

La position de l'adjectif en français : approche quantitative sur corpus et expériences psycholinguistiques

Juliette Thuilier¹

¹Université Toulouse Jean Jaurès & CLLE-ERSS
juliette.thuilier@univ-toulouse2.fr

26 mars 2015

Séminaire ERSS
Université Toulouse Jean Jaurès

Collaborations



Gwendoline Fox
Université Paris 3 – Sorbonne
Nouvelle



Benoît Crabbé
ALPAGE
Université Paris Diderot



Margaret Grant
University of Toronto

Introduction

- En français, l'adjectif épithète peut apparaître en position pré-nominale ou post-nominale

- (1)
- a. une **agréable** soirée
 - b. une soirée **agréable**

Introduction

- En français, l'adjectif épithète peut apparaître en position pré-nominale ou post-nominale

- (1)
- a. une **agréable** soirée
 - b. une soirée **agréable**

- ⇒ Phénomène largement étudié en linguistique française (Forsgren, 1978; Wilmet, 1981; Nølke, 1996; Noailly, 1999; Abeillé et Godard, 1999, parmi d'autres)
- ⇒ De nombreux facteurs interviennent dans le choix de la position

Introduction

- En français, l'adjectif épithète peut apparaître en position pré-nominale ou post-nominale

- (1) a. une **agréable** soirée
b. une soirée **agréable**

- ⇒ Phénomène largement étudié en linguistique française (Forsgren, 1978; Wilmet, 1981; Nølke, 1996; Noailly, 1999; Abeillé et Godard, 1999, parmi d'autres)
- ⇒ De nombreux facteurs interviennent dans le choix de la position

Objectifs du travail présenté aujourd'hui

- ▶ Proposer une modélisation statistique multifactorielle qui rende compte de la position de l'adjectif pour des données attestées
- ▶ Explorer le rôle d'un facteur syntaxique à l'aide de données expérimentales

- 1 La position de l'adjectif
 - Aspects sémantiques
 - Aspects lexicaux
 - Aspects syntaxiques
 - Combinaisons particulières de N et de A

- 2 Modélisation des données de corpus
 - Méthodologie
 - Observations
 - Modélisation statistique multifactorielle
 - Résultats

- 3 Expérience de rappel de phrases
 - Motivation
 - Méthodologie
 - Résultats
 - Discussion

- 4 Conclusion et perspectives

1. Aspects sémantiques

- Les adjectifs subsectifs (2-a) et intensionnels (2-b) ont tendance à être antéposés

- (2) a. une **petite** souris
b. un **vrai** complot

1. Aspects sémantiques

- Les adjectifs subsectifs (2-a) et intensionnels (2-b) ont tendance à être antéposés

- (2) a. une **petite** souris
b. un **vrai** complot

- Les adjectifs intersectifs tendent à apparaître en postposition

- (3) a. une table **ronde**
b. un vase **fragile**

1. Aspects sémantiques

- Les adjectifs subsectifs (2-a) et intensionnels (2-b) ont tendance à être antéposés

- (2) a. une **petite** souris
b. un **vrai** complot

- Les adjectifs intersectifs tendent à apparaître en postposition

- (3) a. une table **ronde**
b. un vase **fragile**

- Cependant, il n'existe pas de propriété sémantique catégoriquement associée à une position (Abeillé et Godard, 1999)

1. Aspects sémantiques

- Les adjectifs subsectifs (2-a) et intensionnels (2-b) ont tendance à être antéposés

- (2) a. une **petite** souris
b. un **vrai** complot

- Les adjectifs intersectifs tendent à apparaître en postposition

- (3) a. une table **ronde**
b. un vase **fragile**

- Cependant, il n'existe pas de propriété sémantique catégoriquement associée à une position (Abeillé et Godard, 1999)

⇒ **Hypothèse de travail** : la dimension sémantique ne permet pas d'expliquer le phénomène dans son ensemble (*contra* Bouchard, 1998) et le choix de la position de l'adjectif peut s'expliquer en grande partie grâce aux propriétés lexicales et aux contraintes syntaxiques

2. Aspects lexicaux

- Les adjectifs présentent des préférences individuelles pour une position qui sont déterminées/façonnées par des propriétés formelles : la longueur, la fréquence et la complexité morphologique (Wilmet, 1981; Thuilier et al., 2010; Fox et Thuilier, 2012)

Antéposition	NOM	Postposition
court		long
fréquent		rare
morphologiquement simple		morphologiquement complexe

3. Aspects syntaxiques

- 1 Unique contrainte catégorique : la présence d'un dépendant post-adjectival rend la postposition obligatoire

- (4)
- a. une musique **agréable à écouter**
 - b. *une **agréable à écouter** musique

3. Aspects syntaxiques

- 1 Unique contrainte catégorique : la présence d'un dépendant post-adjectival rend la postposition obligatoire

- (4) a. une musique **agréable à écouter**
b. *une **agréable à écouter** musique

- 2 La présence d'autres dépendants du nom favorise l'antéposition de l'adjectif (Grevisse et Goosse, 2007)

- (5) a. un **récent** recueil [de textes grecs]_{SP}
b. l'**habituel** air [que joue Paul]_{REL}
c. un **étrange** animal [indomptable]_A

3. Aspects syntaxiques

- 3 La présence d'un modifieur pré-adjectival et la coordination de plusieurs adjectifs sont possibles dans les deux positions

- (6) a. une **très agréable** soirée
b. une soirée **très agréable**
- (7) a. une **belle et longue** table
b. une table **belle et longue**

3. Aspects syntaxiques

3 La présence d'un modifieur pré-adjectival et la coordination de plusieurs adjectifs sont possibles dans les deux positions

- (6) a. une **très agréable** soirée
- b. une soirée **très agréable**
- (7) a. une **belle et longue** table
- b. une table **belle et longue**

