importante puisqu'il indique qu'en 'temps normal', i.e. lorsque la phrase est correcte au niveau syntaxique, les locuteurs ne produisent que peu d'erreurs d'accord conceptuel, et que celles-ci ne surviennent réellement que dans des conditions particulières, i.e. lorsque la phrase comporte une anomalie.

Les auteurs de ces deux papiers ne discutent toutefois pas les conséquences de ces interactions pour l'interprétation théorique de l'effet de distributivité ; ils concluent que l'opération d'accord du verbe est sensible à la représentation conceptuelle du nombre du sujet, sans plus mentionner la restriction de cet effet à la présence conjointe d'une anomalie syntaxique dans la phrase. Il nous semble que l'interaction observée entre un effet conceptuel et une incohérence syntaxique est justement un élément de réponse particulièrement intéressant à la question de la perméabilité de l'opération syntaxique d'accord aux informations conceptuelles. En effet, si la présence d'une incohérence syntaxique est un facteur déterminant, voire nécessaire, dans l'apparition d'un contrôle conceptuel sur l'accord, cela peut indiquer que les locuteurs, français et italiens, peuvent en quelque sorte 'moduler' le recours à l'information conceptuelle en fonction d'éléments variables, inattendus.

Dans une recherche exploratoire, nous avons testé dans quelle mesure la présence d'une incohérence non plus syntaxique mais sémantique au sein de l'énoncé pouvait influencer l'effet de distributivité en français. Nous avons suivi la même procédure que celle utilisée par Vigliocco et al. en français, mais l'adjectif présenté pouvait soit qualifier le nom sujet sur le plan sémantique (ex. *La robe de la mariée est décousue*) soit qualifier le nom local, auquel cas le préambule contenait une incohérence sémantique (ex. *La robe de la mariée est heureuse*). L'adjectif était présenté oralement, de sorte qu'aucune information de nombre n'était fournie. Les résultats sont illustrés dans le Tableau 32. Nous n'avons pas observé d'effet principal de la distributivité.

Trop peu d'erreurs ont été récoltées pour tirer une conclusion tranchée de ces résultats, toutefois, le pattern d'erreurs suggère que l'incongruence sémantique semble prévenir tout effet de la distributivité alors qu'une tendance à cet effet semble pointer dans la condition de congruence sémantique.

Tableau 32 : Distribution des erreurs d'accord sujet-verbe en fonction de la distributivité du sujet et de la congruence sémantique en français

	Sujet non distribué	Sujet distribué
Congruence sémantique	0	5
Incongruence sémantique	8	7
Total	8	12

Il apparaît de ces recherches, d'une part, que la cohérence de la phrase joue un rôle prépondérant dans l'apparition d'un effet conceptuel sur l'accord : tandis qu'une incohérence *syntactique* semble favoriser l'apparition d'un effet de la distributivité, une incohérence *sémantique* semble bloquer tout effet de cette variable. L'interaction rapportée entre la congruence syntaxique et la distributivité du sujet constitue un argument en faveur d'un fonctionnement 'intelligent' de la syntaxe puisqu'il indique que le système peut estimer 'on line' la pertinence d'une information par rapport à une autre dans une situation d'élocution particulière. L'opérateur syntaxique est capable d'intégrer des informations syntaxiques, de repérer une anomalie et d'agir en conséquences. La présence d'une anomalie syntaxique précédant le verbe remettrait en question la fiabilité de l'information syntaxique, ce qui amènerait le locuteur à considérer d'autres informations susceptibles de guider l'opération d'accord du verbe, par exemple l'information conceptuelle. Autrement dit, l'opérateur syntaxique serait occasionnellement perméable aux informations conceptuelles lorsque l'information syntaxique est reconnue non fiable. Cette hypothèse reste à vérifier, entre autre par la récolte de données supplémentaires sur la base de notre matériel en vue de déterminer dans quelle mesure les variables relatives à l'incongruence sémantique de la phrase influencent effectivement l'effet de distributivité du sujet.

Une remarque s'impose par ailleurs quant à la validité 'écologique' des études de la distributivité. Les différentes recherches ont certes inclus des épreuves d'évaluation du matériel afin de s'assurer de l'interprétation plurielle des sujets distribués. Il n'en reste pas moins que ces items apparaissent peu naturels, et que l'interprétation plurielle semble relever d'une construction requérant un véritable effort cognitif plutôt qu'une interprétation spontanée. Prenons par exemple quelques préambules utilisés dans l'expérience en français de Vigliocco et al., tels que *Le prénom des enfants*, *La tache sur les pantalons*, ou encore *Le vase sur les tables*. Il nous semble qu'en production spontanée, un locuteur français choisira d'utiliser des noms sujets pluriels (*Les prénoms des enfants*, etc.) plutôt que de se compliquer la tâche ainsi que celle de son interlocuteur devant déchiffrer la cardinalité de tels syntagmes. Autrement dit, la capacité à déceler l'interprétation distribuée d'un sujet semble relever d'une compétence potentielle du compreneur mis en situation expérimentale plutôt que d'une réelle compétence mise en action

par le locuteur. Il est d'ailleurs assez révélateur que nous ayons observé un effet principal de la distributivité sur l'occurrence des erreurs de répétition des préambules. Les locuteurs avaient tendance à modifier le nombre du nom sujet ou celui du nom local de sorte à ce qu'ils coïncident (deux singuliers ou deux pluriels). Le caractère artificiel de la tâche nous semble limiter considérablement la validité 'écologique' des résultats, et donc leur pertinence.

Un autre effet semble ne pas pouvoir non plus s'interpréter dans le cadre de l'hypothèse lexicale qui est l'important taux d'erreurs d'accord observées lorsque le genre de surface du nom sujet (ex. *le témoin*) est en discordance avec le sexe du référent présent dans le contexte (ex. *Pauline*) (voir Chapitre 3). Cette information conceptuelle relative au sexe du référent constitue une propriété variable qui dépend entièrement de l'intention communicative du locuteur et ne peut donc pas être stockée comme propriété sémantique permanente du nom dans le lexique mental. Ce résultat suggère donc que de l'information non lexicale, externe au système linguistique, peut influencer le traitement syntaxique de la relation d'accord. Un examen attentif du matériel expérimental suggère toutefois qu'en manipulant le nom propre dans la phrase contexte, nous n'avons pas seulement introduit une entité sexuée mais aussi un nom possédant des propriétés morphophonologiques et syntaxiques. Nous avons montré que l'effet conceptuel rapporté ne pouvait pas être réduit aux caractéristiques morphophonologiques du nom propre (voir Chapitre 4). Il reste toutefois que la propriété lexico-syntaxique de genre de chaque nom propre coïncidait toujours avec sa propriété conceptuelle (le sexe), ce qui rend impossible la dissociation de ces deux facteurs au sein de cette expérience. Nous ne pouvons donc pas, pour le moment, exclure que l'effet rapporté soit en fait lié à la discordance entre l'information lexico-syntaxique véhiculée par le nom propre de la phrase contexte et l'information lexico-syntaxique du nom sujet présent dans le préambule.

8.1.2. Les contraintes relatives à la morphophonologie

L'hypothèse maximale postule également que des facteurs de nature morphophonologique influencent les opérations syntaxiques. Les résultats en français et en italien sur l'accord en genre présentés dans le Chapitre 5 vont dans ce sens. En français, nous avons montré que la présence d'un article marqué pour le genre ainsi que d'une terminaison prédictive du genre contribuent toutes deux à une réalisation plus correcte de l'accord que l'absence de ces informations de nature morphophonologique. En italien, le suffixe du nom constitue également un facteur favorable à la construction de l'accord. Le mécanisme d'unification proposé pour rendre compte de l'effet conceptuel est muet quant aux effets morphophonologiques sur les opérations syntaxiques. Les modèles interactifs (ex. Bates & MacWhinney, 1989) permettent, par contre, de rendre compte à la fois des effets conceptuels et des effets morphophonologiques. Le mécanisme responsable de ces effets d'input maximal est l'activation en parallèle des niveaux de traitement

conceptuel, lexical, syntaxique et phonologique. Comme nous l'avons discuté dans le Chapitre 5, la perspective interactive explique l'effet morphophonologique par l'existence dans le réseau de connexions biunivoques entre les niveaux de traitement syntaxique et phonologique. La rétroaction du niveau phonologique sur le niveau syntaxique assure un contrôle direct sur le déroulement du traitement syntaxique.

L'effet morphophonologique que nous rapportons dans le Chapitre 5 peut toutefois également s'interpréter dans une perspective strictement séquentielle de la production si l'on ajoute au modèle une composante de contrôle *externe au système de production*, assurée par le système de compréhension du langage (voir Figure 11). Cette hypothèse, avancée par Levelt (1989), postule que les différents aspects de la production, i.e. conceptuel, lexical, syntaxique et phonologique, sont contrôlés juste avant l'articulation de la phrase de sorte que la survenue d'une erreur est automatiquement renvoyée au 'module' qui en est responsable. La présence d'un corrélat morphophonologique de genre viendrait renforcer la probabilité de détection d'une erreur et ainsi diminuer le risque de production effective d'une telle erreur. Autrement dit, une même quantité d'erreurs d'accord seraient produites à l'origine, mais la présence d'un corrélat morphophonologique augmenterait la probabilité de détecter et ainsi de corriger cette erreur. L'intervention de cette composante dite de monitoring viendrait en quelque sorte 'mimer' une rétroaction du niveau phonologique sur le niveau syntaxique au sein du système de production.

Cette interprétation minimaliste de l'effet morphophonologique s'oppose à l'interprétation interactionniste maximaliste puisqu'elle suppose un déroulement strictement séquentiel du processus de production, préservant ainsi le postulat de cloisonnement informationnel du traitement syntaxique. Les deux interprétations sont difficile à départager par l'observation : les diverses études montrant un impact d'une variable phonologique sur les processus syntaxiques semblent en effet toujours pouvoir s'interpréter dans une perspective séquentielle (voir Levelt et al., 1999). Levelt et al. optent pour la perspective séquentielle par principe méthodologique (rasoir d'Ockham) : un tel modèle est plus simple sur le plan computationnel, il est donc préférable puisqu'il rend aussi bien compte des observations qu'un modèle connexionniste, qu'ils considèrent comme plus complexe.

En résumé, les effets conceptuels que nous rapportons dans les Chapitres 2 et 3 sont incompatibles avec une perspective modulaire stricte selon laquelle le traitement syntaxique opère sur la base des informations strictement lexico-syntaxiques. Dans cette optique, défendue par des auteurs tels que Garrett, Bock et Levelt, les processus syntaxiques subséquents à l'assignation des rôles syntaxiques sont aveugles aux informations conceptuelles extra-linguistiques, mais aussi aux informations lexico-sémantiques (voir Figure 4). L'hypothèse lexicale que nous

avançons peut être conçue comme une explication intermédiaire entre l'hypothèse modulaire stricte et l'hypothèse de l'input maximal. Comme cette dernière, elle suggère que le système syntaxique utilise un maximum d'informations pour effectuer ses opérations. Elle place toutefois une limite plus forte sur cet input en le réduisant aux informations internes au système de représentation linguistique et donc encodées de façon permanente dans les représentations du système, i.e. les informations lexico-sémantiques et morphophonologiques. Le fonctionnement de la syntaxe se fait en interaction avec les autres niveaux de traitement proprement linguistiques, à l'exclusion des aspects conceptuels, contextuels, pragmatiques, variables, relevant du contexte extra-linguistique. Pour en revenir à la question du cloisonnement informationnel, nous proposons d'appeler cette perspective *modulaire élargie* ou *d'interaction limitée* parce qu'elle postule que la composante syntaxique calcule ses représentations sur la base de différentes sources d'information pour autant que celles-ci soient spécifiques au système linguistique, i.e. les propriétés lexico-sémantiques, lexico-syntaxiques et morphophonologiques des unités lexicales, ainsi que les règles syntaxiques. Le *cloisonnement informationnel* se fait par rapport aux informations externes au système linguistique. La production est considérée comme *spécifique à un domaine* en ce qu'elle opère sur des stimuli excentriques, de nature particulière : les stimuli linguistiques.

8.1.3. Les contraintes relatives à la syntaxe

Les contraintes syntaxiques combinatoires

Les contraintes syntaxiques combinatoires relèvent de règles relatives à l'arrangement des mots dans la phrase qui ne sont pas réductibles à des propriétés inhérentes aux unités lexicales (i.e. lexico-syntaxiques). Bock & Cutting (1992) ont ainsi montré que la position du nom local dans la structure propositionnelle de la phrase était un facteur déterminant de l'effet d'attraction. Un nom local situé dans la même proposition que le nom sujet est potentiellement interférent parce qu'il entre en concurrence avec ce dernier. Cette concurrence est liée au fait que les deux noms se situent dans la même unité d'encodage et sont par conséquent traités en même temps par l'encodeur grammatical. Dans le Chapitre 6, nous avons suggéré qu'une décomposition plus fine de la structure hiérarchique élaborée au stade fonctionnel du traitement syntaxique s'imposait pour pouvoir rendre compte de l'effet rapporté. Nous avons ainsi montré que deux noms potentiellement interférents, bien qu'ils soient tous deux situés dans la même proposition que le nom sujet, n'avaient pas le même impact sur l'accord. Le nom situé le plus haut dans la structure hiérarchique de la phrase s'est avéré être une source d'erreurs considérable tandis que celui situé plus bas n'entre pas en concurrence pour contrôler l'accord. La position du nom local, au sein même de la proposition, est donc un facteur déterminant de l'influence qu'il aura sur l'opération d'accord.

Les résultats rapportés dans le Chapitre 6 ne permettent toutefois pas de déterminer dans quelle mesure l'effet syntaxique d'attraction repose strictement sur des éléments présents dans le syntagme nominal ou s'il concerne également des segments du syntagme verbal. A notre connaissance, seulement trois études ont étudié l'impact sur l'accord du verbe d'une marque présente dans le syntagme verbal. Anton-Méndez (1996) ne rapporte aucun effet du pronom clitique objet sur l'accord du verbe en Espagnol (ex. *La jardiniera-S los-P planta-S*, Le jardinier les plante). Par contre, Fayol, Largy & Lemaire (1994) ont montré qu'un pronom clitique objet de nombre différent du sujet constituait un attracteur potentiel dans la réalisation de l'accord par écrit (ex. **Il les promènent*). Hartsuiker, Anton-Méndez & van Zee (submitted) ont également montré que le pronom clitique objet direct pouvait entraîner un effet d'attraction significatif en néerlandais, toutefois dans une moindre mesure qu'un nom local modifiant le syntagme sujet (ex. **Karin zegt dat het meisje-S de kransen-P winnen-P*, *Karin dit que la fille-S les guirlandes-P gagnent-P).

Nous avons mené une recherche en français dans laquelle la marque potentiellement interférente de nombre se situait sur le syntagme nominal objet direct du verbe (Franck, 1998). Des préambules du type (Ex. 60) et (Ex. 61) ont été manipulés.

Ex. 60 : C'est l'adolescente que les boxeurs SEDUIRE
 Ex. 61 : C'est l'adolescente que ___ les boxeurs SEDUIRE

Les participants devaient compléter ces préambules au moyen du verbe présenté. Une condition de contrôle contenant un syntagme prépositionnel attaché au sujet a été ajoutée (ex. *Le chauffeur des ministres*) afin de disposer d'une ligne de base. L'utilisation de structures clivées nous a non seulement permis de tester l'attraction du nom objet mais également de tester l'effet de sa position dans la structure linéaire de la phrase. La position syntaxique de l'attracteur (dans le syntagme verbal vs. dans le syntagme nominal sujet) ainsi que sa position dans la chaîne linéaire, i.e. l'ordre des mots (Sujet-Attracteur-Verbe dans la condition contrôle vs. Attracteur-Sujet-Verbe dans l'Ex. 60 vs. Attracteur-Verbe-Sujet dans l'Ex. 61) ont été manipulées. Les résultats indiquent que les attracteurs objets génèrent un taux d'erreurs d'attraction considérable, suivant le même pattern que les attracteurs classiques, i.e. effet du mismatch avec le nom sujet et pas de différence entre les conditions SP et PS, comme illustré dans la Figure 18.

Ces résultats montrent clairement qu'une marque de nombre ne doit pas spécialement se situer dans le syntagme sujet pour interférer avec l'accord du verbe. Le taux d'erreurs avec les attracteurs objets peut toutefois difficilement être comparé à celui des attracteurs du syntagme prépositionnel étant donné que les structures clivées sont plus rares et plus complexes. Une comparaison entre les deux conditions de structures clivées est néanmoins pertinente et intéressante : au taux significativement plus élevé d'erreurs ont été produites dans la condition

Attracteur-Verbe-Sujet que dans la condition Attracteur-Sujet-Verbe, comme l'illustre la Figure 19.

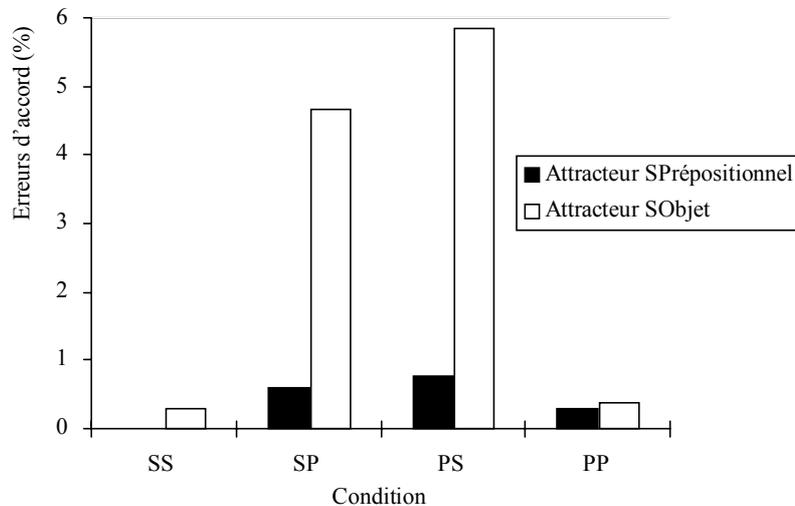


Figure 18. Distribution des erreurs d'accord du verbe en fonction du nombre de N1 et N2 (en abscisse) et du type d'attracteur (Franck, 1998). Note : Les deux conditions contenant un attracteur objet ont été moyennées.