→ La présence d'un adverbe ainsi que la coordination favorisent la mobilité du SAdj (Abeillé et Godard, 1999)

- (8) a. un bon poulet / ?un poulet bon
- b. un **très bon** poulet / un poulet **très bon**
- (9) a. des faux coupables / ?des coupables faux
- b. des vrais coupables / ?des coupables vrais
- c. des **vrais ou faux** coupables / des coupables **vrais ou faux**

3. Aspects syntaxiques

3 La présence d'un modifieur pré-adjectival et la coordination de plusieurs adjectifs sont possibles dans les deux positions

- (6) a. une **très agréable** soirée
- b. une soirée **très agréable**
- (7) a. une **belle et longue** table
- b. une table **belle et longue**

→ La présence d'un adverbe ainsi que la coordination favorisent la mobilité du SAdj (Abeillé et Godard, 1999)

- (8) a. un bon poulet / ?un poulet bon
- b. un **très bon** poulet / un poulet **très bon**
- (9) a. des faux coupables / ?des coupables faux
- b. des vrais coupables / ?des coupables vrais
- c. des **vrais ou faux** coupables / des coupables **vrais ou faux**

→ Les données de corpus montrent que les adjectifs coordonnés ou modifiés sont majoritairement postposés (Forsgren, 1978; Thuilier et al., 2012)

3. Aspects syntaxiques

- 4 Les déterminants démonstratifs et possessifs, ainsi que l'article défini favorisent l'antéposition (Forsgren, 1978)

- (10) a. un **éblouissant** spectacle
b. cet **éblouissant** spectacle
- (11) a. un **habituel** refrain
b. son **habituel** refrain
- (12) a. une très **attendue** décision
b. la très **attendue** décision

4. Combinaisons particulières de N et de A

- Un certain nombre de paires Adjectif - Nom sont fortement collocationnelles

- (13)
- a. un **vibrant** hommage
 - b. un **lourd** tribut

4. Combinaisons particulières de N et de A

- Un certain nombre de paires Adjectif - Nom sont fortement collocationnelles

- (13) a. un **vibrant** hommage
b. un **lourd** tribut

- On observe également que pour certaines paires Adjectif - Nom, la position de l'adjectif est corrélée à une interprétation particulière

- (14) a. un **gros** fumeur (*personne qui fume beaucoup*)
b. un fumeur **gros** (*personne qui fume et qui est grosse*)

- (15) a. un **gros** coiffeur
b. un coiffeur **gros**

4. Combinaisons particulières de N et de A

- Un certain nombre de paires Adjectif - Nom sont fortement collocationnelles

- (13) a. un **vibrant** hommage
b. un **lourd** tribut

- On observe également que pour certaines paires Adjectif - Nom, la position de l'adjectif est corrélée à une interprétation particulière

- (14) a. un **gros** fumeur (*personne qui fume beaucoup*)
b. un fumeur **gros** (*personne qui fume et qui est grosse*)

- (15) a. un **gros** coiffeur
b. un coiffeur **gros**

⇒ Le nom qui se combine avec l'adjectif peut avoir une influence sur le choix de la position

- 1 La position de l'adjectif
 - Aspects sémantiques
 - Aspects lexicaux
 - Aspects syntaxiques
 - Combinaisons particulières de N et de A

- 2 **Modélisation des données de corpus**
 - **Méthodologie**
 - **Observations**
 - **Modélisation statistique multifactorielle**
 - **Résultats**

- 3 Expérience de rappel de phrases
 - Motivation
 - Méthodologie
 - Résultats
 - Discussion

- 4 Conclusion et perspectives

1. Méthodologie

- L'idée est de tester les facteurs rencontrés dans la littérature sur des données attestées
 - ▶ Idée inspirée de travaux sur l'alternance dative en anglais (Bresnan et al., 2007; Bresnan et Hay, 2008; Bresnan et Ford, 2010)
 - ▶ Modélisation statistique : régression logistique (Agresti, 2007) et modèles à effets-mixtes (Gelman et Hill, 2006)
 - ▶ Hypothèse : ces outils statistiques nous permettent de généraliser à partir d'un échantillon de langue, car ils permettent de s'affranchir des variations dues à l'échantillonnage du corpus
 - ▶ Avantage : modèle prédictif

1. Méthodologie

- L'idée est de tester les facteurs rencontrés dans la littérature sur des données attestées
 - ▶ Idée inspirée de travaux sur l'alternance dative en anglais (Bresnan et al., 2007; Bresnan et Hay, 2008; Bresnan et Ford, 2010)
 - ▶ Modélisation statistique : régression logistique (Agresti, 2007) et modèles à effets-mixtes (Gelman et Hill, 2006)
 - ▶ Hypothèse : ces outils statistiques nous permettent de généraliser à partir d'un échantillon de langue, car ils permettent de s'affranchir des variations dues à l'échantillonnage du corpus
 - ▶ Avantage : modèle prédictif
- Les données sont extraites de deux corpus :
 - ▶ *French Treebank* (FTB, Abeillé et Barrier, 2004) : corpus journalistique
 - extraction automatique des adjectifs épithètes qui apparaissent dans les deux positions
 - élimination des adjectifs apparaissant avec un dépendant post-adjectival
 - ▶ C-ORAL-Rom (CORAL, Cresti et Moneglia, 2005) : corpus d'oral
 - extraction manuelle des mêmes adjectifs que ceux de FTB