Il pourrait être avancé que l'assignation des fonctions syntaxiques dans les préambules OVS est plus complexe que dans les préambules OSV, auquel cas l'effet rapporté ne relèverait pas de l'encodage grammatical mais bien du décodage grammatical lors de la phase de compréhension du matériel présenté. Une étude de Schelstraete & Degand (1998) en français a cependant montré que la compréhension de structures OSV était, au contraire, significativement plus ralentie au niveau du verbe que celle de structures OVS. Cela suggère que des difficultés de compréhension ne sont pas à l'origine des patterns d'erreurs que nous rapportons.

Notre recherche confirme les observations de Fayol et al. (1994) en français écrit et de Hartsuiker, Anton-Méndez & van Zee (submitted) en néerlandais selon lesquelles un nom local situé dans le syntagme verbal peut interférer avec la procédure d'accord. Elle indique également que deux structures ayant des représentations fonctionnelles (hiérarchiques) supposées identiques mais se différenciant quant à leur organisation positionnelle (linéaire) des mots n'entraînent pas des taux d'erreurs similaires : lorsque le sujet suit le verbe dans la chaîne linéaire, les locuteurs ont plus de difficultés à accorder correctement le verbe que lorsque le sujet le précède. Selon le modèle séquentiel développé par Garrett, l'unité d'encodage au niveau fonctionnel est supposée être la proposition (voir Introduction générale). Cela signifie que l'ensemble des éléments présents dans la proposition doivent être structurés dans l'arbre hiérarchique avant d'être envoyés au niveau positionnel. Certes, des contraintes lexico-syntaxiques sont à l'œuvre, mais le degré d'accessibilité des unités lexicales n'influence pas le

déroulement de la construction fonctionnelle. Dans ce cadre théorique, la structure positionnelle ne reflète pas l'accessibilité des unités lexicales puisque, peu importe l'ordre dans lequel les éléments de la proposition sont sélectionnés, l'ensemble est envoyé comme un tout, une unité, au niveau positionnel. Le niveau fonctionnel joue donc un rôle de 'filtre' en ce qu'il supprime les effets d'accessibilité des unités lexicales présentes dans la proposition sur le niveau positionnel. Pourtant, comme le suggèrent par exemple les travaux sur les structures passives (ex. Costermans & Hupet, 1977), l'utilisation d'une structure passive reflète la mise en avant du patient de l'action parce que celui-ci occupe une position clé au niveau conceptuel/pragmatique qui le rend plus accessible (par exemple parce que le patient est présent dans le modèle du discours, ce qui n'est pas le cas de l'agent). L'effet de l'ordre des mots sur les erreurs d'accord que nous rapportons dans la présente expérience relève selon nous, lui aussi, de différences quant aux représentations conceptuelles des structures OSV et OVS. La position du verbe avant le sujet dans les phrases OVS suggère qu'il a été sélectionné avant le sujet, et l'important taux d'erreurs dans cette condition par rapport aux phrases OSV reflète le fait que le nombre du sujet n'est pas encore disponible lors de l'encodage grammatical du verbe, qui a dès lors tendance à se baser erronément sur l'information de nombre déjà encodée, i.e. celle de l'objet.

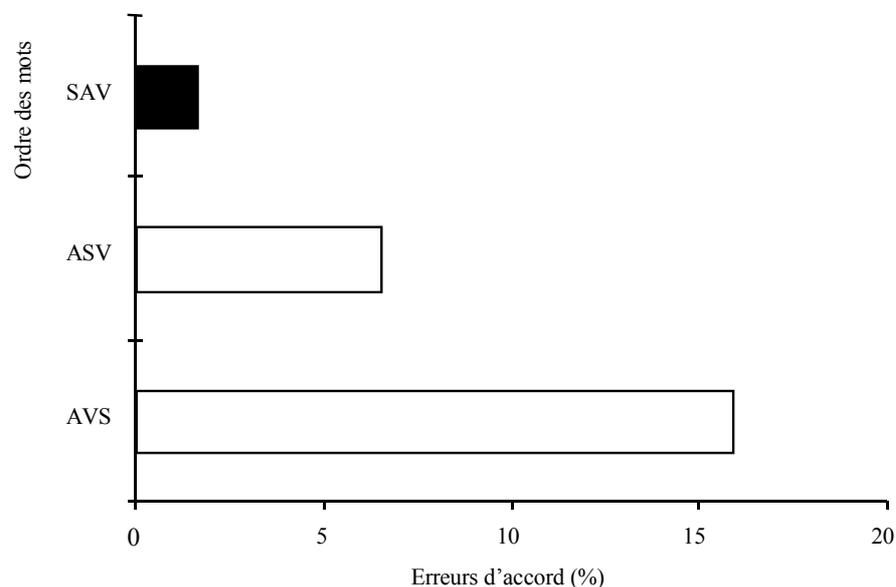


Figure 19. Pourcentages d'erreurs d'accord en fonction de l'ordre du Sujet, de l'Attracteur et du Verbe (Franck, 1998).

Cette interprétation ne remet pas en question le rôle des relations hiérarchiques entre les segments syntaxiques au niveau fonctionnel. Elle va dans le même sens que la conclusion avancée pour rendre compte de l'effet de distance syntaxique puisqu'elle suggère un découpage fonctionnel plus fin de la structure hiérarchique

que celui en termes de propositions, tel que le suggèrent Bock & Cutting. Nous suggérons que de plus petites unités, i.e. de la taille des unités lexicales, peuvent être encodées et envoyées du niveau fonctionnel au niveau positionnel pour être placées dans la chaîne linéaire, ce qui assure un mapping plus direct entre la structure conceptuelle et la structure positionnelle de la phrase via un processus de production incrémental plus fin. Une telle conception permet de rendre compte des effets d'accessibilité lexicale sur l'ordre des mots dans la phrase, mais aussi de ces effets sur un processus syntaxique comme l'accord : un élément accessible plus tôt sera encodé plus tôt au niveau syntaxique et aura d'autant plus de chances d'influencer la réalisation de l'accord, en particulier si l'élément source n'est pas encore disponible.

Les contraintes lexico-syntaxiques

Ces contraintes concernent les propriétés syntaxiques internes aux unités lexicales, et donc stockées dans le lexique. L'effet d'attraction, largement exploité dans l'ensemble des travaux sur l'accord, repose sur la présence d'un nom local de nombre ou de genre différent du nom sujet. Même si certaines études ont montré que la présence de corrélats conceptuel ou morphophonologique de nombre ou de genre sur l'attracteur peut renforcer l'effet d'attraction, ces informations ne semblent *ni nécessaires ni suffisantes* à elles seules pour susciter cet effet (voir Chapitres 1, 4 et 5). L'effet d'attraction apparaît reposer essentiellement sur la présence d'information lexico-syntaxique véhiculée par l'attracteur et différente de celle du nom sujet.

De nombreuses études sur l'accord du verbe rapportent par ailleurs une asymétrie dans l'effet d'attraction (voir Chapitre 6). L'effet d'attraction est souvent plus important dans les préambules de type SP que dans ceux de type PS, cette dernière condition n'entraînant même parfois aucune attraction. Bock & Eberhard (1993) ont suggéré d'expliquer cette asymétrie par le fait qu'un nom local « marqué » constituerait un attracteur plus fort qu'un nom local « non marqué ». La notion de marque, introduite par Jakobson (1957, cité dans Eberhard, 1997), recouvre toutefois plusieurs aspects ; l'élément marqué (pluriel/féminin) est caractérisé par la présence d'une propriété sémantique, d'une propriété syntaxique, d'une propriété morphologique, ainsi que par une plus grande fréquence d'usage (forme *par défaut*). Mais le faible impact des corrélats conceptuel et morphophonologique du nom local, comparativement à l'impact de sa propriété syntaxique de nombre, semble indiquer que cette dernière est la principale responsable de l'asymétrie entre les noms locaux pluriels et singuliers.

Cette interprétation fait reposer l'effet d'asymétrie sur des propriétés syntaxiques de l'attracteur. Il n'est toutefois pas impossible que l'asymétrie relève des propriétés de la source, i.e. le nom sujet, ou encore de la cible de l'accord. Nous avons ainsi favorisé l'explication en termes de propriétés relatives à la *cible* pour rendre compte de l'asymétrie observée dans l'accord en genre. Les recherches

que nous avons menées ont mis en évidence une tendance en italien, voire même un effet significatif en français, à un plus grand taux d'erreurs dans la condition FM que dans la condition MF. L'effet de la marque de genre apparaît donc dans le sens opposé de celui de la marque de nombre puisque la condition la plus propice à l'erreur est celle où le nom sujet est considéré comme marqué (féminin) et le nom local non marqué (masculin). Par ailleurs, dans les conditions de concordance entre les noms sujet et local, un taux d'erreurs plus élevé a été rapporté avec les noms féminins. Nous avons avancé l'hypothèse que ces effets reflètent l'utilisation de la forme la plus simple de l'adjectif sur le plan morphophonologique. Cette hypothèse explique donc l'asymétrie par des propriétés morphophonologiques (et non syntaxiques) de la cible (et non de l'attracteur). Dans ce cadre, l'asymétrie ne découle pas de l'effet d'attraction, contrairement à l'interprétation en termes de marque avancée par Bock et ses collaborateurs. Notons que cette interprétation peut également rendre compte de l'asymétrie pour le nombre en anglais : le verbe est morphophonologiquement plus simple au pluriel (ex. *they eat*, ils mangent) qu'au singulier (ex. *he eats*, il mange). Cependant, l'asymétrie pour le nombre a également été rapportée dans d'autres langues ne présentant pas cette particularité (certaines études en français, en italien et en espagnol), auxquels cas cette explication n'est pas pertinente. Rappelons toutefois que c'est en anglais que l'asymétrie apparaît être la plus forte. L'ensemble de ces données suggère que les asymétries rapportées pour le nombre et pour le genre relèvent peut-être de facteurs différents. Tandis que l'asymétrie dans l'accord en nombre reposerait sur une combinaison des propriétés syntaxiques de l'attracteur et des propriétés morphophonologiques de la cible, l'asymétrie dans l'accord en genre reposerait essentiellement sur les propriétés morphophonologiques de la cible.

Nous ouvrons ici une parenthèse sur la pertinence de la notion de marque en ce qui concerne l'opposition de genre, et plus généralement sur l'implication psycholinguistique de cette notion. Les linguistes considèrent traditionnellement que, parallèlement à l'opposition de nombre, l'opposition de genre implique une forme marquée, le féminin, et une forme non marquée, le masculin (ex., Dubois, 1965, p. 52). Sur le plan sémantique, tandis qu'une entité plurielle réfère au même concept que l'entité singulière correspondante (ex. dans *un chien* vs. *des chiens*, il s'agit toujours de *chiens*), ce n'est pas la cas pour les deux termes opposés par le genre (ex. *un homme* et *une femme* constituent deux concepts différents). Autrement dit, le lien conceptuel unissant les deux termes de l'opposition de nombre est de nature différente du lien conceptuel unissant les deux termes de l'opposition de genre. Par ailleurs, des recherches expérimentales sur la représentation lexicale du nombre et du genre semblent également indiquer que ces deux types d'oppositions ne sont pas représentées de la même manière dans le lexique. Dans une recherche récente en espagnol, Dominguez, Cuetos & Segui (1999) ont présenté des noms et des adjectifs au masculin vs. au féminin (de genre conceptuel) et au singulier vs. au pluriel dans le cadre d'une tâche de décision

lexicale. La fréquence de ces différentes formes était manipulée. Les auteurs rapportent des résultats différents pour le genre et pour le nombre. Un nom masculin est jugé plus rapidement que son correspondant féminin si la forme au masculin prédomine dans la langue, et inversement, un nom féminin est jugé plus rapidement s'il est plus fréquent que son correspondant masculin. Par contre, tandis que, comme pour le genre, un nom singulier est jugé plus rapidement que son correspondant pluriel si la forme au singulier est prédominante, un nom pluriel n'est pas jugé plus rapidement que la forme au singulier y correspondant, même si le pluriel de ce mot est plus fréquent. Cette asymétrie dans le traitement du singulier et du pluriel réplique des observations déjà obtenues en anglais et en néerlandais (Taft, 1979 ; Baayen, Dijkstra & Schreuder, 1997). Ces résultats comparatifs suggèrent que tandis que le singulier constitue bel et bien une forme 'par défaut', i.e. une base lexicale à partir de laquelle le pluriel est formé, le féminin serait par contre représenté comme une unité lexicale à part entière, distincte du masculin. Par conséquent, tandis que la distinction proposée par les linguistes entre une forme marquée et une forme non marquée semble fondée pour rendre compte de l'opposition entre le pluriel et le singulier sur le plan psycholinguistique, la notion de marque n'apparaît pas aussi pertinente en ce qui concerne l'opposition de genre.

Un dernier facteur lexico-syntaxique mis en évidence concerne la catégorie grammaticale de l'attracteur. La plupart des études portent sur des attracteurs nominaux, ceux-ci pouvant par conséquent constituer des sources grammaticalement plausibles de l'accord. La recherche menée en espagnol par Anton-Méndez (1996) sur l'attraction au sein du syntagme verbal a par contre montré une absence d'effet du pronom local objet direct. Le pronom accusatif n'est jamais une source d'accord en espagnol, ce qui pourrait avoir bloqué tout effet d'attraction. En français, Fayol, Largy & Lemaire (1994) ont toutefois rapporté un effet d'attraction du pronom clitique objet (**Il les promèment*), alors que ce dernier n'est pas un sujet potentiel du verbe sur le plan syntaxique. Néanmoins, le pronom clitique objet peut parfois être à la source d'un accord dans la langue française puisqu'il impose l'accord en nombre et en genre du participe passé (ex. *Le boxeur les* (les adolescentes-F,P) *a séduites*-F,P). Le système de traitement syntaxique est par conséquent parfois amené à devoir considérer l'information lexico-syntaxique de nombre et de genre du pronom clitique objet pour réaliser un accord. Trop peu de données sont actuellement disponibles pour pouvoir tirer une conclusion univoque de ces résultats. Toutefois, ces observations suggèrent que la catégorie grammaticale de l'attracteur constitue une contrainte lexico-syntaxique pesant sur l'accord : un mot n'étant jamais la source d'un accord dans la langue ne représenterait, dans ce cadre, pas un attracteur potentiel.

8.1.4. Les contraintes extra-linguistiques

Des contraintes plus générales de nature non linguistique sont également à l'œuvre dans l'effet d'attraction. Une première contrainte concerne les limitations de la mémoire de travail. Plusieurs travaux menés sur l'accord écrit en français par Fayol, Hupet et leurs collaborateurs indiquent en effet qu'une surcharge de la mémoire de travail contribue à augmenter le nombre d'erreurs d'accord, qu'il s'agisse d'une tâche de mémorisation de matériel linguistique, de comptage de clicks sonores ou de résolution d'additions (Fayol et al., 1994 ; Hupet et al., 1998). Par ailleurs, les auteurs ont montré que l'ampleur de cette augmentation n'est pas identique dans toutes les conditions : elle touche particulièrement les conditions de mismatch entre le nom sujet et le nom local. Hupet et al. ont aussi montré que la réduction du taux d'erreurs d'accord liée à la présence d'un nom local ne constituant pas un sujet plausible du verbe s'accompagnait d'une baisse de la performance à une tâche ajoutée de mémorisation ou de calcul. Ils en concluent que les ressources de la mémoire de travail sont mises à contribution lors du processus de vérification de l'accord.

Enfin, comme nous le discutons dans le Chapitre 6, la notion de marque semble étroitement liée à la complexité du traitement du pluriel. Cette hypothèse permet de rendre compte de l'observation d'une absence d'asymétrie entre les conditions SPS et PSP du Chapitre 6, mais aussi des observations concernant la réalisation de l'accord en situation de double tâche, le développement de la maîtrise des marques chez l'enfant et les difficultés observées chez les personnes âgées.

Les différentes recherches ont mis en évidence l'implication de contraintes relatives au sens, à la morphophonologie et à la syntaxe de la phrase dans la réalisation de l'accord. Il apparaît de cet aperçu global des travaux sur l'accord qu'une première source d'hétérogénéité provient de l'interprétation théorique des résultats. Tandis que l'influence conceptuelle sur une opération syntaxique telle que l'assignation des fonctions syntaxiques est largement reconnue et s'explique par un cadre théorique clair, ce n'est pas le cas de l'opération syntaxique d'accord. Nous avons ainsi montré que plusieurs cadres théoriques peuvent être avancés pour rendre compte des effets conceptuel, morphophonologique et syntaxique sur cette opération, pouvant se situer dans des perspectives radicalement opposées telles que la perspective interactive et la perspective modulaire stricte. Il nous semble que bon nombre des difficultés rencontrées pour interpréter les résultats émergent du manque de spécificité des modèles de production. Ceux-ci étant trop peu précis, les prédictions qui en découlent ne sont pas univoques. Un exemple particulièrement illustratif de ce problème est repris du récent article de Eberhard (1999) dans lequel

l'auteur rapporte un effet de distributivité en anglais. La conclusion tirée par l'auteur est la suivante :

Thus, evidence that a subject phrase's grammatical number and conceptual number can reliably influence verb agreement places this operation among those occurring during functional processing. (Eberhard, 1999, p. 577)

Le raisonnement logique suivi par Eberhard va exactement dans le sens opposé de celui que nous avons suivi tout au long de nos recherches, ainsi que d'autres auteurs tels que Bock & Miller (1991), Bock & Cutting (1992) ou même Bock & Eberhard (1993). Selon Eberhard, l'impact conceptuel sur le niveau fonctionnel du traitement syntaxique est un élément qu'elle considère comme acquis (et elle argumente cela sur la base de données empiriques). L'objet de la manipulation n'est donc pas de déterminer si l'opération syntaxique d'accord est sensible à un facteur conceptuel, mais bien de spécifier le stade auquel l'accord est effectué. Le point de départ théorique des autres auteurs est, par contre, la supposition que l'accord a lieu au stade fonctionnel (et ils argumentent en ce sens sur la base d'autres données empiriques). L'étude d'un effet conceptuel permet donc selon ces derniers de déterminer si oui ou non des opérations syntaxiques ayant lieu à ce stade sont perméables à de l'information conceptuelle. La prédiction porte non pas sur le stade auquel a lieu l'accord mais sur l'influence conceptuelle potentielle à ce stade. Les conclusions que nous tirons d'un effet conceptuel sur l'accord sont donc fondamentalement différentes de celles d'Eberhard, bien que les recherches soient interprétées au sein du même modèle de référence.

L'origine de cette discordance vient du manque de précision dans la définition de ce qui est considéré comme acquis et de ce qu'il reste à démontrer, autrement dit, du manque de spécification du modèle. Comme l'écrit Fodor, on ne peut pas séparer la question de l'évaluation du degré de cloisonnement informationnel d'un système de la question de l'individuation de ce système et de la caractérisation des représentations finales qu'il fournit. Dans les recherches futures, il s'agira de détailler le modèle et en particulier de préciser le statut qu'il réserve aux différentes opérations syntaxiques ainsi qu'aux représentations lexicales qu'elles impliquent comme le nombre et le genre. Cette démarche doit se faire en interaction avec l'acquisition des résultats empiriques en respectant le principe de base de la démarche scientifique, i.e. énoncer les prédictions du modèle *a priori* et l'amender *a posteriori*, sous la réserve de nouvelles données.

8.2 La variabilité dans l'application des contraintes

8.2.1. La variabilité au sein des langues

L'accord d'une cible est en général réalisé sur une base syntaxique, i.e. l'information lexico-syntaxique présente sur la source, et cela dans toutes les langues étudiées. La contrainte syntaxique semble donc être de loin la plus forte,

prédominant largement sur les contraintes de nature conceptuelle/lexico-sémantique et morphophonologique. Toutefois, comme ces différentes informations coïncident la plupart du temps, il est relativement difficile de départager leurs rôles respectifs en temps normal. Les recherches expérimentales ont tenté de séparer l'impact des contraintes syntaxique, conceptuelle et morphophonologique sur l'accord en étudiant le comportement des locuteurs dans deux types de situations.

Premièrement, une série d'études ont été menées dans lesquelles les informations conceptuelle ou morphophonologique concordait avec l'information syntaxique. L'effet de la contrainte conceptuelle est dans ce cas déterminé sur la base de la comparaison entre les taux d'erreurs en présence d'information conceptuelle et celui en l'absence de cette information (genre conceptuel vs. genre grammatical/de surface). L'effet de la contrainte morphophonologique est quant à lui estimé sur base de la comparaison entre la présence et l'absence d'information morphophonologique (terminaison prédictive ou non du genre, régularité du pluriel, article marqué ou non pour le genre ou pour le nombre). Deuxièmement, une série d'autres études ont consisté à manipuler la discordance entre l'information conceptuelle et l'information syntaxique (genre de surface opposé au sexe du référent, sujet distribué mais syntaxiquement singulier, noms singuliers collectifs, noms pluriels à référence unique) ou entre l'information morphophonologique et l'information syntaxique (noms singuliers possédant une pseudo marque du pluriel).

Une vision globale de l'ensemble des recherches met en évidence différents facteurs de variabilité au sein des langues. Ces facteurs concernent : la variabilité liée à la catégorie grammaticale de la cible, la variabilité liée à la distance séparant la cible et la source, la variabilité dans l'effet des contraintes de nature différente, la variabilité dans l'effet d'indices relevant de contraintes de même nature, et enfin la variabilité quant à l'effet des propriétés du nom sujet et celui des propriétés de l'attracteur. Nous examinons ces différentes variations intra-langue ci-dessous, à la lumière des notions de validité et de coût des indices telles qu'elles sont développées dans le modèle de compétition (Bates & MacWhinney, 1989).

L'impact de facteurs conceptuels a été étudié sur trois types de cibles d'accord : le verbe, l'adjectif prédicatif et le pronom (*tag* et réflexif). Nous avons vu qu'au sein d'une même langue comme l'anglais, la représentation conceptuelle n'influçait pas ou peu l'accord en nombre du verbe avec le sujet, tandis que l'accord du pronom était fort sensible à la numérosité conceptuelle. En fait, si l'on considère l'ensemble des langues imposant une relation d'accord entre les mots, on constate qu'elles présentent toutes certains accords syntaxiques et certains accords conceptuels. Une analyse des langues slaves a amené Corbett (1979, 1983) à suggérer que l'accord conceptuel, basé sur l'information conceptuelle plutôt que grammaticale, dépend d'une hiérarchie universelle dans laquelle sont situées les

cibles de l'accord : épithète < prédicat < pronom relatif < pronom personnel. Si une langue permet l'accord conceptuel pour l'une de ces unités syntaxiques dans la hiérarchie (par exemple le pronom relatif), alors toutes les autres unités syntaxiques se situant à la droite de celle-ci peuvent, elles aussi, s'accorder conceptuellement (le pronom personnel dans l'exemple donné).

En étudiant l'utilisation du pluriel de politesse dans différentes langues slaves et romanes, Comrie (1975) a quant à lui relevé que l'impact conceptuel sur l'accord du prédicat avec le sujet dépendait du type de prédicat (nominal, adjectival, verbal, etc.). Dans les langues slaves comme le polonais, le prédicat nominal s'accorde conceptuellement, au singulier, tandis que le prédicat verbal s'accorde avec la marque grammaticale plurielle. Le pluriel et le singulier sont tolérés pour l'accord du prédicat adjectival ainsi que pour le participe ; cependant, si le participe est accordé au singulier, l'adjectif qui le suit devra l'être aussi. Certaines langues slaves telles que le russe, le tchèque ou le serbo-croate diffèrent légèrement du polonais en ce qu'elles n'acceptent pas qu'un segment puisse prendre les deux formes : l'adjectif et le participe doivent être au pluriel, comme le verbe. Comrie en déduit une échelle dans l'accord conceptuel du prédicat : verbe < participe < adjectif < syntagme nominal. Autrement dit, plus un prédicat se comporte comme un verbe, plus il subira un accord grammatical, tandis que plus un prédicat se comporte comme un nom, plus il s'accordera conceptuellement. Cette hiérarchie se vérifie également dans les langues romanes. Par exemple, en français (mais aussi en italien et en roumain) le verbe s'accorde au pluriel avec le pronom pluriel de politesse *vous*, tandis que le participe, l'adjectif et le nom s'accordent conceptuellement au singulier (ex. *Vous êtes venu, Vous êtes loyal, Vous êtes professeur*). Comrie propose de généraliser cette hiérarchie à tous les cas où le nombre grammatical et le nombre référentiel ne concordent pas. Le verbe s'accorde toujours au singulier avec le pronom indéfini singulier *on*, tandis que le participe, l'adjectif et le syntagme nominal peuvent s'accorder au pluriel si le pronom réfère à une pluralité (ex. *On est perdus, On est fatigués, On est des bons amis*).

Un second facteur de variabilité intra-linguistique, lié au précédent, est la distance syntaxique entre les éléments impliqués dans l'accord. Nous avons vu qu'une plus longue distance syntaxique entre la cible et un élément potentiellement interférent réduit le risque d'erreur d'accord (Chapitre 6, Expérience sur les attracteurs objets directs, Bock & Cutting, 1992, Vigliocco & Nicol, 1998). Parallèlement au fait qu'une plus longue distance réduit l'impact des propriétés lexico-syntaxiques d'un nom interférent (et donc aussi probablement celles de la source), une plus longue distance semble favoriser l'utilisation de propriétés conceptuelles dans la phrase. Les travaux sur les pronoms anaphoriques (ex. Gernsbacher, 1991 ; Meyer & Bock, 1998) ont montré un recours plus important à l'information conceptuelle de l'antécédent lorsque le pronom se situe dans une autre phrase que son référent que lorsque le pronom et le référent se trouvent dans

la même phrase. La hiérarchie avancée par Corbett rend bien compte de cet effet de distance entre la source et la cible sur l'utilisation de contraintes conceptuelles pour contrôler l'accord : le pronom, qui est aussi la cible souvent la plus éloignée du nom, se montre le plus sensible à l'influence conceptuelle, tandis que l'épithète, accolé au nom, y est le moins sensible. Le prédicat verbal est quant à lui plus imperméable que le participe, lui même plus imperméable que l'adjectif. Cet ordre, défini sur la base de la catégorie grammaticale des mots, semble coïncider avec la proximité entre la source et la cible dans la structure de la phrase.

Il est intéressant de noter que l'hypothèse d'une augmentation du contrôle conceptuel avec la distance séparant la source et la cible entre parfaitement dans le cadre du principe de limitation des coûts dans l'utilisation d'un indice. En tant que locuteur, il nous arrive à tous de ne plus nous souvenir de la source d'un accord, celle-ci ayant été produite trop longtemps avant la cible. Pour produire un accord syntaxique dans ce cas, l'information lexico-syntaxique de la source doit avoir été stockée en mémoire de travail. Le coût d'un tel stockage augmente rapidement avec la distance, d'autant plus que la probabilité d'interférences s'accroît au fur et à mesure de la production de matériel linguistique entre la source et la cible. L'utilisation de l'indice lexico-syntaxique de la source devient dès lors fort coûteuse. Par contre, le locuteur a toujours à l'esprit l'idée qu'il est occupée à exprimer ; l'information conceptuelle est donc toujours disponible. Le système responsable de l'accord se basera par conséquent plus facilement sur un indice conceptuel puisqu'il est plus disponible et utilisable à un coût moindre que l'information lexico-syntaxique. La contrainte psycholinguistique de limitation des coûts rend ainsi compte des analyses objectives faites par les linguistes sur les occurrences d'accord conceptuel dans la langue.

Tandis que la distance entre la source et la cible favorise l'accord conceptuel, elle semble par contre réduire l'utilisation d'indices morphophonologiques. Meyer & Bock (1998) rapportent un effet des propriétés morphophonologiques de l'antécédent sur l'accord du pronom seulement lorsque ce dernier se situe dans la même phrase que l'antécédent. L'information morphophonologique requiert, comme l'information lexico-syntaxique, d'être stockée en mémoire de travail si la cible n'apparaît pas immédiatement. Ces deux informations, contrairement à l'information conceptuelle, sont donc particulièrement sensibles à un facteur tel que la distance puisque celle-ci augmente la charge en mémoire et donc le coût de leur utilisation. Autrement dit, tandis que l'augmentation de la distance entre la source et la cible réduit l'impact de variables linguistiques (lexico-syntaxiques et morphophonologiques), elle augmente l'impact de variables extra-linguistiques (conceptuelles).

Une série d'autres observations indiquent que le recours à l'information conceptuelle varie en fonction de la validité de l'indice lexico-syntaxique. Premièrement, nous avons souligné que les études sur la distributivité menées en

italien et en français suggèrent que la présence d'une discordance syntaxique au sein de la phrase puisse moduler l'effet de ce facteur conceptuel, voire même être entièrement responsable de l'apparition de cet effet (voir Tableau 31). Par contre, une incongruence sémantique préviendrait tout contrôle conceptuel sur l'accord. Autrement dit, l'opérateur syntaxique fonctionne comme s'il pouvait 'choisir' l'information à transmettre à la cible en fonction de la fiabilité relative des indices syntaxique et conceptuel dans la phrase. La présence d'une anomalie syntaxique amènerait le système à recourir plus facilement à l'information conceptuelle, en vue de construire l'accord sur des bases plus fiables. Deuxièmement, nous avons rapporté dans le Chapitre 3 une utilisation abondante de l'indice conceptuel en situation de discordance entre l'information grammaticale de genre et le sexe du référent (*la victime* référant à un homme). Parallèlement à cet effet conceptuel massif, nous avons constaté la disparition de l'effet syntaxique d'attraction : un nom local de même genre que le nom sujet entraînait autant d'erreurs qu'un nom local de genre différent. Autrement dit, l'importance de l'indice conceptuel relatif au nom sujet semble prédominer sur l'indice lexico-syntaxique présent sur le nom local. Troisièmement, Bock & Eberhard (1993) ont observé une tendance envers un effet conceptuel du nom local mais uniquement dans la condition où le nom local et le nom sujet étaient en mismatch, autrement dit quand une difficulté syntaxique se présentait. Enfin, les propriétés sémantiques d'animation et de concrétude du nom local manipulées par Bock & Miller (1991) n'ont eu d'influence sur l'accord que lorsque la structure syntaxique des préambules était suffisamment complexe (propositions relatives). Ces quatre observations concordent pour suggérer que le système syntaxique responsable de réaliser l'accord module son recours aux informations conceptuelles en fonction de la validité de l'information syntaxique ; les indices conceptuels sont plus facilement considérés lorsque la syntaxe de la phrase comporte une anomalie ou simplement lorsque le traitement syntaxique est rendu difficile. Cette validité n'est pas représentée dans le système de façon permanente ; elle est estimée 'on-line' dans une situation d'élocution particulière.

L'information conceptuelle est donc utilisée lorsque l'information lexico-syntaxique est trop coûteuse ou encore lorsque le traitement syntaxique de la phrase est délicat. Il est également possible d'augmenter le recours à cette information, sur le plan expérimental, en lui donnant plus d'importance. Eberhard (1999) a ainsi constaté un effet de la distributivité du sujet en anglais lorsque les préambules étaient illustrés par des images. Cette observation doit attirer notre attention sur le fait que les productions étudiées en laboratoire sont probablement peu propices à mettre en évidence l'impact des contraintes conceptuelles sur le traitement syntaxique. Les représentations conceptuelles des préambules sont probablement fort peu influentes étant donné qu'elles doivent être construites en dehors de tout contexte de communication, et sur la base d'un matériel expérimental souvent abstrait. Cette remarque ne remet pas en question les

différentes conclusions tirées plus haut quant à l'impact respectif des différentes contraintes. Elle suggère simplement que la contrainte conceptuelle joue peut-être un rôle plus important en situation d'élocution spontanée qu'il apparaît dans les situations expérimentales manipulées.

Les notions de validité et de coût d'un indice permettent également de rendre compte de la variabilité rapportée quant à l'effet morphophonologique (voir Chapitre 5). Nous avons montré que l'indice morphophonologique sur le suffixe avait un impact important sur l'accord en italien, tandis qu'un autre indice de même nature, présent sur l'article, n'exerçait aucun contrôle. Par contre, ces deux indices se sont avérés jouer un rôle en français, bien que l'indice sur l'article semble même légèrement prédominer sur l'indice nominal. Ces différences constatées au sein de ces deux langues peuvent s'expliquer par le fait que tandis que l'information sur le nom constitue un indice à la fois valide et fiable en italien, la validité et la fiabilité de cet indice sont moindres en français. Les locuteurs italiens peuvent par conséquent se permettre, pour limiter les coûts, de ne considérer que cet indice pour construire l'accord, tandis que les locuteurs français doivent également recourir à l'indice sur l'article, fort valide mais peu fiable, pour compenser la faiblesse de l'indice nominal.

Enfin, une dernière source de variabilité intra-linguistique concerne l'effet différentiel d'indices de même nature selon qu'ils se situent sur le nom sujet ou sur le nom local. Les comparaisons rendues possibles par la manipulation des mêmes indices suggèrent que l'effet conceptuel du nom local est de moindre ampleur que celui du nom sujet. De même, les manipulations des propriétés morphophonologiques du nom local ont mis en évidence des effets moindres à ceux des propriétés du nom sujet, voire même inexistantes tandis que les propriétés du nom sujet s'avéraient significatives. La relation privilégiée entre la source et la cible est plus que probablement l'explication de cette différence. La relation unissant le sujet et la cible est plus forte, à la fois au niveau conceptuel et au niveau syntaxique, que la relation entre le nom local et la cible. Il est évident que les indices portés par la source, de quelque nature qu'ils soient, doivent assurer un plus grand contrôle de l'accord que ceux du nom local, ce qui explique l'ampleur de leurs effets.

L'ensemble des travaux que nous avons rapportés dans la thèse concernent l'accord tel qu'il est réalisé par des adultes. Quelques travaux réalisés chez les enfants semblent indiquer que la procédure syntaxique suivie n'est pas la même. Les enfants commettent bien sûr plus d'erreurs, mais la nature de ces erreurs semble également se modifier au cours du développement langagier (Fayol, Hupet & Largy, 1999). Ce résultat indique que l'apprentissage de l'accord ne consiste pas simplement à progressivement mieux maîtriser une procédure mais bien à en transformer la nature au fil du développement. Une perspective strictement modulaire de la syntaxe, souvent innéiste (ex. Chomsky, 1957 ; Pinker, 1997),

soutient qu'une telle opération doit s'effectuer indépendamment des facteurs conceptuels et phonologiques, peu importe le stade de son développement. On sait pourtant que le développement du langage commence par l'acquisition de compétences sémantiques et phonologiques, et que par conséquent il n'est pas impossible que le développement syntaxique, suivant celui des deux autres composantes, s'appuie sur elles. Dans le cadre d'une perspective fonctionnaliste (ex. Bowerman, 1973), on pourrait envisager que l'acquisition de la procédure syntaxique d'accord serait au départ guidée par les liens conceptuels entre les éléments à accorder, pour progressivement se détacher de la représentation conceptuelle et se 'grammaticaliser' au cours du développement. De même, les régularités phonologiques de nombre et de genre pourraient, chez les jeunes enfants, servir d'indices pour la production syntaxique (pour autant que ces régularités soient maîtrisées), et cela dans une plus grande mesure que chez les adultes. La procédure d'accord, au départ entièrement contrôlée par des facteurs non syntaxiques, deviendrait progressivement une routine fonctionnant sur la base d'un nombre moins important d'informations, et privilégiant l'information lexicosyntaxique. Cette évolution du développement syntaxique, allant dans le sens d'une plus grande autonomie de fonctionnement, assurerait ainsi une plus grande efficacité dans bon nombre de situations d'accord.

Dans les langues étudiées, l'accord est généralement contrôlé par des contraintes d'ordre syntaxique. Cette information est non seulement la plus valide, puisqu'un accord, pour être correct, doit s'appuyer sur elle, elle est également utilisable à un moindre coût. Toutefois, des contraintes conceptuelles et/ou morphophonologiques peuvent occasionnellement contrôler cette opération, voire même l'emporter sur l'indice syntaxique. Nous avons montré que la variabilité observée, au sein des langues, entre des indices de même nature semble reposer d'une part sur des contraintes cognitives, extra-linguistiques, de nature attentionnelle et mnésique liées au coût de l'indice (la catégorie grammaticale du mot cible, la distance qui le sépare de la source, i.e. la disponibilité de l'indice), d'autre part sur des contraintes linguistiques, propres au fonctionnement du système syntaxique (la validité et la fiabilité des indices). Les recherches futures devront s'attacher à quantifier de façon plus précise et dans des situations d'élocution plus naturelles l'impact relatif de chacun des indices conceptuel, syntaxique et morphophonologique sur le contrôle de l'accord. Des études développementales pourraient également s'avérer éclairantes en ce qu'elles permettront de mettre en évidence l'évolution de l'utilisation des différentes contraintes qui pèsent sur l'accord. Nous suggérons également, à titre d'hypothèse, que le stade de développement du langage constitue lui aussi un facteur de

variabilité intra-linguistique reposant sur des contraintes à la fois cognitives et linguistiques.