1. Methodologie

- 11 variables annotées

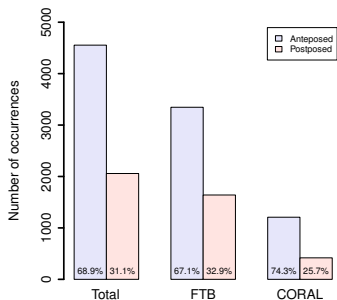
	VARIABLES	DESCRIPTION
1	COORD	l'adjectif est coordonné ou non
2	MOD	l'adjectif apparaît avec un modifieur pré-adjectival ou non
3	DEMDDET	le SN est introduit par un démonstratif ou non
4	POSSDET	le SN est introduit par un possessif ou non
5	DEFART	le SN est introduit par un article défini ou non
6	SP	le SN contient un SP ou non
7	REL	le SN contient une relative ou non
8	POSTADJ	le SN contient un adjectif postposé ou non
9	COLLOCAN	score de collocation pour le bigramme A+N ($\log(\chi^2)$)
10	COLLOCNA	score de collocation pour le bigramme N+A ($\log(\chi^2)$)
11	MODALITÉ	oral (s) ou écrit (w)

1. Methodologie

- Différentiation de 2 lemmes adjectivaux pour les 5 formes suivantes
 - ▶ *seul*
(16) un seul homme / un homme seul
 - ▶ *simple*
(17) une simple phrase / une phrase simple
 - ▶ *pur*
(18) un pur produit / un produit pur
 - ▶ *propre*
(19) son propre pantalon / son pantalon propre
 - ▶ *ancien*
(20) un coffre ancien / un ancien coffre
- Parmi ces 10 lemmes, seuls 5 apparaissent dans les deux positions dans les données et ont donc été gardés pour la modélisation

2. Observations

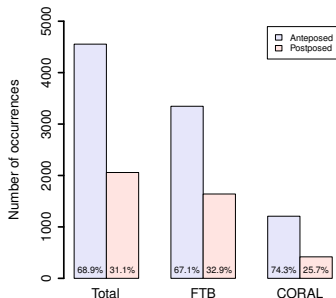
- 6612 occurrences d'adjectifs alternants
 - ▶ 4494 dans le FTB
 - ▶ 1627 dans CORAL



2. Observations

- 6612 occurrences d'adjectifs alternants

- ▶ 4494 dans le FTB
- ▶ 1627 dans CORAL



- Répartition des lemmes dans les données

	FTB	CORAL
Nombre de lemmes	170	130
Lemmes alternants	170	56

- Seuls 43% des lemmes apparaissent dans les deux positions dans les données de CORAL

⇒ Moins d'alternance en français oral

3. Modélisation statistique multifactorielle

Regression logistique (Agresti, 2007)

- modélisation du comportement d'une variable binaire en fonction de plusieurs variables prédictives
- position de l'adjectif = variable binaire
 - postposition = 0
 - antéposition = 1

Fonction logistique

Fonction à valeur dans l'intervalle $[0,1]$

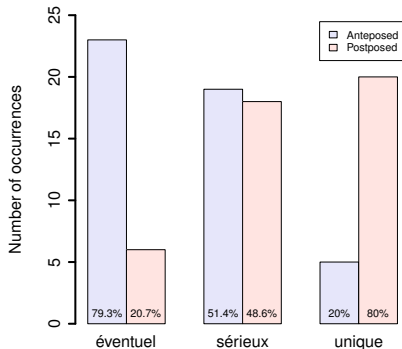
$$\pi_{\text{ante}} = \frac{e^{\beta\mathbf{X}}}{1 + e^{\beta\mathbf{X}}} \quad (1)$$

où

- π_{ante} = probabilité d'antéposition de l'adjectif
- β = coefficients de régression $\alpha, \beta_0 \dots \beta_n$
- \mathbf{X} = variables prédictives $X_0 \dots X_n$

3. Modélisation statistique multifactorielle

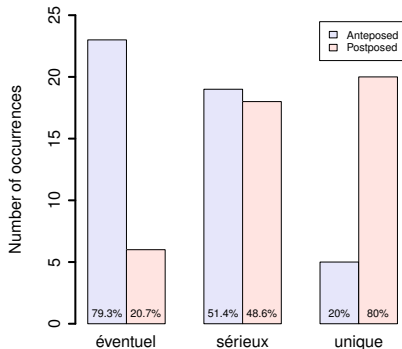
- Les adjectifs comme source de variation dans les données



Comment capturer les caractéristiques de chaque groupe de données ?

3. Modélisation statistique multifactorielle

- Les adjectifs comme source de variation dans les données

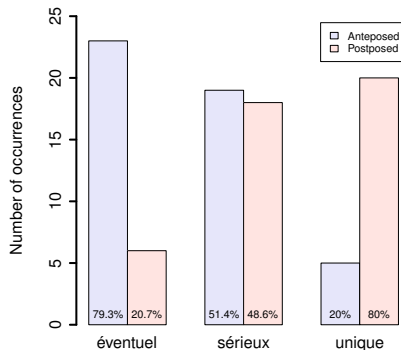


Comment capturer les caractéristiques de chaque groupe de données ?

⇒ Modèles à effets-mixtes (Gelman et Hill, 2006)

3. Modélisation statistique multifactorielle

- Les adjectifs comme source de variation dans les données

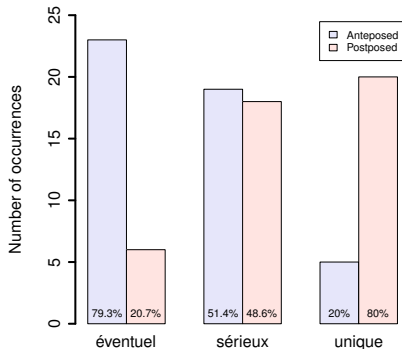


Comment capturer les caractéristiques de chaque groupe de données ?

- ⇒ Modèles à effets-mixtes (Gelman et Hill, 2006)
 - ▶ L'idée : au-delà du modèle général (effet fixe), chaque groupe de données reçoit son propre coefficient (effet aléatoire)

3. Modélisation statistique multifactorielle

- Les adjectifs comme source de variation dans les données



Comment capturer les caractéristiques de chaque groupe de données ?