8.2.2. La variabilité entre les langues

Les recherches sur l'accord ont également mis en évidence une relative disparité entre les langues, qui concerne essentiellement l'impact des contraintes non syntaxiques sur la construction de l'accord.

Sur le plan linguistique, si le continuum entre le verbe et le nom proposé par Comrie pour rendre compte de l'impact conceptuel sur l'accord du prédicat semble bien être universel, la *frontière* où commence l'accord conceptuel varie par contre d'une langue à l'autre. Par exemple, les langues slaves tolèrent un accord grammatical de l'adjectif et du participe avec le pronom pluriel de politesse, tandis que les langues romanes réservent l'accord grammatical au seul verbe. Le grec moderne, contrairement aux autres langues, tolère même un accord conceptuel du verbe. Au niveau psycholinguistique, tandis que les locuteurs italiens, espagnols, français et néerlandais commettent facilement des erreurs d'accord conceptuel, les locuteurs anglophones y sont nettement plus résistants. Une explication plausible de ces différences dans le comportement des locuteurs est qu'elles reposent sur la délimitation de la frontière entre accord conceptuel et grammatical au niveau linguistique. Les locuteurs d'une langue tolérant de nombreux accords conceptuels, et pour reprendre la hiérarchie de Comrie, ayant placé la frontière de l'accord conceptuel le plus près du verbe, auront davantage tendance à produire des erreurs conceptuelles que les locuteurs d'une langue dont la frontière se situe près du syntagme nominal. En effet, il est fort probable que lorsque l'opération d'accord doit fréquemment tenir compte de la représentation conceptuelle du message pour produire une phrase correcte, elle procède occasionnellement de la sorte dans des situations non appropriées. Un premier facteur de variation serait donc relatif à la position de la frontière de la référence conceptuelle dans le continuum verbe-syntagme nominal. Cette frontière détermine la validité générale de l'information conceptuelle ; la validité augmentant à mesure que la frontière se rapproche du verbe, les locuteurs seraient d'autant plus enclins à recourir à cet indice.

Par ailleurs, les langues ont développé des systèmes de représentation flexionnelle du nombre et du genre fort différents. Considérons l'anglais. L'anglais ne marque le genre que sur certains pronoms, et se limite à marquer le nombre sur le nom, le pronom et le verbe, et encore, seulement dans certaines conjugaisons. Les marques de nombre sur le nom et sur le verbe se limitent à l'opposition présence vs absence de *-s* final³¹. Le faible impact des facteurs conceptuel et morphophonologique en anglais peut s'expliquer par le fait que la procédure d'accord, étant relativement simple puisqu'elle se limite essentiellement à l'accord en nombre entre le sujet et le verbe, peut se contenter de l'information lexico-

³¹ Ces caractéristiques sont explicitées dans la section 1.2.1.

syntaxique pour deux raisons. Premièrement, l'indice lexico-syntaxique est extrêmement *valide*, i.e. il est fiable et toujours présent. Il est toujours correct d'utiliser l'information lexico-syntaxique de la cible pour construire l'accord. Seuls quelques rares cas d'accord du pronom tolèrent un accord conceptuel, mais aucun cas ne l'impose jamais dans la langue. Autrement dit, la fiabilité de l'indice lexico-syntaxique est maximale. Par ailleurs, la source de l'accord, i.e. le sujet de la phrase, est toujours présente dans la structure de la phrase. L'information lexico-syntaxique est donc toujours accessible.

L'information conceptuelle est également un indice disponible et relativement fiable puisqu'elle coïncide la plupart du temps avec l'information lexico-syntaxique. Néanmoins, son utilisation comporte un facteur de risque puisque seules quelques situations syntaxiques tolèrent un accord conceptuel non syntaxique en anglais ; dans tous les autres cas, un désaccord entre l'indice conceptuel et l'indice lexico-syntaxique ne peut donc que générer une erreur. En ce qui concerne l'information morphophonologique sur le nom, elle est relativement fiable puisque la plupart des noms pluriels se marquent de la même façon, i.e. par le *-s* final. Toutefois, une forte ambiguïté réside dans la représentation lexicale de cette marque : tandis qu'elle indique de manière fiable le pluriel des noms, elle constitue également un indice extrêmement fiable de la singularité du verbe. Autrement dit, la fiabilité du *-s* final est restreinte à la catégorie grammaticale du mot qui le porte. Son utilisation comporte par conséquent un facteur de risque considérable. Par ailleurs, la disponibilité de l'indice morphophonologique est limitée puisque la langue comporte des noms dont le pluriel est irrégulier (ex. *child*-S/*children*-P, enfant/enfants) et des noms portant une pseudo marque de pluriel. En résumé, l'indice lexico-syntaxique dans la langue anglaise est donc non seulement un indice d'une grande validité, il ne possède en plus pas de réel concurrent puisque la validité des indices conceptuel et morphophonologique est nettement inférieure.

Deuxièmement, l'utilisation de l'indice lexico-syntaxique en anglais se fait à un *moindre coût*. Cette information est directement et obligatoirement spécifiée lors de l'accès lexical, et cela sur un sujet toujours lexicalisé et qui occupe toujours la même position par rapport au verbe. De plus, comme le sujet précède toujours le verbe, sa marque grammaticale est disponible au moment où celle du verbe doit être spécifiée. Par conséquent, on voit mal pourquoi les locuteurs anglophones mobiliseraient des ressources supplémentaires pour prendre en compte des indices conceptuel ou morphophonologique dont l'utilisation n'entraînerait pas de véritable bénéfice étant donné que l'indice lexico-syntaxique est à la fois nécessaire, suffisant, et utilisable à un moindre coût.

Contrairement à l'anglais, des langues comme le français, l'italien, l'espagnol ou, dans une moindre mesure, le néerlandais, ont développé des systèmes flexionnels fort complexes. Outre l'accord en nombre, ces langues doivent

également traiter l'accord en genre, et les accords en nombre et en genre ne concernent pas seulement le sujet et le verbe mais une série d'autres segments de la phrase comme les déterminants, les adjectifs, ou les participes. La quantité de marques d'accord à gérer représente une double difficulté : la première, liée à la complexité de la tâche, la seconde, liée au risque de perte d'information. Il est raisonnable d'imaginer que les locuteurs de ces langues, pour pallier à ces deux difficultés, gardent à leur disposition la possibilité de recourir à d'autres sources d'information que l'information grammaticale.

A nouveau, une série de facteurs influencent la validité et le coût des différents indices dans ces langues. Par exemple, l'effet de distributivité semble plus important en espagnol et en italien qu'en français ou en néerlandais. Dans ces deux dernières, l'information lexico-syntaxique est toujours disponible puisque la source d'un accord est systématiquement lexicalisée. Cela n'est pas le cas de l'italien ou de l'espagnol dans lesquelles le sujet de la phrase n'est pas obligatoire (langues dites *pro-drop*). Pour accorder le verbe ou n'importe quel autre segment s'accordant avec le sujet, les locuteurs de ces langues doivent donc nécessairement recourir à l'information référentielle pour connaître les spécifications notionnelles de la source, ce qui peut contribuer à augmenter l'effet conceptuel. Par ailleurs, l'effet de distributivité rapporté en néerlandais est plus net que celui rapporté en français. Tandis que l'ordre de mots est relativement flexible en néerlandais, de même qu'en italien et en espagnol, ce n'est pas le cas du français qui, comme l'anglais, impose souvent l'ordre dit canonique (Sujet-Verbe-Objet). Autrement dit, les locuteurs francophones, comme les anglophones, disposent toujours de l'information lexico-syntaxique du sujet au moment de produire le verbe, ce qui n'est pas le cas des trois autres langues. Les facteurs de *pro-drop* et de flexibilité structurale peuvent donc, lorsqu'on les combine, rendre compte de différences entre les quatre langues dans l'effet de distributivité en ce qu'ils influencent la validité des indices lexico-syntaxique et conceptuel.

L'impact différentiel des variables morphophonologiques étudiées a été discuté en détails dans le Chapitre 5. Il peut lui aussi se résumer à des différences entre les langues quant à la validité (fiabilité et disponibilité) et au coût (perceptibilité et charge mnésique) des indices morphophonologiques par rapport à l'indice lexico-syntaxique.

Outre des différences entre les langues quant à la nature des erreurs produites, des différences sont également apparues dans la quantité de ces erreurs. Il va de soi que des différences méthodologiques peuvent partiellement expliquer cette variabilité quantitative. Toutefois, les recherches que nous avons menées parallèlement en français et en italien tentaient de limiter ces différences méthodologiques. Les locuteurs francophones présentaient malgré tout un plus grand taux d'erreurs d'accord en genre que les italophones, et cela de façon systématique dans l'ensemble des recherches. Dans le Chapitre 6, nous rapportons

également que, mis dans les mêmes conditions expérimentales, les participants francophones produisent plus d'erreurs d'accord en nombre que les anglophones. Une première explication plausible de cette différence est la grande diversité des flexions de nombre et de genre en français, contrairement à l'italien ou à l'anglais. Par ailleurs, cette diversité se retrouve également d'un segment à l'autre en français (le pluriel nominal ne se marque pas de la même manière que le pluriel verbal), tandis que l'italien et l'anglais présentent une forte homogénéité des marques sur les différents segments. Le choix étant plus limité en italien et en anglais, la probabilité que ces locuteurs produisent une erreur est moindre. Par exemple, en italien, le suffixe de genre de l'adjectif au singulier peut être *-o*, *-a* ou *-e* ; la probabilité d'erreur lors d'une sélection au hasard est donc de 2/3. Par contre, en français, des dizaines de flexions opposent le masculin et le féminin (voir Chapitre 5), ce qui rend la probabilité d'erreur nettement supérieure, tandis que la probabilité de produire au hasard une réponse correcte est minime. Cette caractéristique du français le rend d'ailleurs particulièrement difficile à apprendre par des locuteurs étrangers.

L'utilisation des indices conceptuel et morphophonologique présents dans la phrase pour construire l'accord d'une cible n'apparaît pas être une question de tout ou rien, i.e. d'isolation stricte ou d'interaction systématique de la syntaxe avec les autres composantes. Le système syntaxique semble au contraire pouvoir moduler le recours à ces différentes informations en fonction de la validité et du coût d'utilisation de ces indices. Dans toutes les langues examinées, l'information lexico-syntaxique est privilégiée lors du traitement syntaxique parce que la plus valide et la moins coûteuse. Toutefois, aucune des langues ne se montre totalement aveugle aux autres informations, même si leur utilisation comme indices pouvant guider une opération comme l'accord est parfois fort limitée. Nous avons dégagé quatre critères propres à chaque langue déterminant le recours à ces autres informations : leur fiabilité, leur disponibilité, leur perceptibilité et la charge attentionnelle et mnésique que leur utilisation mobilise. Des différentes observations interlinguistiques, nous suggérons que le traitement syntaxique possède une propriété de *perméabilité potentielle*, plutôt qu'*actuelle*. Il est potentiellement ouvert sur les informations conceptuelles et morphophonologiques, mais il ne les utilise que lorsque les bénéfices d'une telle perméabilité l'emportent sur les coûts.

Conclusion générale

Tout modèle, qu'il soit modulaire ou interactif, doit rendre compte à la fois de la rapidité et de la précision de la production d'une phrase. Le modèle séquentiel modulaire conçoit la production syntaxique comme une composante autonome, dont le fonctionnement repose sur un minimum de contraintes strictement nécessaires. La restriction du flux d'information tel que le suggère la perspective modulaire limite les risques d'interférence entre les différents niveaux, tandis que la prise en compte d'un minimum d'informations à chaque niveau assure une rapidité maximale au système. Une limitation importante pèse toutefois sur le fonctionnement d'un tel système syntaxique modulaire, liée aux risques de perte d'information. Si, au cours d'un transfert d'un niveau à l'autre, une information se perd, le traitement à ce niveau est sérieusement compromis étant donné la contrainte de séquentialité stricte empêchant tout retour en arrière ou tout saut en avant.

Les différentes observations rapportées dans notre thèse plaident plutôt en faveur d'une perméabilité *potentielle* du système syntaxique aux informations conceptuelles et morphophonologiques. Cette perméabilité est assurée par un fonctionnement interactif de la production, selon lequel chaque portion de phrase peut être traitée simultanément par les différents niveaux linguistiques. Ce fonctionnement en parallèle des différentes composantes de la production permet à chacune d'entre elles de puiser dans la réserve d'information des autres, pour autant que les bénéfices l'emportent sur les coûts. Une telle ouverture du traitement syntaxique sur les indices de nature conceptuelle et morphophonologique limite les risques d'erreurs liées à la perte de l'information lexico-syntaxique.

L'avantage de cette perméabilité repose néanmoins sur la nécessaire concordance entre l'information syntaxique et les informations conceptuelle et morphophonologique. En cas de discordance, le recours à l'information conceptuelle ou morphophonologique peut mener à l'erreur, comme il a été montré. On remarque toutefois que la fréquence de ces situations de discordance entre les indices dans la langue est extrêmement faible. Par ailleurs, l'exploitation par le système de production de la *redondance au niveau psycholinguistique*, i.e. la prise en considération d'informations de natures différentes, est fortement liée à la présence de *redondance au niveau linguistique*, i.e. la spécification de plusieurs segments dans la phrase. Les langues qui présentent une plus grande perméabilité dans le traitement syntaxique sont également les langues qui contiennent le plus grand nombre de segments à accorder dans la phrase. La redondance au niveau linguistique constitue un facteur de cohésion syntagmatique ; les marques assurent

la jonction des syntagmes. Mais elle est également nécessaire à la conservation d'une information que les 'bruits' ou l'inertie du système linguistique peuvent masquer. La redondance au niveau psycholinguistique joue le même rôle en ce sens qu'elle assure au système la possibilité de recourir à une information alternative en cas de perte de l'information lexico-syntaxique, ou dans les cas où certaines caractéristiques syntaxiques rendent le traitement délicat. Comme la redondance se fait au prix d'un certain coût, son introduction dans le système de production est limitée aux situations où elle procure un réel bénéfice pour assurer une formulation syntaxique correcte.

Le cadre théorique sur lequel nous avons basé nos interprétations, en particulier celles concernant la variabilité intra- et inter-langues, s'articule autour de la notion de validité de l'indice d'accord. Cette notion a été avancée, à l'origine, au sein du modèle de compétition et reprise par les théories probabilistes basées sur la satisfaction des contraintes, pour rendre compte de l'influence de facteurs sémantiques et pragmatiques dans le traitement syntaxique en compréhension de phrases. Quelle est la pertinence de ce cadre théorique pour l'étude de la production de la phrase ? Cette question repose sur la façon dont on conçoit les relations entre ces deux systèmes, et plus particulièrement entre le système responsable de l'encodage syntaxique (en production) et celui responsable du décodage syntaxique (en compréhension). Nous partageons le point de vue que la production et la compréhension opèrent sur les mêmes *représentations* linguistiques, mais que les *processus* effectués par ces deux systèmes sont différents. En production, ils consistent à partir d'une structure de sens pour en dégager la séquence sonore correspondante, tandis que ce déroulement se fait en sens inverse en compréhension. Toutefois, dans les deux cas, le système est responsable d'établir les relations entre les différents niveaux linguistiques. Ainsi, la redondance facilite la compréhension en ce qu'elle offre au compreneur un maximum d'informations sur lesquelles il peut se baser pour construire l'interprétation de la phrase. Elle assure par ailleurs une certaine flexibilité au système de compréhension puisqu'elle permet au compreneur de laisser tomber un indice sans que cela n'aie de conséquence sur l'interprétation de la phrase. La redondance représente par contre un certain coût pour le locuteur. Ce dernier doit s'assurer que l'ensemble des informations redondantes soient non seulement produites, mais aussi qu'elles soient concordantes, sous peine de construire une phrase incorrecte. Cet aspect majeur des différences entre production et compréhension est souligné par Bates & Devescovi :

Comprehension is governed by uncertainty, a condition in which vertical mappings are paramount whereas horizontal constraints on coherence and well-formedness are relaxed. Production is governed by commitment, a condition in which the goodness-of-fit between form and meaning (i.e. vertical constraints) may be relaxed while horizontal constraints (i.e. the drive to produce coherent and well-formed speech) are correspondingly increased. (Bates & Devescovi, 1989, p. 229)

Autrement dit, les indices d'accord constituent des contraintes indispensables à respecter pour assurer une production correcte, tandis qu'ils ne sont la plupart du temps pas indispensables pour réaliser une compréhension correcte. Cet aspect se traduit sur le plan développemental par un décalage dans l'apparition des indices d'accord en production et en compréhension. Tandis que les enfants produisent des accords corrects dès l'âge de trois ans, il a été montré qu'ils ne font usage de ces indices pour désambigüiser la compréhension de phrases que vers l'âge de six ou sept ans, et cela tant dans les langues riches en marques d'accord (en français, Kail, 1989 ; en italien, Bates & Devescovi, 1989) que dans les langues plus pauvres (en anglais, Keeney & Wolfe, 1972). Par ailleurs, chez les adultes, l'importance de ces indices en production est étroitement liée à leur importance en compréhension ; dans les langues pauvres comme l'anglais, les compreneurs ne font qu'un faible usage des indices d'accord, tandis que ces indices jouent un rôle important pour le décodage syntaxique dans les langues plus riches comme le français ou l'italien. Parallèlement à cela, nous avons montré qu'en production, la redondance des informations de nombre ou de genre aux différents niveaux linguistiques n'est que faiblement exploitée par les locuteurs anglophones, en comparaison avec les locuteurs francophones ou italophones.

Il apparaît des différentes recherches rapportées que la régulation des flux d'informations conceptuelles, syntaxiques et morphophonologiques lors de l'encodage grammatical ne peut se résumer à une propriété universelle de la production langagière. Cette régulation repose au contraire sur une série de variations intra-linguistiques et interlinguistiques en vertu desquelles la perméabilité remplit une fonction plus ou moins adaptée à une production correcte et rapide de la phrase. Ces variations reflètent des disparités dans la balance entre bénéfiques et coûts liés à l'utilisation de chacun de ces indices, elles reflètent aussi des différences dans la valeur communicative des indices, à savoir le rôle qu'ils jouent dans les processus de compréhension que l'interlocuteur doit effectuer. Les observations suggèrent que tant que ces facteurs de variabilité ne seront pas sérieusement examinés, le débat opposant les modèles modulaires et interactifs restera stérile.

Références bibliographiques

Akmajian, A., & Heny, F. (1975). *An introduction to the principles of transformational syntax*. Cambridge, MA, MIT Press.

Altmann, G. T. M. (1989). Parsing and interpretation: An introduction. *Journal of Language and Cognitive Processes*, 4, 1-19.

Antinucci, F., Duranti, A., & Gebert, L. (1979). Relative clause structure, relative clause perception, and the change from SOV to SVO. *Cognition*, 7(2), 145-176.

Anton-Méndez, I. (1996). Clitics and attraction errors: An experimental study on language production. Unpublished Master Thesis, Dpt. Of Cognitive Science, University of Arizona.

Baayen, R.H., Dijkstra, T., & Schreuder, R. (1997). Singular and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language*, 37, 94-117.

Barbaud, P., Ducharme, C., & Valdois, D. (1982). D'un usage particulier du genre en canadien-français: la féminisation des noms à initiale vocalique. *Journal of Linguistics*, 27(2), 103-133.

Barlow, M. (1993). *A situated theory of agreement*. Garland, Outstanding Dissertations in Linguistics. New York, Garland.

Bates, E., & Devescovi, A. (1989). Crosslinguistic studies of sentence production, in B. MacWhinney, B. & E. Bates (Eds), *The crosslinguistic study of sentence processing*. Cambridge, Cambridge University Press, 225-253.

Bates, E., & MacWhinney, B. (1982). Functionalist approaches to grammar, in E. Wanner and L.R. Gleitman (Eds), *Language acquisition: The state of the art*. Cambridge, Cambridge University Press, 173-218.

Bates, E., & MacWhinney, B. (1989). Functionalism and the Competition Model, in B. MacWhinney, B. & E. Bates (Eds), *The crosslinguistic study of sentence processing*. Cambridge, Cambridge University Press, 3-73.

Berg, T. (1992). Prelexical and postlexical features in language production. *Applied Psycholinguistics*, 13, 199-235.

Bock, J.K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review*, 89, 1-47.

- Bock, J.K. (1986). Meaning, sound, and syntax: lexical priming in sentence production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 12, 575-586.
- Bock, J.K. (1987). An effect of the accessibility of word forms on sentence structures. *Journal of Memory and Language*, 26, 119-137.
- Bock (1991). A sketchbook of production problems. *Journal of Psycholinguistic Research*, 20, 141-160.
- Bock, J.K. (1995). Producing agreement. *Current Directions in Psychological Science*, 4, 56-61.
- Bock, J.K. (1995). Sentence production: From mind to mouth, in J. Miller & P. Eimas (Eds.), *Handbook of Perception and Cognition (Vol. 11): Speech, Language, and Communication*. New York, Academic Press, 181-216.
- Bock, J. K., & Cutting, J.C. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language*, 31, 99-127.
- Bock, J.K., & Eberhard, K.M. (1993). Meaning, sound and syntax in English number agreement. *Language and Cognitive Processes*, 8(1), 57-99.
- Bock, J.K., Eberhard, K.M., & Cutting, J.C. (1992). Controlling Number agreement on verbs and anaphors. Paper presented at the Meeting of the Psychonomic Society, St. Louis, Missouri.
- Bock, J.K., & Kroch, A.S. (1989). The isolability of syntactic processing, in G.N. Carlson and M.K. Tanenhaus (Ed.), *Linguistic Structure in Language Processing*. Dordrecht, Kluwer, 157-196.
- Bock, J.K., & Levelt, W. (1994). Language production. Grammatical encoding, in M.A. Gernsbacher (Ed.), *Handbook of Psycholinguistics*. Academic Press.
- Bock, J.K., & Miller, C.A. (1991). Broken agreement. *Cognitive Psychology*, 23, 45-93.
- Bock, J.K., Nicol, J.L., & Cutting, J.C. (1999). The ties that bind: creating number agreement in speech. *Journal of Memory and Language*, 40, 330-346.
- Bock, J.K., & Warren, R.K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition*, 21, 47-67.
- Bowerman, M. (1973). Structural relationships in children's utterances: syntactic or semantic? in T.E. Moore (Ed.), *Cognitive development and the acquisition of language*. New York, Academic Press, 197-213.
- Bresnan, J.(1982). *The mental representation of grammatical relations*. Cambridge MA, MIT Press.
- Briet, H. (1996). *Savoir accorder le verbe*. Belgique : Duculot.

- Butterworth, B. (1983). Lexical representations, *in* B. Butterworth (Ed.), *Language production (Vol. 2)*. London, Academic Press, 257-294.
- Cacciari, C. Carreiras, M., & Barbolini-Cionini, C. (1997). Anaphora resolution in Italian. *Journal of Memory and Language*, 37, 517-532.
- Carreiras, M., Garnham, A., Oakhill, J., & Cain, K. (1996). The use of stereotypical information in constructing a mental model: Evidence from English and Spanish. *Quarterly Journal of Experimental Psychology*, 49A, 639-663.
- Charvillat, A. (manuscrit non publié). Analyse structurale en temps réel et compétition d'indices: Etude du traitement des relatives ambiguës en français.
- Chomsky, N. (1957). *Syntactic structures*. The Hague, Mouton.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA, MIT Press.
- Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht: Foris.
- Clark, E.V. (1973). What's in a word? On the child's acquisition of semantics in his first language, *in* T.E. Moore (Ed.), *Cognitive development and the acquisition of language*. New York, Academic Press, 65-110.
- Comrie, B. (1975). Polite plurals and predicate agreement. *Language*, 51, 406-418.
- Content, A., Mousty, P., & Radeau, M. (1990). Brulex: Une base de données lexicales informatisée pour le français écrit et parlé. *L'Année Psychologique*, 90, 551-566.
- Cooper, W.E., Paccia, J.M., & Lapointe, S.G. (1978). Hierarchical coding in speech timing. *Cognitive Psychology*, 10, 154-177.
- Corbett, G.G. (1983). *Hierarchies, targets and controllers: Agreement patterns in Slavic Languages*. University Park, PA, Pennsylvania State University Press.
- Corbett, G. G. (1991). *Gender*. Cambridge, Cambridge University Press.
- Costermans, J. (1980). *Psychologie du langage*. Bruxelles, Mardaga.
- Costermans, J., & Hupet, M. (1977). The other side of Johnson-Laird's interpretation of the passive voice. *British Journal of Psychology*, 68, 107-111.
- Cuetos, F., & Mitchell, D. (1988). Cross-linguistic differences in parsing: Restrictions on the use of the late closure strategy in Spanish. *Cognition*, 30, 73-105.
- Cutler, A. (1983). Lexical complexity and sentence processing, *in* G.B. Flores d'Arcais & J.R. Jarvella (Eds.), *The process of Language Understanding*. Chichester: Wiley, 43-79.

Deese, J. (1984). *Thought into Speech: The Psychology of Language*. Englewood Cliffs, NJ: Prentice-Hall.

Dell, G.S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93(3), 283-321.

Dell, G.S., Burger, L.K., & Svec, W.R. (1997). Language production and serial order : A functional analysis and a model. *Psychological Review*, 104(1), 123-147.

Dell, G.S., & O'Seaghdha, P.G. (1992). Stages of lexical access in language production. *Cognition*, 42, 287-314.

Dell, G.S., & Reich, P.A. (1981). Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior*, 20, 611-629.

De Smedt, K. (1990). IPF : An incremental parallel formulator, in R. Dale, C. Mellish and M. Zock (Ed.), *Current research in natural language generation*. London, Academic Press, 167-192.

De Smedt, K.J. (1994). Parallelism in incremental sentence generation, in G. Adriaens and U. Hahn (Eds.), *Parallel models of natural language computation*. Norwood, NJ: Ablex.

Dominguez, A., Cuetos, F., & Segui, J. (1999). The processing of grammatical gender and number in Spanish. *Journal of Psycholinguistic Research*, 28(5), 485-498.

Dubois, J. (1965). *Grammaire structurale du français*. Paris, Larousse.

Eberhard, K.M. (1993a). The specification of grammatical number in English. Unpublished doctoral dissertation, Michigan State University.

Eberhard, K.M. (1993b). The specification of grammatical number in English. Poster presented at the Sixth Annual Meeting of the CUNY Sentence Processing Conference.

Eberhard, K.M. (1997). The marked effect of number on subject-verb agreement. *Journal of Memory and Language*, 36, 147-164.

Eberhard, K.M. (1999). The accessibility of conceptual number to the process of subject-verb agreement in English. *Journal of Memory and Language*, 41, 560-578.

Fayol, M. & Got, C. (1991). Automatisme et contrôle dans la production écrite. *L'Année Psychologique*, 91, 187-205.

Fayol, M., Hupet, M., & Largy, P. (1999). The acquisition of subject-verb agreement in written French: From novices to experts. *Reading and Writing: An Interdisciplinary Journal*, 11, 153-174.

Fayol, M., Largy, P., & Lemaire, P. (1994). When cognitive overload enhances subject-verb agreement errors : A study in French written language. *The Quarterly Journal of Experimental Psychology*, 47, 437-464.

Fodor, J.A. (1983). *The Modularity of Mind*. Cambridge, MA, MIT Press.

Ford, M., & Holmes, V.M. (1978). Planning units and syntax in sentence production. *Cognition*, 6, 35-53.

Franck, J. (1998). Object attraction in the production of subject-verb agreement in French. 4th Conference on Architectures and Mechanisms for Language Processing, Freiburg, Germany.

Franck, J., & Butterworth, B. (1997). Syntactic distance as a key factor in the construction of subject-verb agreement in French. Tenth Annual CUNY Conference on Human Sentence Processing, Santa Monica, USA.

Franck, J., Collina, S., & Vigliocco, G. (1999). Gender agreement in language production II: When morpho-phonology hits syntax, in A. Vandierendonck, M. Brysbaert, K. Van der Gooten (Eds), Eleventh Conference of the European Society for Cognitive Psychology, Gent, ESCoP/Academia Press, 245.

Franck, J., & Hupet, M. (en révision). La régulation du flux d'information entre les niveaux de production: L'accord grammatical comme outil d'investigation. *L'Année Psychologique*.

Franck, J., & Vigliocco, G. (1998). On the conflict between conceptual and grammatical gender. Eleventh Annual CUNY Conference on Human Sentence Processing, New Brunswick, USA.

Franck, J., Vigliocco, G., & Nicol, J. (submitted). Attraction in Sentence Production: The Role of Syntactic Structure and Processing Complexity. *Language and Cognitive Processes*.

Frazier, L. (1990). Exploring the architecture of the language system, in G.T. Altman (Ed.), *Cognitive models of speech processing : Psycholinguistics and computational perspectives*. Cambridge, MA, MIT Press, 409-433.

Fromkin, V.A. (1971). The non-anomalous nature of anomalous utterances. *Language*, 47, 27-52.

Garnham, A, Oakhill, J., & Ehrlich, M.F. (1995). Representations and processes in the interpretation of pronouns: New evidence from Spanish and French. *Journal of Memory and Language*, 34, 41-62.

Garrett, M.F. (1976). Syntactic processes in sentence production, in R. Wales and E. Walker (Ed.), *New Approaches to Language Mechanisms*. Amsterdam, North-Holland, 133-177.

Garrett, M.F. (1980). Levels of processing in sentence production, in B. Butterworth (Ed.), *Language production, Volume 1: Speech and Talk*. London, Academic Press.

Garrett, M.F. (1982). Production of Speech: Observations from normal and pathological language use, in A.W. Ellis (Ed.), *Normality and pathology in cognitive functions*. London, Academic Press.

Garrett, M.F. (1988). Processes in language production, in F.J. Nieuwmeyer (Ed.), *Linguistics : The Cambridge survey, Vol III, Biological and psychological aspects of language*. Cambridge, Harvard University Press, 69-96.

Garrett, M.F. (1990). Sentence Processing, in D.N. Osherson and H. Lasnik (Eds.), *Language: An invitation to cognitive science, Volume 1*. Cambridge MA, MIT Press.

Garrett, M.F. (1992). Disorders of lexical selection, *Cognition*, 42, 143-180.

Garrett, M.F. (1993). Errors and their relevance for models of language production, in G. Blanken, J. Dittman, H. Grim, J. Marshall & C. Wallech (Eds.), *Linguistic disorders and pathologies*. Berlin, de Gruyter.

Garrett, M.F. (in press). Remarks on the architecture of language processing systems, in Grodzinsky, Shapiro and Swinney (Eds), *Language and Brain*.

Gazdar, G., Klein, E., Pullum, G.K., & Sag, I.A. (1985). *Generalized phrase structure grammar*. Oxford, Blackwell, and Cambridge, MA, Harvard University Press.

Gernsbacher, M.A. (1991). Comprehending conceptual anaphors. *Language and Cognitive Processes*, 6, 81-105.

Gilboy, E., Sopena, J.-M., Clifton, C., & Frazier, L. (1995). Argument structure and association preferences in Spanish and English complex NPs. *Cognition*, 54, 131-167.

Greenberg, J. (1966). *Language universals*. The Hague: Mouton.

Grévisse, M. (1993). *Le bon usage*. Paris-Louvain-la-Neuve, Duculot.

Hanse, J. (1987). *Nouveau dictionnaire des difficultés du français moderne*. Paris-Louvain-la-Neuve, Duculot.

Harley, T.A. (1993). Connectionist approaches to language disorders. *Aphasiology*, 7, 221-249.

Harris, M. (1990). French, in M. Harris and N. Vincent (Eds.), *The Romance languages*. London, Croom Helm.

Harris, M., & Vincent, N. (1990). *The Romance languages*. London: Croom Helm.

Hartsuiker, R.J., Anton-Méndez, I., & van Zee, M. (submitted). Object attraction in subject-verb agreement construction.

Hartsuiker, R.J., Schriefers, H.J., Bock, K., & Kikstra, G.M. (in prep.). Morphophonological influences on the construction of subject-verb agreement.

Haviland, S.E., & Clark, H.H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, *13*, 512-521.

Hupet, M., Fayol, M., & Schelstraete, M.-A. (1998). Effects of semantic variables on the subject-verb agreement processes in writing. *British Journal of Psychology*, *89*, 59-75.

Hupet, M., Schelstraete, M.-A., Demaeght, N., & Fayol, M. (1996). Les erreurs d'accord sujet-verbe en production écrite. *L'Année Psychologique*, *96*, 587-610.

Jespersen, O. (1924). *The philosophy of grammar*. London, Allen & Unwin.

Kail, M. (1989). Cue validity, cue cost, and processing types in sentence comprehension in French and Spanish, in B. MacWhinney, B. & E. Bates (Eds), *The crosslinguistic study of sentence processing*. Cambridge, Cambridge University Press, 77-117.

Keenan, E.L. (1976). Towards a universal definition of 'subject of', in C. Li (Ed.), *Subject and topic*. New York, Academic Press, 303-333.

Keeney, T.J., & Wolfe, J. (1972). The acquisition of agreement in English. *Journal of Verbal Learning and Verbal Behavior*, *11*, 698-705.

Kempen, G. (1997). Grammatical performance in human sentence production and comprehension. Manuscript in preparation.

Kempen, G., & Hoenkamp, E. (1987). Incremental procedural grammar for sentence formulation. *Cognitive Science*, *11*, 201-258.

Kempen, G., & Vosse, T. (1989). Incremental syntactic tree formation in human sentence processing : A cognitive architecture based on activation decay and simulated annealing. *Connection Science*, *1*, 275-292.

Lackner, J., & Garrett, M.F. (1972). Resolving ambiguity: Effects of biasing context in the unattended ear. *Cognition*, *1(4)*, 359-372.

Lackner, J., & Tuller, B. (1976). The influence of syntactic segmentation on perceived stress. *Cognition*, *4(4)*, 303-307.

LaHeij, W., Mak, P., Sander, J., & Willeboordse, E. (1998). The gender congruency effect in picture-word tasks. *Psychological Research*, *61*, 209-219.

Lapointe, S.G. (1986). Markedness, the organisation of linguistic information in speech production, and language acquisition, in F.R. Eckman, E.A. Moravcsik and J.R. Wirth (Eds.), *Markedness*. New York, Plenum Press, 219-239.

Levelt, W.J.M. (1983). Monitoring and self-repair in speech. *Cognition*, 14, 41-104.

Levelt, W.J.M. (1989). *Speaking : From intention to articulation*. Mass., MIT Press.

Levelt, W.J.M., & Maasen, R. (1981). Lexical search and order of mention in sentence production, in W. Klein and W.J.M. Levelt (Eds.), *Crossing the boundaries in linguistics*. Dordrecht, Reidel, 221-252.

Levelt, W.J.M., Roelofs, A., & Meyer, A.S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1-75.

MacDonald, M.C., Pearlmutter, N.J., & Seidenberg, M.S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101(4), 676-703.

MacKay (1987). *The organisation of perception and action: A theory for language and other cognitive skills*. New York: Springer-Verlag.

Marcantonio, A, & Pretto, A.M. (1988). Il nome, in L. Renzi (Ed.) *Grande grammatica di consultazione della lingua Italiana*. Vol. 1 Bologna, Italy, Il Mulino.

Marslen-Wilson, W. (1973). Speech shadowing and speech perception. PhD Thesis, MIT.

Marslen-Wilson, W., & Tyler, L.K. (1987). Against modularity, in J. Garfield (Ed.), *Modularity in knowledge representation and natural-language understanding*. Mass., MIT. Press, 37-62.

McClelland, J.L. (1987). The case for interactionism in language processing, in M. Coltheart Max (Ed.), *Attention and performance 12: The psychology of reading*. Hove, England UK: Lawrence Erlbaum, 3-36.