- ⇒ Modèles à effets-mixtes (Gelman et Hill, 2006)
 - ▶ L'idée : au-delà du modèle général (effet fixe), chaque groupe de données reçoit son propre coefficient (effet aléatoire)
 - ▶ Chaque lemme adjectival a un coefficient captant son comportement individuel

3. Modélisation statistique multifactorielle

Le modèle

- Le modèle est construit avec les 11 variables prédictrices et un effet aléatoire
- Tous les effets sont statistiquement significatifs et participent donc à la prédiction de la position de l'adjectif épithète
- La qualité de prédiction du modèle est très bonne
 - ▶ exactitude moyenne : 0.88 (validation croisée 10 passes)
 - ▶ *concordance probability* moyenne : $C=0.947$ (validation croisée 10 passes)
- Colinéarité faible : indice de conditionnement (*condition number*)
 $\kappa = 8.15$

Effets aléatoires

Variance Ecart-type

ADJECTIF 2.3938 1.5472

Nombre d'obs: 6612, groupes: ADJECTIF, 170

Effets fixes

	Estimation	Err-type	valeur z	P(> z)
Intercept	-0.782	0.182	-4.304	1.68e-05 ***
DEMDET=1	1.226	0.246	4.99	5.9e-07 ***
POSSDET=1	1.185	0.235	5.04	4.7e-07 ***
DEFART=1	0.370	0.107	3.47	0.0005 ***
POSTADJ=1	0.587	0.154	3.82	0.0001 ***
SP=1	0.840	0.104	8.04	8.6e-16 ***
REL=1	0.714	0.210	3.40	0.0007 ***
COLLOCAN	0.378	0.018	20.52	<2e-16 ***
MOD=1	-1.957	0.174	-11.26	<2e-16 ***
COORD=1	-1.266	0.266	-4.76	1.9e-06 ***
COLLOCNA	-0.443	0.020	-22.12	< 2e-16 ***
MODALITÉ = w	0.458	0.121	3.78	0.0002 ***

4. Résultats

- Chaque coefficient associé à une variable prédictrice peut être interprété comme une préférence pour une position
 - ▶ coefficient positif = préférence pour l'antéposition
 - ▶ coefficient négatif = préférence pour la postposition

4. Résultats

- Chaque coefficient associé à une variable prédictrice peut être interprété comme une préférence pour une position
 - ▶ coefficient positif = préférence pour l'antéposition
 - ▶ coefficient négatif = préférence pour la postposition
- L'antéposition est favorisée par :
 - ▶ les déterminants démonstratifs, possessifs et définis
 - ▶ la présence d'un autre dépendant nominal (relative, SP et adjectif postposé)
 - ▶ haut score de collocation pour le bigramme A+N

4. Résultats

- Chaque coefficient associé à une variable prédictrice peut être interprété comme une préférence pour une position
 - ▶ coefficient positif = préférence pour l'antéposition
 - ▶ coefficient négatif = préférence pour la postposition
- L'antéposition est favorisée par :
 - ▶ les déterminants démonstratifs, possessifs et définis
 - ▶ la présence d'un autre dépendant nominal (relative, SP et adjectif postposé)
 - ▶ haut score de collocation pour le bigramme A+N
- La postposition est favorisée par :
 - ▶ un SAdj complexe et « lourd » (adjectifs modifiés et coordonnés)
 - ▶ haut score de collocation pour le bigramme N+A

4. Résultats

1 Moins d'alternance dans les données d'oral

- ⇒ Dans le discours non-planifié, les locuteurs pourraient avoir tendance à respecter les préférences lexicales de façon plus importante (Benzitoun, 2013)

4. Résultats

1 Moins d'alternance dans les données d'oral

⇒ Dans le discours non-planifié, les locuteurs pourraient avoir tendance à respecter les préférences lexicales de façon plus importante (Benzitoun, 2013)

2 Préférences lexicales et combinaisons Nom - Adjectif

- ▶ Chaque adjectif a une préférence plus ou moins forte pour une position (effet aléatoire)
- ▶ Le nom avec lequel se combine l'adjectif peut favoriser la position non-préférée :
 - *fort* a une préférence pour l'antéposition (60%)
 - combiné avec le nom *point*, la postposition est fortement préférée

4. Résultats

1 Moins d'alternance dans les données d'oral

⇒ Dans le discours non-planifié, les locuteurs pourraient avoir tendance à respecter les préférences lexicales de façon plus importante (Benzitoun, 2013)

2 Préférences lexicales et combinaisons Nom - Adjectif

- ▶ Chaque adjectif a une préférence plus ou moins forte pour une position (effet aléatoire)
- ▶ Le nom avec lequel se combine l'adjectif peut favoriser la position non-préférée :
 - *fort* a une préférence pour l'antéposition (60%)
 - combiné avec le nom *point*, la postposition est fortement préférée

3 Effet des déterminants : déterminants définis favorisent l'antéposition

⇒ Dans les contextes anaphoriques, quand la relation entre le référent du nom et la propriété dénotée par l'adjectif est établie, l'antéposition est facilitée (Waugh, 1977)

4. Résultats

- 4 Effet du poids (Abeillé et Godard, 1999, 2000, 2004; Wasow, 1997, 2002)
 - ▶ Syntagmes adjectivaux longs et complexes tendent à être postposés
 - ▶ Résultat en accord avec la généralisation selon laquelle dans les langues SVO, les constituants lourds tendent à apparaître en dernier (Hawkins, 1994)
 - ▶ Interaction significative entre les variables MOD et MODALITÉ : l'effet de la présence d'un modifieur pré-adjectival est plus important dans les données orales (Thuilier, 2013)
- ⇒ Dans le discours non-planifié, les locuteurs auraient tendance à suivre plus souvent des principes généraux tels que « court avant long »