McClelland, J.L., & Rumelhart, D.E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, 88, 375-407.

McDonald, M.C., Pearlmutter, N.J., & Seidenberg, M.S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101, 676-703.

Meyer, A., & Bock, J.K. (1998). Representations and processes in the production of pronouns: Some perspectives from Dutch. *Journal of Memory and Language*, 41, 281-301.

Miller, G.A. (1962). Some psychological studies of grammar. *American Psychologist*, 17, 748-762.

Morton, J. (1969). The interaction of information in word recognition. *Psychological Review*, 76, 165-178.

- Motley, M.T., & Baars, B.J. (1979). Effects of cognitive set upon laboratory induced verbal (Freudian) slips. *Journal of Speech and Hearing Research*, 22(3), 421-432.
- Motley, M.T., Baars, B.J., & Camden, C.T. (1981). Syntactic criteria in prearticulatory editing: Evidence from laboratory-induced slips of the tongue. *Journal of Psycholinguistic Research*, 10(5), 503-522.
- Nicol, J. (1995). Effects of clausal structure on subject-verb agreement errors. *Journal of Psycholinguistic Research*, 24(6), 507-516.
- Nooteboom, S.G. (1980). Speaking and unspeaking: Detection and correction in phonological and lexical errors in spontaneous speech, in V.A. Fromkin (Ed.), *Errors in linguistic performance*. New York: Academic Press, 87-95.
- Perfetti, C.A. (1990). The cooperative language processors: Semantic influences in an autonomous syntax, in D.A. Balota, G.B. Flores d'Arcais & K. Rayner (Eds), *Comprehension processes in reading*. Hillsdale, Lawrence Erlbaum Associates, 205-230.
- Pillon, A. (1993). *La mémoire des mots*. Liège, Mardaga.
- Pinker, S. (1999). *The language instinct*. New York, NY, USA: William Morrow and Co.
- Pollard, C., & Sag, I.A. (1994). *Head-driven phrase structure grammar*. Chicago, The University of Chicago Press.
- Quirk, R., & Greenbaum, S. (1973). *A university grammar of English*. London, Longman.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1972). *A grammar of contemporary English*. London, Longman.
- Sapir E. (1921). *Language*. New York, Harcourt, Brace & World.
- Schelstraete, M.A. (1993). Le traitement syntaxique en compréhension de phrases. *L'Année Psychologique*, 93, 543-582.
- Schelstraete, M.-A., & Degand, L. (1998). Assignment of grammatical functions in French relative clauses. *Language Sciences*, 20(2), 163-188.
- Shieber, S.M. (1986). *An introduction to unification-based approaches to grammar*. CSLI Lecture Notes 4, Stanford:CSLI.
- Schriefers, H. (1993). Syntactic processes in the production of noun phrases. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(4), 841-850.
- Schriefers, H., & Jescheniak, J. (1999). Representation and processing of grammatical gender in language production: A review. *Journal of Psycholinguistic Research*, 28(6), 575-600.

Schriefers, H., Meyer, A.S., & Levelt, W.J. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language*, 29(1), 86-102.

Schriefers, H., & Teruel, E. (1999). The production of noun phrases: A cross-linguistic comparison of French and German. Proceedings of the 21st Annual Conference of the Cognitive Science Society.

Sternberg, S., Knoll, R.L., Monsell, S., & Wright, C.E. (1988). Motor programs and hierarchical organization in the control of rapid speech. *Phonetica*, 45, 175-197.

Stemberger, J.P. (1982). Syntactic errors in speech. *Journal of Psycholinguistic Research*, 11, 313-345.

Stemberger, J.P. (1985a). An interactive activation model of language production, in A.W. Ellis (Ed.), *Progress in the Psychology of Language*. London, Lawrence Erlbaum, 143-186.

Stemberger, J.P. (1985b). *The lexicon in a model of speech production*. New York, Garland.

Stemberger, J.P., & MacWhinney, B. (1986). Form oriented inflectional errors in language processing. *Cognitive Psychology*, 18, 329-354.

Stevenson, S. (1994). *A competitive attachment model for resolving syntactic ambiguities in natural language parsing*. (Tech. Report RuCCS TR-18). New York, Rutgers University, Center for Cognitive Science.

Strang, B.M.H. (1966). Some features of S-V concord in present day English in I. Cellini and G. Melchiori (Eds.), *English studies today: Forth series*. Rome, Edizioni di Storia e Letteratura, 73-87.

Taft, M. (1979). Recognition of affixed words and the word frequency effect. *Memory and Cognition*, 7, 263-272.

Tanenhaus, M.K., Spivey-Knowlton, M.J., Eberhard, K.M., & Sedivy, J.C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268, 1632-1634.

Tasmowski-De Rick, L., & Verluyten, P. (1982). Linguistic control of pronouns. *Journal of Semantics*, 1, 323-346.

Tiesma, P.M. (1982). Local and general markedness. *Language*, 58, 832-848.

Trueswell, J.C., Tanenhaus, M.K., & Garnsey, S.M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, 33, 285-318.

Tucker, G.R., Lambert, W.E., & Rigault, A.A. (1977). *The French speaker's skill with grammatical gender. An example of rule-governed behavior*. The Hague-Paris, Mouton.

Van Berkum, J.-A. (1997). Syntactic processes in speech production: The retrieval of grammatical gender. *Cognition*, 64(2), 115-152.

Van der Molen, H., & Morton, J. (1979). Remembering plurals: Unit of coding and form of coding during serial recall. *Cognition*, 7(1), 35-47.

Van Turenhout, M., Hagoort, P., & Brown, C.M. (1997). Electrophysiological evidence on the time course of semantic and phonological processes in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(4), 787-806.

Van Turenhout, M., Hagoort, P., & Brown, C.M. (1999). The time course of grammatical and phonological processing during speaking: Evidence from event-related brain potentials. *Journal of Psycholinguistics Research*, 28(6), 649-676.

Vigliocco, G., Anton-Méndez, I., Franck, J. & Collina, S. (1999). Sound and syntax in sentence production: The effect of phonological predictability on syntactic encoding. 5th Conference on Architectures and Mechanisms for Language Processing, Edinburgh, UK.

Vigliocco, G., Antonini, T., & Garrett, M.F. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological Science*, 8, 314-317.

Vigliocco, G., Butterworth, B., & Garrett, M.F. (1996a). Subject-verb agreement in Spanish and English: The role of conceptual factors. *Cognition*, 51, 261-298.

Vigliocco, G., Butterworth, B., & Semenza, C. (1995). Constructing subject-verb agreement in Speech: The role of semantic and morphological factors. *Journal of Memory and Language*, 34, 186-215.

Vigliocco, G., & Fava, E. (1993). Diglossia, forme fonologiche e sistemi sintattici: Alcune considerazioni di non-isomorfismo, in E. Cresti e M. Moneglia (Ed.), *Ricerche recenti sull'acquisizione dell'italiano in bambini normali e patologici*. Bulzoni, Roma, 291-329.

Vigliocco, G., & Franck, J. (1999). When sex and syntax go hand in hand: Gender agreement in language production. *Journal of Memory and Language*, 40, 455-478.

Vigliocco, G., & Franck, J. (submitted). When sex hits syntax: Syntactic accuracy in language production. *Journal of Memory and Language*.

Vigliocco, G., Franck, J., & Collina S. (1999). Gender agreement in language production I: When sex hits syntax, in A. Vandierendonck, M. Brysbaert, K. Van

der Goten (Eds), Eleventh Conference of the European Society for Cognitive Psychology, ESCoP/Academia Press, Gent, 231.

Vigliocco, G., Hartsuiker, R., & Franck, J. (1999). The interplay of meaning, sound and syntax in sentence production. Manuscript in preparation.

Vigliocco, G., Hartsuiker, R.J., Jarema, G., & Kolk, H.H.J. (1996b). How many labels on the bottles? Notional concord in Dutch and French. *Journal of Language and Cognitive Processes*, 11, 407-421.

Vigliocco, G., & Nicol, J. (1998). Separating hierarchical relations and word order in language production. Is proximity concord syntactic or linear? *Cognition*, 68(1), 13-29.

Vigliocco, G., Verès, C., & Franck, J. (1997). When there is only one baby on the blankets: Conceptual number effects in subject-verb agreement. Tenth Annual CUNY Conference on Human Sentence Processing, Santa Monica, USA.

Vigliocco, G., Vinson, D.P., Martin, R.C., & Garrett, M.F. (1999). Is "count" and "mass" information available when the noun is not? An investigation of tip of the tongue states and anomia. *Journal of Memory and Language*, 40(4), 534-558.

Vigliocco, G. & Zilli, T. (in press). Syntactic accuracy in sentence production: Gender disagreement in Italian language impaired and unimpaired speakers. *Journal of Psycholinguistic Research*.

Vincent, N. (1990). Italian, in M. Harris and N. Vincent (Eds.), *The Romance Languages*. London, Croom Helm, 279-313.

Zagar, D., Pynte, J., & Rativeau, S. (1994). Evidence for early-closure attachment on first-pass reading times in French. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 50A(2), 421-438.

Zandvoort, R. W. (1961). *Varia syntactica. Language and Society: Essays presented to Arthur M. Jensen on his Seventieth Birthday*. Copenhagen, Det Berlingske Bogtrykkeri.

Annexes

Annexe 1

Liste des items expérimentaux utilisés dans l'Expérience 1 du Chapitre 2.

La bambina nel lettino	SIMPATICA/ SIMPATICO
La vicina nel giardino	SUDATA/ SUDATO
La nonna sul divano	RILASSATA/ RILASSATO
La cuoca dell'albergo	BRAVA/ BRAVO
L'impiegata nell'ufficio	BIONDA/ BIONDO
La contadina nell'orto	ANZIANA/ ANZIANO
La zingara nel parcheggio	FURIOSA/ FURIOSO
La bidella nell'atrio	ARRABBIATA/ ARRABBIATO
La piscina del campeggio	LIBERA/ LIBERO
La domanda al convegno	COMPLICATA/ COMPLICATO
La scrivania nello studio	IMPOLVERATA/ IMPOLVERATO
La valigia per il viaggio	STRAPIENA/ STRAPIENO
L'entrata dello stabilimento	AMPIA/ AMPIO
L'anteprima al teatro	ATTESA/ ATTESO
La vetrina del negozio	DECORATA/ DECORATO
La panchina nel parco	AZZURRA/ AZZURRO
Il maestro di scuol	RAFFREDDATO/ RAFFREDDATA
Il mago con la carta	FAMOSO/ FAMOSA
Il proprietario della villa	RICCO/ RICCA
Il fidanzato in gioiellieri	ANSIOSO/ ANSIOSA
Lo sposo in chiesa	NERVOSO/ NERVOSA
Il suocero in banca	ANNOIATO/ ANNOIATA
Il cognato nella sala	STUFO/ STUFA
Il sarto in tintoria	SILENZIOSO/ SILENZIOSA
L'antipasto alla cena	MISTO/ MISTA
L'antipasto della trattoria	SAPORITO/ SAPORITA
Il cesto con la merenda	ORNATO/ ORNATA
Lo sgabuzzino della casa	SERRATO/ SERRATA
L'applauso alla conferenza	STENTATO/ STENTATA
Il materasso nella tenda	MORBIDO/ MORBIDA
Il macchinario dell'officina	VECCHIO/ VECCHIA
Il pagamento alla posta	ANNULLATO/ ANNULLATA
La compagna per il viaggio	ALLEGRO/ ALLEGRA
La scienziate al convegno	PREPARATO/ PREPARATA
La psicologa nello studio	INDAFFARATO/ INDAFFARATA
La ragazza nel parco	CONTENTO/ CONTENTA
L'amica del campeggio	PARTITO/ PARTITA
La bagnina dello stabilimento	GRASSO/ GRASSA
La commessa del negozio	SCORBUTICO/ SCORBUTICA
La ballerina al teatro	ESPERTO/ ESPERTA
La porta dell'albergo	ROSSO/ ROSSA
La macchina nel parcheggio	LAVATO/ LAVATA
La piantina nel giardino	CRESCIUTO/ CRESCIUTA

La melanzana nell'orto
La macchia sul divano
La libreria nell'ufficio
La coperta nel lettino
La lampada nell'atrio
Il bimbo con la merenda
L'inquilino della casa
L'invitato alla cena
L'operaio dell'officina
Il filosofo alla conferenza
Il pensionato alla posta
L'indovino nella tenda
Lo sgattero della trattoria
Il gioco con la carta
Il banco di scuola
L'armadio nella sala
Lo sportello in banca
Il conto in tintoria
L'antifurto della villa
Il braccialetto in gioielleria
Il cero in chiesa

MATURO/ MATURA
OLEOSO/ OLEOSA
NERO/ NERA
COLORATO/ COLORATA
FULMINATO/ FULMINATA
CONTENTA/ CONTENTO
TIMIDA/ TIMIDO
MAGRA/ MAGRO
PIGRA/ PIGRO
ESAUSTA/ ESAUSTO
PERMALOSA/ PERMALOSO
CURIOSA/ CURIOSO
ENERGICA/ ENERGICO
NOIOSA/ NOIOSO
ROVINATA/ ROVINATO
ALTA/ ALTO
OCCUPATA/ OCCUPATO
ECESSIVA/ ECESSIVO
ROTTA/ ROTTO
PREZIOSA/ PREZIOSO
ACCESA/ACCESO

Annexe 2

Liste des items expérimentaux utilisés dans l'Expérience 2 du Chapitre 2.

La soeur de la pharmacienne	DISCRÈTE/DISCRET
La chatte de la voisine	PEUREUSE/PEUREUX
La couleur de la robe	LAIDE/LAID
La housse de la couette	VIEILLE/VIEUX
La réponse de la secrétaire	IDIOTE/IDIOT
La perruque de la dentiste	BLONDE/BLOND
La présentatrice de la soirée	AMUSANTE/AMUSANT
La serveuse de la taverne	ATTIRANTE/ATTIRANT
Le copain du coiffeur	COQUET/COQUETTE
Le chien du concierge	AGRESSIF/AGRESSIVE
Le récit du témoin	INTÉRESSANT/INTÉRESSANTE
Le cadre du vélo	LOURD/LOURDE
Le visage du voleur	EFFRAYANT/EFFRAYANTE
Le discours du sénateur	POMPEUX/POMPEUSE
Le responsable du projet	PRÉTENTIEUX/PRÉTENTIEUSE
Le délégué du groupe	STIMULANT/STIMULANTE
La collaboratrice du patron	AUDACIEUSE/AUDACIEUX
La femme du médecin	SCRUPULEUSE/SCRUPULEUX
La sortie du souterrain	GLISSANTE/GLISSANT
La cage du canari	LÉGÈRE/LÉGER
La cuisine du chef	GRASSE/GRAS
La famille du professeur	ACCUEILLANTE/ACCUEILLANT
La présidente du jury	ABSENTE/ABSENT
La doyenne du village	CHARMANTE/CHARMANT
Le gardien de la prisonnière	DÉPRESSIF/DÉPRESSIVE
Le mari de la boulangère	GÉNÉREUX/GÉNÉREUSE
Le vol de la mésange	ÉLÉGANT/ÉLÉGANTE
Le châssis de la fenêtre	BLANC/BLANCHE
Le chapeau de la bigotte	CURIEUX/CURIEUSE
Le héros de la pièce	ÉTONNANT/ÉTONNANTE
Le réalisateur de la publicité	PROVOQUANT/PROVOQUANTE

Annexe 3

Liste des items expérimentaux utilisés dans l'Expérience 3 du Chapitre 2.

Il gabbiano sul molo	AFFAMATA/AFFAMATO
Il gabbiano sulla chiatta	AFFAMATA/AFFAMATO
Il merlo sul ramo	ALLEGRO/ALLEGRA
Il merlo sulla pianta	ALLEGRO /ALLEGRA
La panna sul gelato	APPETITOSA/APPETITOSO
La panna sulla torta	APPETITOSA/APPETITOSO
Il radicchio al forno	APPETITOSO/APPETITOSA
Il radicchio alla griglia	APPETITOSO/APPETITOSA
Il criceto nel negozio	ASSOPITO/ASSOPITA
Il criceto nella scatola	ASSOPITO/ASSOPITA
Il capanno del villaggio	BASSA/BASSO
Il capanno nella foresta	BASSA/BASSO
La coccinella sul filo	BELLO/BELLA
La coccinella sulla foglia	BELLO/BELLA
Il delfino nella vasca	BIANCA/BIANCO
Il delfino nell'acquario	BIANCA/BIANCO
Il ciottolo nella vasca	BIANCO/BIANCA
Il ciottolo nell'acquario	BIANCO/BIANCA
La talpa nella buca	CIECA/CIECO
La talpa nell'orto	CIECA/CIECO
La tenda nella radura	COLORATO/COLORATA
L'altalena nel giardino	COLORATO/COLORATA
La barca nell'acqua	DANNEGGIATO/DANNEGGIATA
La barca nello stagno	DANNEGGIATO/DANNEGGIATA
Il branzino al forno	DELIZIOSO/DELIZIOSA
Il branzino alla griglia	DELIZIOSO/DELIZIOSA
Il carretto in campagna	FERMO/FERMA
Il carretto nel recinto	FERMO/FERMA
La mosca sul gelato	DISGUSTOSA/DISGUSTOSO
La mosca sulla torta	DISGUSTOSA/DISGUSTOSO
Il quadro nella stanza	FAMOSA/FAMOSO
Il quadro sul muro	FAMOSA/FAMOSO
La zanzara nella cucina	FASTIDIOSO/FASTIDIOSA
La zanzara nell'ufficio	FASTIDIOSO/FASTIDIOSA
La foca con il birillo	FERITA/FERITO
La foca con la palla	FERITA/FERITO
Il fagiano in campagna	FERMO/FERMA
Il fagiano nel recinto	FERMO/FERMA
Il falco sul braccio	FIERA/FIERO
Il falco sulla spalla	FIERA/FIERO
La gara con il birillo	FINITO/FINITA
La gara con la palla	FINITO/FINITA
La rugiada sul filo	FRESCA/FRESCO
La rugiada sulla foglia	FRESCA/FRESCO
La bacinella sopra il cesto	GIALLO/GIALLA
La bacinella sopra la panca	GIALLO/GIALLA
La calzamaglia in cantina	GRIGIA/GRIGIO
La calzamaglia nel cassetto	GRIGIA/GRIGIO
La lucertola in cantina	GRIGIA/GRIGIO
La lucertola nel cassetto	GRIGIA/GRIGIO
La carota col pomodoro	GUSTOSO/GUSTOSA