- ⇒ Méthodes quantitatives utilisées pour décrire et modéliser un phénomène syntaxique
- ⇒ Quelques pistes de travail à partir des résultats obtenus :
 - ▶ mieux comprendre les « préférences lexicales »
 - longueur et fréquence (Thuilier et al., 2010; Thuilier, 2012)
 - dimension sémantique (projet ANR SPAM en cours de soumission)
 - ▶ mieux comprendre l'effet de certaines contraintes syntaxiques
 - Effet de la présence d'un dépendant du nom sur la position de l'adjectif

- 1 La position de l'adjectif
 - Aspects sémantiques
 - Aspects lexicaux
 - Aspects syntaxiques
 - Combinaisons particulières de N et de A

- 2 Modélisation des données de corpus
 - Méthodologie
 - Observations
 - Modélisation statistique multifactorielle
 - Résultats

- 3 **Expérience de rappel de phrases**
 - **Motivation**
 - **Méthodologie**
 - **Résultats**
 - **Discussion**

- 4 Conclusion et perspectives

1. Motivation

- Les résultats de la modélisation statistique montre que la présence d'un dépendant post-nominal favorise l'antéposition de l'adjectif :

- (21) il me manquait un **énorme** sac [où c'est qu'il y avait le caméscope dedans] (CORAL)
- (22) c' est la **principale** étape [de la création d'une pièce] (CORAL)
- (23) il privilégie une série de grandes priorités [...] dont découlent d'**importantes** dépenses [nouvelles] (CORAL)

1. Motivation

- Les résultats de la modélisation statistique montre que la présence d'un dépendant post-nominal favorise l'antéposition de l'adjectif :

- (21) il me manquait un **énorme** sac [où c'est qu'il y avait le caméscope dedans] (CORAL)
- (22) c' est la **principale** étape [de la création d'une pièce] (CORAL)
- (23) il privilégie une série de grandes priorités [...] dont découlent d'**importantes** dépenses [nouvelles] (CORAL)

Grevisse et Goosse (2007) : « *Quand le nom est suivi d'un complément, on préfère souvent, du moins dans la langue écrite, ne pas les séparer, et par conséquent, on antépose l'adjectif* »

1. Motivation

- Les résultats de la modélisation statistique montre que la présence d'un dépendant post-nominal favorise l'antéposition de l'adjectif :

- (21) il me manquait un **énorme** sac [où c'est qu'il y avait le caméscope dedans] (CORAL)
- (22) c' est la **principale** étape [de la création d'une pièce] (CORAL)
- (23) il privilégie une série de grandes priorités [...] dont découlent d'**importantes** dépenses [nouvelles] (CORAL)

Grevisse et Goosse (2007) : « *Quand le nom est suivi d'un complément, on préfère souvent, du moins dans la langue écrite, ne pas les séparer, et par conséquent, on antépose l'adjectif* »

Questions

- Cet effet concerne-t-il principalement l'écrit, le discours planifié ?
- Cet effet peut-il émerger dans des conditions expérimentales, lorsque l'intégralité des paramètres est contrôlée ?

2. Méthodologie : Expérience de rappel de phrases

Principe

On demande à des sujets de répéter des phrases qu'ils ont préalablement lues, et on compare l'ordre des mots dans leurs productions par rapport à l'ordre présenté dans l'*input*

2. Méthodologie : Expérience de rappel de phrases

Principe

On demande à des sujets de répéter des phrases qu'ils ont préalablement lues, et on compare l'ordre des mots dans leurs productions par rapport à l'ordre présenté dans l'*input*

Items expérimentaux

ANT	sans SP	Depuis quelques années, ce produit connaît une considérable évolution.
POST	sans SP	Depuis quelques années, ce produit connaît une évolution considérable.
ANT	avec SP	Depuis quelques années, ce produit connaît une considérable évolution de qualité.
POST	avec SP	Depuis quelques années, ce produit connaît une évolution considérable de qualité.

2. Méthodologie : Expérience de rappel de phrases

Principe

On demande à des sujets de répéter des phrases qu'ils ont préalablement lues, et on compare l'ordre des mots dans leurs productions par rapport à l'ordre présenté dans l'*input*

Items expérimentaux

ANT	sans SP	Depuis quelques années, ce produit connaît une considérable évolution.
POST	sans SP	Depuis quelques années, ce produit connaît une évolution considérable.
ANT	avec SP	Depuis quelques années, ce produit connaît une considérable évolution de qualité.
POST	avec SP	Depuis quelques années, ce produit connaît une évolution considérable de qualité.

- 24 items expérimentaux de ce type, avec 12 adjectifs différents

2. Méthodologie : Expérience de rappel de phrases

Principe

On demande à des sujets de répéter des phrases qu'ils ont préalablement lues, et on compare l'ordre des mots dans leurs productions par rapport à l'ordre présenté dans l'*input*

Items expérimentaux

ANT	sans SP	Depuis quelques années, ce produit connaît une considérable évolution.
POST	sans SP	Depuis quelques années, ce produit connaît une évolution considérable.
ANT	avec SP	Depuis quelques années, ce produit connaît une considérable évolution de qualité.
POST	avec SP	Depuis quelques années, ce produit connaît une évolution considérable de qualité.

- 24 items expérimentaux de ce type, avec 12 adjectifs différents
- *véritable, mauvais, meilleur, nouveau, redoutable, éventuel, rapide, possible, traditionnel, considérable, regrettable, irrésistible*

2. Méthodologie : Sujets et procédure

Sujets

- 36 sujets
- francophones natifs, entre 18 et 30 ans
- payés pour leur participation

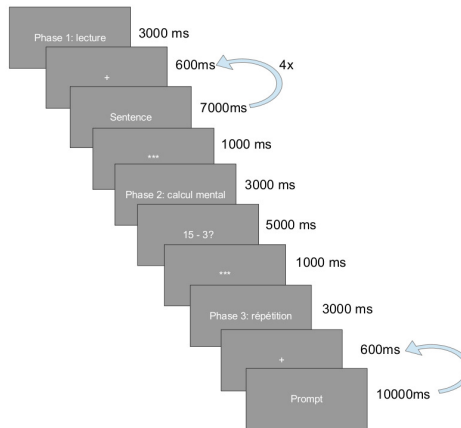
2. Méthodologie : Sujets et procédure

Sujets

- 36 sujets
- francophones natifs, entre 18 et 30 ans
- payés pour leur participation

Procédure

- l'expérience s'est déroulée en séries successives, organisées en 3 phases
- **Phase 1** : Lecture de 4 phrases (1 item expérimental + 3 distracteurs)
- **Phase 2** : Calcul mental
- **Phase 3** : Rappel des 4 phrases (randomisées)
- Chaque sujet a complété 24 séries + 2 séries d'entraînement
- Chaque sujet lit et répète 6 items expérimentaux par condition



2. Méthodologie : Codage et traitement des réponses

- Réponses transcrites et codées (deuxième codage en cours)
- Codage selon les catégories suivantes
 - ▶ Correct (ordre respecté, changements lexicaux acceptés sauf pour l'adjectif)
 - ▶ Inversion
 - ▶ Élément absent
 - ▶ Non réponse
 - ▶ 2 ordres

2. Méthodologie : Codage et traitement des réponses

- Réponses transcrites et codées (deuxième codage en cours)
- Codage selon les catégories suivantes
 - ▶ Correct (ordre respecté, changements lexicaux acceptés sauf pour l'adjectif)
 - ▶ Inversion
 - ▶ Autres

2. Méthodologie : Codage et traitement des réponses

- Réponses transcrites et codées (deuxième codage en cours)
 - Codage selon les catégories suivantes
 - ▶ Correct (ordre respecté, changements lexicaux acceptés sauf pour l'adjectif)
 - ▶ Inversion
 - ▶ Autres
 - Exclusion de résultats
 - ▶ Sujets présentant moins de 30% de réponses utilisables
- 32 sujets restants

2. Méthodologie : Codage et traitement des réponses

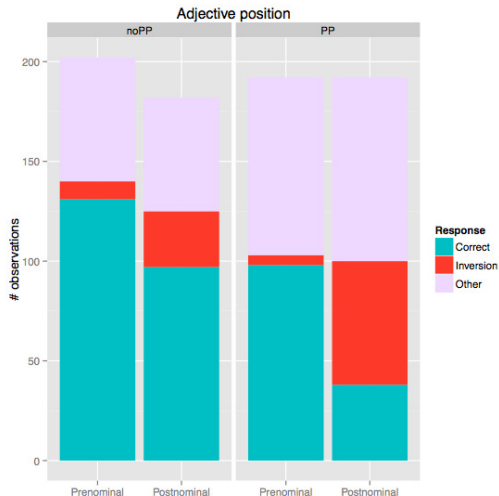
- Réponses transcrites et codées (deuxième codage en cours)
 - Codage selon les catégories suivantes
 - ▶ Correct (ordre respecté, changements lexicaux acceptés sauf pour l'adjectif)
 - ▶ Inversion
 - ▶ Autres
 - Exclusion de résultats
 - ▶ Sujets présentant moins de 30% de réponses utilisables
- 32 sujets restants

Prédiction

Les sujets vont produire plus d'inversions '*postposition vers antéposition*' lorsqu'un SP sera présent dans le SN

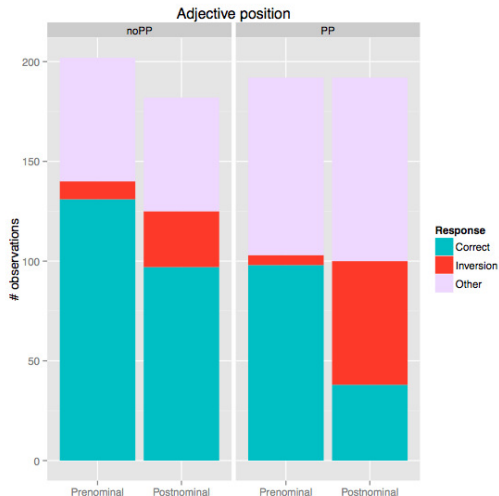
3. Résultats

- proportion de réponses utilisables plus basse pour les SN plus complexes (contenant un SP)



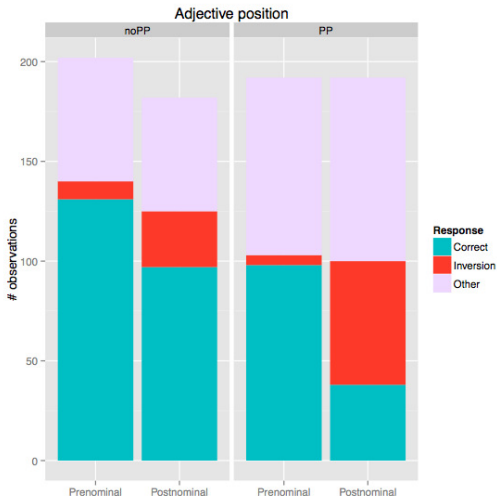
3. Résultats

- proportion de réponses utilisables plus basse pour les SN plus complexes (contenant un SP)
- seulement 6.4% d'inversions '*antéposition vers postposition*'