La carota nell'insalata	GUSTOSO/GUSTOSA
Il pulcino al mercato	IMPAURITA/IMPAURITO
Il pulcino nella vetrina	IMPAURITA/IMPAURITO
La formica nello zucchero	INDAFFARATO/INDAFFARATA
La formica sulla credenza	INDAFFARATO/INDAFFARATA
Il calamaro nella pentola	INTERA/INTERO
Il calamaro sul vassoio	INTERA/INTERO
Il cetriolo nella pentola	INTERA/INTERO
Il cetriolo sul vassoio	INTERA/INTERO
La lucciola nel parco	LUMINOSO/LUMINOSA
La lucciola nella piazza	LUMINOSO/LUMINOSA
L'aquila in cielo	MAESTOSA/MAESTOSO
L'aquila sulla cima	MAESTOSA/MAESTOSO
Il frutto sul ramo	MATURA/MATURO
Il frutto sulla pianta	MATURA/MATURO
La nuvola in cielo	MINACCIOSA/MINACCIOSO
La nuvola sulla cima	MINACCIOSA/MINACCIOSO
Lo scarafaggio nel bagno	MORTO/MORTA
Lo scarafaggio sulla scala	MORTO/MORTA
La gazza sulla tovaglia	NERO/NERA
La gazza sull'albero	NERO/NERA
Il ragno nella stanza	PELOSA/PELOSO
Il ragno sul muro	PELOSA/PELOSO
Il drago del racconto	PICCOLO/PICCOLA
Il drago della favola	PICCOLO/PICCOLA
Il libro con il racconto	PICCOLO/PICCOLA
Il libro con la favola	PICCOLO/PICCOLA
Il graffio sul braccio	PROFONDO/PROFONDA
Il graffio sulla spalla	PROFONDO/PROFONDA
Il ciclamino nel prato	PROFUMATA/PROFUMATO
Il ciclamino nell'erba	PROFUMATA/PROFUMATO
La cannella nello zucchero	PROFUMATA/PROFUMATO
La cannella sulla credenza	PROFUMATA/PROFUMATO
Il sigaro nel negozio	PROFUMATO/PROFUMATA
Il sigaro nella scatola	PROFUMATO/PROFUMATA
La scimmia con il cappello	RIDICOLO/RIDICOLA
La scimmia con la gonna	RIDICOLO/RIDICOLA
Il gambero sulla spiaggia	ROSSA/ROSSO
Il gambero sullo scoglio	ROSSA/ROSSO
La mela sulla tovaglia	ROSSA/ROSSO
La mela sull'albero	ROSSA/ROSSO
Il secchiello sulla spiaggia	ROSSO/ROSSA
Il secchiello sullo scoglio	ROSSO/ROSSA
La zappa nella buca	ROTTA/ROTTO
La zappa nell'orto	ROTTA/ROTTO
Il binocolo nella tasca	ROVINATA/ROVINATO
Il binocolo sul libro	ROVINATA/ROVINATO
Il cartello sul molo	SCOLORITO/SCOLORITA
Il cartello sulla chiatta	SCOLORITO/SCOLORITA
La piega del cappello	SCUCITA/SCUCITO
La piega della gonna	SCUCITA/SCUCITO
Il cappotto al mercato	SGUALCITA/SGUALCITO
Il cappotto nella vetrina	SGUALCITA/SGUALCITO
La cavalletta nel giardino	SILENZIOSA/SILENZIOSO
La cavalletta nella radura	SILENZIOSA/SILENZIOSO
Lo scoiattolo nel prato	SIMPATICA/SIMPATICO
Lo scoiattolo nell'erba	SIMPATICA/SIMPATICO
La rana nell'acqua	SOLITARIA/SOLITARIO
La rana nello stagno	SOLITARIA/SOLITARIO

La casetta nel bosco
La casetta sulla vetta
La finestra della cucina
La finestra dell'ufficio
L'uccellino nella tasca
L'uccellino sul libro
La lampada nel parco
La lampada nella piazza
La tartaruga sopra il cesto
La tartaruga sopra la panca
Il ghepardo nel villaggio
Il ghepardo nella foresta
La farfalla nel bosco
La farfalla sulla vetta
La lumaca sul pomodoro
La lumaca sull'insalata
Il barattolo nel bagno
Il barattolo sulla scala

SOLITARIO/SOLITARIA
SOLITARIO/SOLITARIA
SPALANCATO/SPALANCATA
SPALANCATO/SPALANCATA
SPAVENTATO/SPAVENTATA
SPAVENTATO/SPAVENTATA
SPENTO/SPENTA
SPENTO/SPENTA
STANCA/STANCO
STANCA/STANCO
TEMUTO/TEMUTA
TEMUTO/TEMUTA
VARIOPINTO/VARIOPINTA
VARIOPINTO/VARIOPINTA
VISCIDO/VISCIDA
VISCIDO/VISCIDA
VUOTA/VUOTO
VUOTA/VUOTO

Annexe 4**Liste des items expérimentaux utilisés dans l'Expérience 4 du Chapitre 2.**

Le saumon de la rivière	FRAIS/FRAÎCHE
Le poussin de la kermesse	BRUYANT/BRUYANTE
Le vautour de la réserve	ÉTONNANT/ÉTONNANTE
Le héron de la lagune	BEAU/BELLE
Le brochet dans la barque	LOURD/LOURDE
Le grillon sur la brindille	LÉGER/LÉGÈRE
Le chaton sur la moquette	DOUX/DOUCE
Le pinson sur la jonquille	DÉLICAT/DÉLICATE
Le gorille de la jungle	EFFRAYANT/EFFRAYANTE
Le matou sur la commode	VILAIN/VILAINE
Le livreur de la pizza	ÉPATANT/ÉPATANTE
Le pirate sur la passerelle	DANGEREUX/DANGEREUSE
Le postier de la bourgade	DISCRET/DISCRETE
Le rodeur dans la forêt	INQUIÉTANT/INQUIÉTANTE
Le prieur dans l'église	SILENCIEUX/SILENCIEUSE
Le sportif sur la piste	RÉSISTANT/RÉSISTANTE
La perruche du marché	FATIGANTE/FATIGUANT
La sauterelle du jardin	PETITE/PETIT
La cigogne du marais	GRANDE/GRAND
L'anguille du vivier	REMUANTE/REMUANT
La crevette dans le filet	GLISSANTE/GLISSANT
La mésange dans le rosier	GRACIEUSE/GRACIEUX
La limace sur le chou-fleur	GROSSE/GROS
La méduse sur le sable	SÈCHE/SEC
La guenon du cirque	MARRANTE/MARRANT
La jument du manège	CRASSEUSE/CRASSEUX
La crémière du quartier	COQUETTE/COQUET
La sorcière du chateau	MAUDITE/MAUDIT
La marraine du projet	PARFAITE/PARFAIT
La serveuse dans le café	CHARMANTE/CHARMANT
La commère dans le fauteuil	CURIEUSE/CURIEUX
La chatelaine sur le balcon	LAIDE/LAID

Annexe 5

Liste des items expérimentaux utilisés dans les Expériences 1 et 3 du Chapitre 3.

A. *Masculine Head Nouns*

1. A metà partita Dino/Carla ha creato confusione per cambiarsi le carte
Half way through the game Dino/Carla created confusion in order to exchange cards
Il baro del gioco
The cheater of the game-masc
Il baro della partita
The cheater of the game-fem
FURBO/FURBA
clever
2. Anche se ha solo sei anni Giovanni/Caterina dà concerti in tutto il mondo
Although (he/she) is only six, Giovanni/Caterina performs concerts around the world
Il prodigio del conservatorio
The prodigy of the music school-masc
Il prodigio dell' accademia
The prodigy of the academy-fem
DOTATA/DOTATO
talented
3. Il programma di riforma illustrato da Silvio/Silvia ha raccolto molti consensi
The program for reform advocated by Silvio/Silvia has received positive feedback
Il capo del governo
The head of the govern-masc
Il capo della giunta
The head of the town council-fem
DETERMINATO/DETERMINATA
determined
4. Il punto forte del programma di Carlo/Carla è il risanamento economico
The strong part of Carlo's/Carla's program is the economic growth
Il membro del Senato
The member of the senate-masc
Il membro della Camera
The member of the Chamber-fem
DINAMICO/DINAMICA
dynamic
5. La lite era accesa ma Simone/Simona si è dimostrata imparziale
The quarrel was acrimonious, but Simone/Simona was impartial
L'arbitro del battibecco
The arbitrator of the quarrel-masc
L'arbitro della contesa
The arbitrator of the contest-fem
EQUILIBRATA/EQUILIBRATO
balanced
6. Le ricerche di biologia molecolare di Michele/Rita sono affascinanti
Michele/Rita's molecular biology research is fascinating
Il genio del Dipartimento
The genius of the department-masc
Il genio della Facolta'
The genius of the College-fem
STRAORDINARIA/STRAORDINARIO
Extraordinary
7. L'imputato è stato scagionato dopo che Sandro/Sandra ha interrogato il testimone

- The defendant has been released after Sandro/Sandra questioned the witness
 L'avvocato del processo
 The lawyer of the trial-masc
 L'avvocato della difesa
 The lawyer of the defense-fem
 METICOLOSA/METICOLOSO
 meticulous
8. Quando va a fare le visite a domicilio Alfredo/Antonella dimentica spesso il ricettario
 When (s/he) visits patients at home, Alfredo/Antonella often forgets the prescription book
 Il medico dell'ospizio
 The doctor of the retirement home-masc
 Il medico della mutua
 The doctor of the health insurance-fem
 SMEMORATO/SMEMORATA
 Absent minded
9. Hanno chiamato Pietro/Piera al posto del commesso e se l'è cavata bene
 (They) called Pietro/Piera to replace a clerk and s/he did well
 Il sostituto nel negozio
 The substitute in the store-masc
 Il sostituto nella bottega
 The substitute in the store-fem
 ESPERTA/ESPERTO
 expert
10. Il migliore della serata è stato Davide/Cristina che ha fatto ridere tutti
 The best one of the night was Davide/Cristina who made everybody laugh
 Il comico del teatro
 The comedian of the theatre-masc
 Il comico della compagnia
 The comedian of the (theatre) company-fem
 SIMPATICA/SIMPATICO
 nice
11. In molti hanno riconosciuto Sergio/Sara che scappava di nascosto dall'abitazione
 Many have recognized Sergio/Sara who was sneaking out the house
 L'individuo del furto
 The individual of the burglary-masc
 L'individuo della rapina
 The individual of the robbery-fem
 SCOPERTO/SCOPERTA
 found
12. La mafia avrebbe fatto l'attentato se Angelo/Angela non avesse presentato denuncia
 The Mafia would have carried out the assassination if Angelo/Angela did not expose them
 Il magistrato del capoluogo
 The magistrate of the capital-masc
 Il magistrato della provincia
 The magistrate of the province-fem
 CORAGGIOSA/CORAGGIOSO
 brave
13. La recensione di Paolo/Paola dava dei giudizi molto severi sul film
 Paolo/Paola's review gave very severe criticisms of the movie
 Il critico del periodico
 The critic of the newspaper-masc
 Il critico della rivista
 The critic of the magazine-fem
 SEVERO/SEVERA
 Harsh
14. Questa mattina la maestra ha interrogato Stefano/Elisa che ha fatto scena muta
 This morning the teacher as questioned Stefano/Elisa who could not answer
 Il somaro del corso
 The idiot of the class-masc

Il somaro della scuola
The idiot of the school-fem
IMPREPARATO/IMPREPARATA
unprepared
15. La storia narra le avventure di Valentino/Valentina attraverso lo spazio e il tempo
The tale tells about the adventures of Valentino/Valentina across space and time
Il personaggio del racconto
The character of the novel-masc
Il personaggio della favola
The character of the tale-fem
CURIOSO/CURIOSA
funny
16. Tutti si fermano a guardare l'immagine di Gianni/Gianna appesa in ingresso
Everybody stopped to look at Gianni/Gianna's portrait in the entrance
Il soggetto del quadro
The subject of the painting-masc
Il soggetto della fotografia
The subject of the photograph-fem
CONOSCIUTA/CONOSCIUTO
Well-known

B. Feminine Head Nouns

1. Ancora una volta il piccolo Lino/Lina ha dato un morso alla torta
Yet again little Lino/Lina had a bite of the cake
La canaglia della mensa
The rascal of the canteen-fem
La canaglia del refettorio
The rascal of the refectory-masc
GOLOSO/GOLOSA
Greedy (with a sweet tooth)
2. C'era molta gente in piazza quando Antonio/Antonia ha lanciato il suo appello
There were many people in the square when Antonio/Antonia made his/her speech
La personalità dell'assemblea
The personality of the assembly-fem
La personalità del partito
The personality of the (political) party-masc
APPLAUDITO/APPLAUDITA
applauded
3. Dopo l'ultimo discorso Guglielmo/Imelda è diventata un simbolo per la gente
After the last speech, Guglielmo/Imelda became a symbol for the people
L'autorità della repubblica
The authority of the republic-fem
L'autorità dello Stato
The authority of the State-masc
STIMATA/STIMATO
renown
4. E' stata assunta Stefano/Patrizia come coordinatrice della campagna elettorale
Stefano/Patrizia has been hired as campaign coordinator
La recluta per la propaganda
The recruit for the campaign-fem
La recluta per il partito
The recruit for the political party-masc
ENERGICO/ENERGICA
energetic
5. I clienti apprezzano Francesco/Francesca che tratta tutti amichevolmente
The costumers like Francesco/Francesca who is friendly with everyone
La persona alla cassa
The person at the register-fem

- La persona nel negozio
The person in the store-masc
EDUCATA/EDUCATO
polite
6. I compagni allontanano Vittorio/Vittoria perché ha raccontato tutto alla maestra
The classmates shunned Vittorio/Vittoria who told what happened to the teacher
La spia della scuola
The spy of the school-fem
La spia dell'istituto
The spy of the institute-masc
ODIATA/ODIATO
hated
7. I fedeli si inchinano ad adorare Mario/Maria durante la celebrazione
The followers bent to adore Mario/Maria during the celebration
La divinita' della setta
The divinity of the sect-fem
La divinità del rito
The divinity of the rite-masc
VENERATO/VENERATA
venerated
8. La sera gli spettatori arrivano insieme, ma Claudio/Claudia accompagna tutti al posto
In the evening, spectators arrive all together, but Claudio/Claudia take each of them to their seat
La maschera dell'opera
The usher/uscherette of the opera-fem
La maschera del teatro
The usher/uscherette of the theatre-masc
INDAFFARATO/INDAFFARATA
busy
9. Le località che Luigi/Luisa ci ha fatto vedere erano tutte molto suggestive
The places that Luigi/Luisa showed us were all very charming
La guida per la gita
The guide for the journey-fem
La guida per il viaggio
The guide for the trip-masc
PREPARATO/PREPARATA
Well-prepared
10. L'ultima rappresentazione di Luciano/Luciana continua a dare repliche
The last show by Luciano/Lucian continues to play
La stella della commedia
The star of the comedy-fem
La stella del palcoscenico
The star of the stage-masc
AMATA/AMATO
loved
11. Nonostante andasse in scena per pochi istanti, Giuseppe/Giuseppa ha fatto faville
Although only on the stage for a few minutes, Giuseppe/Giuseppa was outstanding
La comparsa della commedia
The appearance of the comedy-fem
La comparsa dello spettacolo
The appearance of the show-masc
SICURO/SICURA
Confident
12. Per far ridere i colleghi Tiziano/Tiziana racconta barzellette a raffica
To make (his/her) colleague laugh, Tiziano/Tiziana tells jokes continuously
La macchietta della compagnia
The character of the company-fem
La macchietta del gruppo
The character of the group-masc

BUFFO/BUFFA

funny

13. Pur di compiacere i superiori Alberto/Alberta si presterebbe a fare qualsiasi cosa

To please (his/her) superiors, Alberto/Alberta would do anything

La marionetta della azienda-fem

The puppet of the firm

La marionetta del reparto-masc

The puppet of the division

SCIOCCA/SCIOCCO

silly

14. Sono scrosciati gli applausi quando Valerio/Valeria ha cominciato il suo assolo

Everybody applauded when Valerio/Valeria started his/her solo

La celebrità dell'orchestra

The celebrity of the orchestra-fem

La celebrità del coro

The celebrity of the choir-masc

APPREZZATA/APPREZZATO

appreciated

15. Un camion ha investito Fabio/Fabiola che correva in bicicletta ascoltando musica

A truck ran over Fabio/Fabiola who was biking fast while listening to music

La vittima della fatalità

The victim of the fatality

La vittima dello scontro

The victim of the crash

DISTRATTA/DISTRATTO

Distracted

16. Un giorno a settimana va Andrea/Arianna a prendere i posti a lezione

Once a week, Andrea/Arianna goes early to class to reserve (good) seats

La matricola dell'università

The freshman of the university

La matricola del corso

The freshman of the class-masc

SCRUPOLOSA/SCRUPOLOSO

diligent

Annexe 6

Liste des items expérimentaux utilisés dans les Expériences 2 et 4 du Chapitre 3.