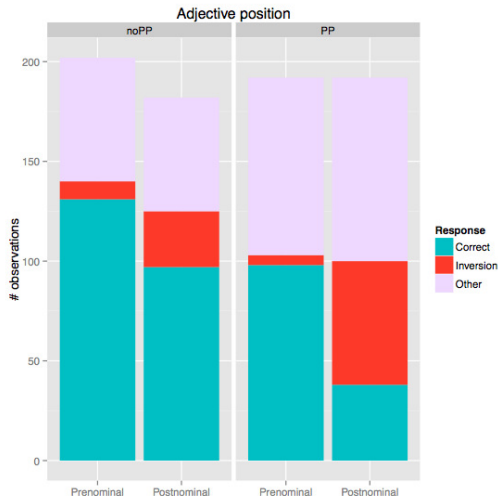
3. Résultats

- proportion de réponses utilisables plus basse pour les SN plus complexes (contenant un SP)
- seulement 6.4% d'inversions '*antéposition vers postposition*'
- 62% d'inversions '*postposition vers antéposition*' avec un SP



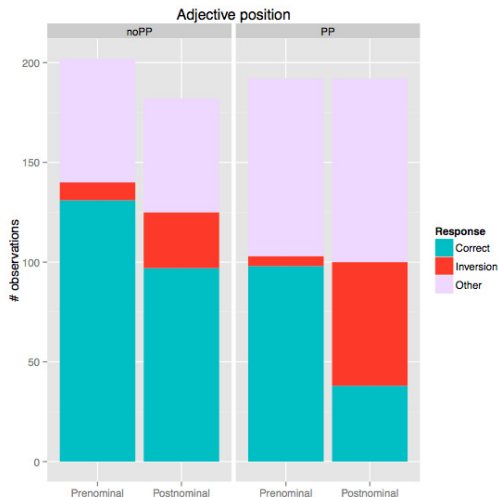
3. Résultats

- proportion de réponses utilisables plus basse pour les SN plus complexes (contenant un SP)
- seulement 6.4% d'inversions '*antéposition vers postposition*'
- 62% d'inversions '*postposition vers antéposition*' avec un SP
- seulement 22.4% sans SP



3. Résultats

- proportion de réponses utilisables plus basse pour les SN plus complexes (contenant un SP)
- seulement 6.4% d'inversions '*antéposition vers postposition*'
- 62% d'inversions '*postposition vers antéposition*' avec un SP
- seulement 22.4% sans SP
- effet de la présence d'un SP sur la position de l'adjectif : interaction entre position de l'adjectif et présence du SP statistiquement significative pour rendre compte des cas d'inversion ($p < .01$, dans un modèle de régression logistique à effets mixtes)



3. Résultats

- Variations importantes des résultats pour certains adjectifs
 - ▶ 9.8% des réponses avec *mauvais* en postposition
 - ▶ 80.8% des réponses avec *considérable* en postposition

3. Résultats

- Variations importantes des résultats pour certains adjectifs
 - ▶ 9.8% des réponses avec *mauvais* en postposition
 - ▶ 80.8% des réponses avec *considérable* en postposition
- Les inversions '*antéposition vers postposition*' concernent des adjectifs longs (3-4 syll) : *redoutable*, *éventuel*, *irrésistible*, *considérable*, *traditionnel*

3. Résultats

- Variations importantes des résultats pour certains adjectifs
 - ▶ 9.8% des réponses avec *mauvais* en postposition
 - ▶ 80.8% des réponses avec *considérable* en postposition
 - Les inversions '*antéposition vers postposition*' concernent des adjectifs longs (3-4 syll) : *redoutable, éventuel, irrésistible, considérable, traditionnel*
- ⇒ Les résultats de l'expérience suggèrent que malgré les comportements particuliers liés au lexique, la présence d'un SP favorise l'antéposition de l'adjectif
- ⇒ Cela confirme les observations faites en corpus

4. Discussion

- Les résultats de l'expérience pourraient s'expliquer à trois niveaux
 - 1 Niveau lexical : il existe des préférences de position plus ou moins fortes pour chaque adjectif

4. Discussion

- Les résultats de l'expérience pourraient s'expliquer à trois niveaux
 - 1 Niveau lexical : il existe des préférences de position plus ou moins fortes pour chaque adjectif
 - 2 Niveau prosodique :
 - **Hypothèse** : placer l'adjectif en antéposition lorsqu'il y a un dépendant post-nominal permet de meilleurs groupements prosodiques (moins de groupes phonologiques, des groupes plus équilibrés et bien calibrés) (Delais-Roussarie, 1996)

4. Discussion

- Les résultats de l'expérience pourraient s'expliquer à trois niveaux
 - 1 Niveau lexical : il existe des préférences de position plus ou moins fortes pour chaque adjectif
 - 2 Niveau prosodique :
 - **Hypothèse** : placer l'adjectif en antéposition lorsqu'il y a un dépendant post-nominal permet de meilleurs groupements prosodiques (moins de groupes phonologiques, des groupes plus équilibrés et bien calibrés) (Delais-Roussarie, 1996)

Items expérimentaux	Co.	Inv.
(un chasseur) ₃ (redoutable) ₃	8	0
(un redoutable) ₄ (chasseur) ₂	3	3
(un chasseur) ₃ (redoutable) ₃ (de cerfs) ₂	3	4
(un redoutable) ₄ (chasseur de cerfs) ₄	6	0

4. Discussion

- Les résultats de l'expérience pourraient s'expliquer à trois niveaux
 - 1 **Niveau lexical** : il existe des préférences de position plus ou moins fortes pour chaque adjectif
 - 2 **Niveau prosodique** :
 - **Hypothèse** : placer l'adjectif en antéposition lorsqu'il y a un dépendant post-nominal permet de meilleurs groupements prosodiques (moins de groupes phonologiques, des groupes plus équilibrés et bien calibrés) (Delais-Roussarie, 1996)

Items expérimentaux	Co.	Inv.
(un chasseur) ₃ (redoutable) ₃	8	0
(un redoutable) ₄ (chasseur) ₂	3	3
(un chasseur) ₃ (redoutable) ₃ (de cerfs) ₂	3	4
(un redoutable) ₄ (chasseur de cerfs) ₄	6	0