A. *Masculine Head Nouns*

1. Il arrive trop souvent que Bernard/Bernadette oublie d'emporter son carnet de prescriptions
It happens too often that Bernard/Bernadette forgets to bring his (her) prescription book
Le médecin de la famille
The doctor of the family-fem
Le médecin du quartier
The doctor of the area-masc
DISTRAIT/DISTRAITE
Distracted
2. Il semble que Bertrand/Brigitte ait humilié ses élèves plus d'une fois dans sa carrière
It seems that Bertrand/Brigitte humiliated his (her) students more than once in his (her) career
Le professeur de la pension
The professor of the boarding school-fem
Le professeur du pensionnat
The professor of the boarding school-masc
ODIEUX/ODIEUSE
odious
3. On raconte que Maurice/Marie, décédé il y a dix ans, se balade la nuit dans les couloirs
People say that Maurice/Marie who died ten years ago walks around in the corridors at night
Le fantôme de la demeure
The ghost of the house-fem
Le fantôme du château
The ghost of the castle-masc
TROUBLANT/TROUBLANTE
Disconcerting
4. Les recherches de Roland/Rachel en biologie moléculaire comportent des enjeux considérables
The research of Roland/Rachel in molecular biology involves very important stakes
Le génie de la faculté
The genius of the faculty-fem
Le génie du département
The genius of the department-masc
PRUDENT/PRUDENTE
Prudent
5. Parmi les fondateurs de la convention, seul Eric/Céline ne pouvait pas venir
Amongst the creators of the Convention, only Eric/Céline was not able to come
Le défenseur de la charte
The defender of the chart-fem
Le défenseur du traité
The defender of the treaty -masc
SOUFFRANT/SOUFFRANTE
Suffering
6. L'histoire raconte les aventures de Valérian/Laureline à travers le temps et l'espace
The story tells the adventures of Valérian/Laureline through time and space
Le personnage de la BD
The character of the comics-fem
Le personnage du livre
The character of the book-masc
CURIEUX/CURIEUSE
Curious

7. Les violences se sont déroulées sous les yeux de Paul/Pauline accoudé(e) à son balcon
The violence occurred in sight of Paul/Pauline who was standing on the balcony
Le témoin de la bagarre
The witness of the fight-fem
Le témoin du conflit
The witness of the conflict-masc
SUSPECT/SUSPECTE
Suspect
8. La silhouette de Marco/Margot sert à merveille les élèves du conservatoire en sculpture
The silhouette of Marco/Margot is ideal for the students of the conservatory in sculpture
Le modèle de la classe
The model of the class-fem
Le modèle du groupe
The model of the group-masc
PARFAIT/PARFAITE
Perfect
9. Le prix a été décerné à Christophe/Claudine pour sa remarquable intrigue policière
The award was given to Christophe/Claudine for his (her) remarkable detective novel
L'auteur de la nouvelle
The author of the fiction-fem
L'auteur du roman
The author of the story-masc
HEUREUX/HEUREUSE
Happy
10. On reproche parfois à Cédric/Chantal d'être un peu trop sévère dans ses jugements
Cedric/Chantal is sometimes reproached for judging things too severely
Le critique de la revue
The critic of the review-fem
Le critique du périodique
The critic of the journal-masc
EXIGEANT/EXIGEANTE
demanding
11. Le point principal du programme de Denis/Denise concerne l'intégration des immigrés
The main point of the program of Denis/Denise concerns the integration of immigrants
Le membre de la Chambre
The member of the Chamber-fem
Le membre du Sénat
The member of the Senate-masc
ACTIF/ACTIVE
Active
12. Les oeuvres de l'exposition que Damien/Delphine commente sont toutes du courant surréaliste
The works in the exhibition that Damien/Delphine describes are all from the surrealist school
Le guide de la galerie
The guide of the gallery-fem
Le guide du musée
The guide of the museum-masc
CAPTIVANT/CAPTIVANTE
Captivating
13. Depuis Alger, Robert/Edith fait parvenir quotidiennement les nouvelles informations
From Algiers, Robert/Edith daily sends new information
Le correspondant de l'agence
The correspondent of the agency-fem
Le correspondant du journal
The correspondent of the newspaper-masc
INQUIET/INQUIETE
Worried
14. La nouvelle collection automne-hiver est présentée par Gaëtan/Gaëtane cette année

The new autumn-winter collection is presented by Gaëtan/Gaëtane this year

Le mannequin de la revue

The mannequin of the review-fem

Le mannequin du catalogue

The mannequin of the catalog-masc

ATTIRANT/ATTIRANT

Attractive

15. Depuis qu'il a six ans, Philippe/Catherine donne des concerts à travers le monde

Since the age of six, Philippe/Catherine has given concerts throughout the world

Le prodige de l'académie

The prodigy of the academy-fem

Le prodige du conservatoire

The prodigy of the conservatory-masc

BRILLANT/BRILLANTE

Brilliant

16. En vacances, on a fait la connaissance de Pierre/Pierrette qui ressemble fort à quelqu'un qu'on connaît

When we were on holiday we met Pierre/Pierrette who really looked like someone we know

Le sosie de la plage

The double of the beach-fem

Le sosie du camping

The double of the camping-masc

B. *Feminine head nouns*

1. L'interview de Martin/Martine qui joue le rôle principal était fort décevante

The interview with Martin/Martine who plays the main role was very disappointing

La vedette du film

The superstar of the film-masc

La vedette de la série

The superstar of the series-fem

IDIOT/IDIOTE

Idiot

2. Un camion a percuté Fabien/Fabienne qui roulait à vélo et ne l'avait pas entendu

A truck hit Fabien/Fabienne who was riding a bike and had not heard it

La victime de l'accident

The victim of the accident-masc

La victime de la collision

The victim of the collision-fem

SOURD/SOURDE

Deaf

3. Un soir sur quatre c'est Alain/Aline qui assure la surveillance des tentes scout

One night in every four it is Alain/Aline who checks the scout tents

La sentinelle du camp

The guard of the camp-masc

La sentinelle de la troupe

The guard of the group-fem

ATTENTIF/ATTENTIVE

Paying attention

4. Les locataires craignent Louis/Louise qui peut les mettre dehors quand bon lui semble

The tenants fear Louis/Louise who can force them to leave whenever he (she) wants to

L'autorité de la maison

The authority of the house-fem

L'autorité de l'immeuble

The authority of the building-masc

AGRESSIF/AGRESSIVE

Aggressive

5. Les clients apprécient François/Françoise qui les accueille à l'hôtel Métropole

Clients like François/Françoise who welcomes them at the Metropole hotel

- La personne au comptoir
The person at the counter-masc
La personne à la réception
The person at the reception desk-fem
SOURIANT/SOURIANTE
Smiling
6. Des centaines de gens ont défilé dans les rues quand Jean/Jeanne a lancé son appel
Hundreds of people walked in the streets when Jean/Jeanne made the appeal
La personnalité du mouvement
The personality of the movement-masc
La personnalité de la manifestation
The personality of the demonstration-fem
CONVAINQUANT/CONVAINQUANTE
Convincing
7. Il arrive que Gilbert/Gilberte agresse un client parce que sa tête ne lui plaît pas
Sometimes Gilbert/Gilberte attacks a customer because he (she) does not like his face
La brute du café
The brute of the cafe-masc
La brute de la buvette
The brute of the refreshment bar-fem
VIOLENT/VIOLENTE
Violent
8. Ce soir, Pietro/Jacqueline sera sur les planches pour une représentation spéciale
Tonight Pietro/Jacqueline will go on the stage for a special show
La star du théâtre
The star of the theater-masc
La star de la scène
The star of the scene-fem
ANXIEUX/ANXIEUSE
Anxious
9. C'est toujours Luc/Lucie qui hésite lorsqu'il s'agit de faire un coup dangereux
It is always Luc/Lucie who hesitates when we are planning to do something dangerous
La mauviette du groupe
The coward of the group-masc
La mauviette de la bande
The coward of the gang-fem
PEUREUX/PEUREUSE
Scared
10. On a choisi Patrick/Patricia pour gérer la publicité et la propagande
We chose Patrick/Patricia to manage advertisement and propaganda
La recrue du parti
The recruit of the party-masc
La recrue de la direction
The recruit of the management-fem
ELOQUENT/ELOQUENTE
Eloquent
11. On fait faire ce qu'on veut à la pauvre Roger/Rita qui vient de se faire engager
One can have poor Roger/Rita, who has just been hired, do anything
La marionnette du bureau
The marionnette of the office-masc
La marionnette de la firme
The marionnette of the firm-fem
SOU MIS/SOUMISE
Obedient
12. Que ce soit en math, en français, ou en histoire, Olivier/Olivia est toujours premier(e)
Olivier/Olivia is always best, whether in Math, French or History
La tête du cours
The brain of the course-masc
La tête de la classe

- The brain of the class-fem
PRETENTIEUX/PRETENTIEUSE
Pretentious
13. Une fois de plus, le petit Lucas/Lola a été mordre dans les tartelettes aux amandes
Once more, little Lucas/Lola stole bites from the almond pies
La canaille du réfectoire
The rascal of the refectory-masc
La canaille de la crèche
The rascal of the creche-fem
GOURMAND/GOURMANDE
Greedy
14. Plusieurs oscars ont été remis à Léon/Laure pour ses triples pirouettes arrière
Many awards were given to Léon/Laure for his (her) triple backward jumps
L'étoile du patin
The star of the skate-masc
L'étoile de la glace
The star of the ice-fem
ELEGANT/ELEGANTE
Elegant
15. La découverte de Christian/Christiane en génétique l'a porté à la une des journaux
The discovery of Christian/Christiane in genetics brought him (her) to the front page of the newspapers
La sommité du laboratoire
The leading figure of the laboratory-masc
La sommité de la médecine
The leading figure of medicine-fem
DISCRÈT/DISCRETE
Discrete
16. Grâce à sa grande renommée, Fidel/Margaret est devenu(e) un symbole pour le peuple
Thanks to his (her) fame, Fidel/Margaret became a symbol for the people
La célébrité du pays
The celebrity of the country-masc
La célébrité de la nation
The celebrity of the nation-fem
INFLUENT/INFLUENTE
Influential

Annexe 7

Liste des items expérimentaux utilisés dans l'expérience du Chapitre 4.

La bague de la fiancée	BRILLANT/BRILLANTE
La bague de la mauviette	BRILLANT/BRILLANTE
La bague du fiancé	BRILLANT/BRILLANTE
La bague du professeur	BRILLANT/BRILLANTE
La blouse de la bouchère	REPUGNANT/RÉPUGNANTE
La blouse de la teigne	REPUGNANT/RÉPUGNANTE
La blouse du boucher	REPUGNANT/RÉPUGNANTE
La blouse du sosie	REPUGNANT/RÉPUGNANTE
La démarche de la comédienne	IRRITANT/IRRITANTE
La démarche de la fripouille	IRRITANT/IRRITANTE
La démarche du comédien	IRRITANT/IRRITANTE
La démarche du personnage	IRRITANT/IRRITANTE
La démission de la célébrité	TROUBLANT/TROUBLANTE
La démission de la directrice	TROUBLANT/TROUBLANTE
La démission du directeur	TROUBLANT/TROUBLANTE
La démission du génie	TROUBLANT/TROUBLANTE
La dissertation de la gagnante	CAPTIVANT/CAPTIVANTE
La dissertation de la recrue	CAPTIVANT/CAPTIVANTE
La dissertation du gagnant	CAPTIVANT/CAPTIVANTE
La dissertation du prodige	CAPTIVANT/CAPTIVANTE
La maison de la personne	FABULEUX/FABULEUSE
La maison de la voisine	FABULEUX/FABULEUSE
La maison du médecin	FABULEUX/FABULEUSE
La maison du voisin	FABULEUX/FABULEUSE
La réflexion de la mère	HUMILIANT/HUMILIANTE
La réflexion de la rosse	HUMILIANT/HUMILIANTE
La réflexion du juge	HUMILIANT/HUMILIANTE
La réflexion du père	HUMILIANT/HUMILIANTE
La remarque de la cliente	IDIOT/IDIOTE
La remarque de la star	IDIOT/IDIOTE
La remarque du client	IDIOT/IDIOTE
La remarque du guide	IDIOT/IDIOTE
Le cahier de la canaille	DECEVANT/DÉCEVANTE
Le cahier de la gamine	DECEVANT/DÉCEVANTE
Le cahier du cancre	DECEVANT/DÉCEVANTE
Le cahier du gamin	DECEVANT/DÉCEVANTE
Le discours de la personnalité	HARGNEUX/HARGNEUSE
Le discours de la présidente	HARGNEUX/HARGNEUSE
Le discours du défenseur	HARGNEUX/HARGNEUSE
Le discours du président	HARGNEUX/HARGNEUSE
Le nez de la brute	OSSEUX/OSSEUSE
Le nez de la voleuse	OSSEUX/OSSEUSE
Le nez du gourmet	OSSEUX/OSSEUSE
Le nez du voleur	OSSEUX/OSSEUSE
Le parapluie de la sentinelle	EPATANT/ÉPATANTE
Le parapluie de la voyageuse	EPATANT/ÉPATANTE
Le parapluie du magistrat	EPATANT/ÉPATANTE
Le parapluie du voyageur	EPATANT/ÉPATANTE
Le récit de la crapule	ACCABLANT/ACCABLANTE
Le récit de la fermière	ACCABLANT/ACCABLANTE
Le récit du fermier	ACCABLANT/ACCABLANTE

Le récit du témoin	ACCABLANT/ACCABLANTE
Le slogan de la militante	SCANDALEUX/SCANDALEUSE
Le slogan de la sommité	SCANDALEUX/SCANDALEUSE
Le slogan du militant	SCANDALEUX/SCANDALEUSE
Le slogan du parasite	SCANDALEUX/SCANDALEUSE
Le teint de la fumeuse	MALADIF/MALADIVE
Le teint de la vedette	MALADIF/MALADIVE
Le teint du fumeur	MALADIF/MALADIVE
Le teint du modèle	MALADIF/MALADIVE
Le témoignage de la boulangère	CONVAINCANT/CONVAINCANTE
Le témoignage de la victime	CONVAINCANT/CONVAINCANTE
Le témoignage du boulanger	CONVAINCANT/CONVAINCANTE
Le témoignage du mannequin	CONVAINCANT/CONVAINCANTE

Annexe 8

Liste des items expérimentaux utilisés dans les Expériences 1 et 3 du Chapitre 6 (dans la condition Singulier, Singulier, Singulier).

L'ordinateur avec le programme de l'expérience
L'exposition avec la peinture de l'enfant
Le trousseau avec la clé de la cellule
Le livre sur la privatisation de l'entreprise
La conférence sur l'histoire de la guerre
L'étude sur l'effet de la drogue
La théorie sur le chant du canari
L'inscription sur la porte de la toilette
Le dépliant sur l'enjeu de la communauté
La ligne sur le bord de la route
La nappe sur la table du banquet
Le récit de la souffrance dans l'hôpital
La visite au palais de l'artisanat
La menace de la réforme dans l'école
L'hélicoptère pour le vol au-dessus du canyon
La soirée dans l'appartement du sous-sol
Le producteur du film sur l'artiste
Le chat sur le toit du voisin
Le chien sur le coussin du bébé
Le gréviste sur l'escalier du patron
L'architecte du château de la millionnaire
La victime du mensonge du policier
Le voleur de la voiture du garagiste
Le responsable du droit de l'immigré
Le gamin du clochard près de la poubelle
L'avocat de l'assassin du village
Le cochon du laboureur du champs
La copine du propriétaire de la villa
L'accompagnatrice du musicien de la rue
Le professeur de mon cousin sur la photo
L'entraîneur du vainqueur du match
Le possesseur du perroquet dans l'arbre

Annexe 9**Liste des items expérimentaux utilisés dans les expériences 2 et 4 du Chapitre 6 (dans la condition Singulier, Singulier, Singulier).**

The advertisement from the office of the real estate agent
The announcement by the director of the foundation
The article by the writer for the magazine
The author of the speech about the city
The computer with the program for the experiment
The contract for the actor in the film
The discussion about the topic of the paper
The dog on the path around the lake
The friend of the editor of the magazine
The gift for the daughter of the visitor
The helicopter for the flight over the canyon
The lesson about the government of the country
The letter from the friend of my cousin
The manual by the developer of the machine
The mast on the deck of the yacht
The meal for the guest of the inn-keeper
The museum with the picture of the poet
The new design for the engine of the plane
The payment for the service to the school
The photo of the girl with the baby
The post in the support for the platform
The prescription by the doctor from the clinic
The producer of the movie about the artist
The publisher of the book about the king
The setting for the movie about the astronomer
The statue in the garden near the mansion
The switch for the light on the stairway
The telegram to the friend of the soldier
The threat to the president of the company
The tour of the museum near the monument
The train to the city on the lake
The truck on the bridge over the stream

Annexe 10

Liste des items expérimentaux utilisés dans l'expérience du Chapitre 7 (dans la condition Singulier, Pluriel, Singulier).

A. Condition Animé, Inanimé, Inanimé

Le producteur des films qui assiste au festival/se jouent au festival
Le compositeur des mélodies qui chantonne dans l'église/résonnent dans l'église
Le créateur des icônes qui s'approche de l'écran/s'effacent de l'écran
L'architecte des bâtiments qui mesure le clocher/dépassent le clocher
Le voleur des diamants qui court sur le trottoir/roulent sur le trottoir
Le fabricant des patins qui danse sur la glace/crissent sur la glace
Le capitaine des bateaux qui hurle dans la tempête/tangent dans la tempête
Le témoin des conflits qui se promène dans le camp/se déroulent dans le camp
La vendeuse des tissus qui apporte le fauteuil/recouvrent le fauteuil
La victime des microbes qui ameute le village/ravagent le village
La journaliste des périodiques qui critique l'article/comportent l'article
L'auteur des livres qui s'appuie sur l'étagère/reposent sur l'étagère
Le livreur des caisses qui sonne à la porte/obstruent la porte
Le gardien des maisons qui marche dans le parc/donnent sur le parc
Le propriétaire des étangs qui habite derrière le bois/stagnent derrière le bois
Le cultivateur des arbustes qui travaille dans le champ/bourgeonnent dans le champ

B. Condition Inanimé, Animé, Inanimé

La marmite des scouts qui chauffe sur le feu/chantent près du feu
L'exposition des artistes qui se termine le dimanche/s'absentent le dimanche
La valise des bandits qui renferme le revolver/exigent le revolver
Le blé des fermiers qui germe sous le soleil/bêchent sous le soleil
La clé des stagiaires qui rouille sous le paillason/fouillent sous le paillason
La voiture des gangsters qui grince dans le virage/braquent dans le virage
Le linge des ménagères qui s'envole sur l'herbe/se couchent sur l'herbe
Le plancher des voisins qui craque la nuit/marchent la nuit
La dentelle des couturières qui s'abîme sur la table/travaillent sur la table
L'affiche des organisateurs qui se décolle du panneau/se chargent du panneau
La grève des réfugiés qui dure depuis un mois/jeunent depuis un mois
La tente des vacanciers qui perce sous l'averse/courent sous l'averse
Le plâtre des plafonniers qui sèche dans le seau/crachent dans le seau
Le potage des gamins qui coule sur le tapis/mangent sur le tapis
Le chocolat des gosses qui colle sur le clavier/tapent sur le clavier
L'alarme des locataires qui s'enclenche la nuit/rentrent la nuit