- 3 **Niveau syntaxique** :
 - **Hypothèse** : placer un adjectif entre le nom et le SP augmente le coût d'intégration syntaxique du SP pour l'interlocuteur (Gibson, 1998; Demberg et Keller, 2008)

4. Discussion

- Les résultats de l'expérience pourraient s'expliquer à trois niveaux
 - 1 **Niveau lexical** : il existe des préférences de position plus ou moins fortes pour chaque adjectif
 - 2 **Niveau prosodique** :
 - **Hypothèse** : placer l'adjectif en antéposition lorsqu'il y a un dépendant post-nominal permet de meilleurs groupements prosodiques (moins de groupes phonologiques, des groupes plus équilibrés et bien calibrés) (Delais-Roussarie, 1996)

Items expérimentaux	Co.	Inv.
(un chasseur) ₃ (redoutable) ₃	8	0
(un redoutable) ₄ (chasseur) ₂	3	3
(un chasseur) ₃ (redoutable) ₃ (de cerfs) ₂	3	4
(un redoutable) ₄ (chasseur de cerfs) ₄	6	0

- 3 **Niveau syntaxique** :
 - **Hypothèse** : placer un adjectif entre le nom et le SP augmente le coût d'intégration syntaxique du SP pour l'interlocuteur (Gibson, 1998; Demberg et Keller, 2008)
 - Mise en place d'une expérience de *self-paced reading*, dans le but d'observer si les temps de lecture sur le SP sont plus longs lorsque l'adjectif est postposé, ce qui constituerait un argument en faveur de notre hypothèse

- 1 La position de l'adjectif
 - Aspects sémantiques
 - Aspects lexicaux
 - Aspects syntaxiques
 - Combinaisons particulières de N et de A
- 2 Modélisation des données de corpus
 - Méthodologie
 - Observations
 - Modélisation statistique multifactorielle
 - Résultats
- 3 Expérience de rappel de phrases
 - Motivation
 - Méthodologie
 - Résultats
 - Discussion
- 4 Conclusion et perspectives

Conclusion et perspectives

- Une approche quantitative et expérimentale du problème de la position de l'adjectif en français
 - ▶ Modélisation précise du phénomène à partir de données de corpus
 - ▶ Données expérimentales permettant de confirmer et d'expliquer certaines observations faites en corpus

Quelques perspectives

1 Corpus

- ▶ Augmenter les données concernant les adjectifs alternants, notamment à l'oral
- ▶ Décomposer les « préférences » lexicales

2 Expériences

- ▶ Tester les 2 hypothèses formulées pour expliquer l'effet de la présence d'un dépendant post-nominal sur la position de l'adjectif

Références bibliographiques I

- Abeillé, A., Barrier, N., 2004. Enriching a French treebank. In : Proceedings of Language Resources and Evaluation Conference (LREC). Lisbonne.
- Abeillé, A., Godard, D., 1999. La position de l'adjectif épithète en français : le poids des mots. *Recherches Linguistiques de Vincennes* 28, 9–32.
- Abeillé, A., Godard, D., 2000. French word order and lexical weigh. In : Borsley, R. (Ed.), *The Nature and Function of Syntactic Categories*, vol. 32 of *Syntax and Semantics*. Academic Press, New York, pp. 325–358.
- Abeillé, A., Godard, D., 2004. De la légèreté en syntaxe. *Bulletin de la Société de Linguistique de Paris* XCIX, 69–106.
- Agresti, A., 2007. *An Introduction to Categorical Data Analysis*. Wiley.

Références bibliographiques II

- Benzitoun, C., 2013. Adjectifs épithètes alternants en français parlé : premiers résultats. TIPA. Travaux interdisciplinaires sur la parole et le langage 29. [Http ://tipa.revues.org/585](http://tipa.revues.org/585).
- Bouchard, D., 1998. The distribution and interpretation of adjectives in French : A consequence of bare phrase structure. *Probus* 10, 139–183.
- Bresnan, J., Cueni, A., Nikitina, T., Baayen., R. H., 2007. Predicting the dative alternation. In : Boume, G., Kraemer, I., Zwarts, J. (Eds.), *Cognitive Foundations of Interpretation*. Royal Netherlands Academy of Science, Amsterdam.
- Bresnan, J., Ford, M., 2010. Predicting syntax : Processing dative constructions in American and Australian varieties of English. *Language* 86, 186–213.
- Bresnan, J., Hay, J., 2008. Gradient grammar : An effect of animacy on the syntax of *give* in New Zealand and American English. *Lingua* 118, 245–259.

Références bibliographiques III

- Cresti, E., Moneglia, M. (Eds.), 2005. C-ORAL-ROM Integrated Reference Corpora for Spoken Romance Languages. John Benjamins, Amsterdam, USA.
- Delais-Roussarie, E., 1996. Phonological Phrasing and Accentuation in French. In : Nespou, M., Smith, N. (Eds.), *Dam Phonology : HIL Phonology Papers II*. Holland Academic Graphics, pp. 1–38.
- Demberg, V., Keller, F., 2008. Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition* 109, 193–210.
- Forsgren, M., 1978. *La place de l'adjectif épithète en français contemporain, étude quantitative et sémantique*. Almqvist & Wilksell, Stockholm.
- Fox, G., Thuilier, J., 2012. Predicting the position of attributive adjectives in the French. In : Slavkovik, M., Lassiter, D. (Eds.), *New Directions in Logic, Language and Computation. Lecture Notes in Computer Science*, Springer.

Références bibliographiques IV

- Gelman, A., Hill, J., 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.
- Gibson, E., 1998. Linguistic complexity : Locality of syntactic dependencies. *Cognition* 68, 1–76.
- Grevisse, M., Goosse, A., 2007. *Le bon usage*. De Boeck Université, 14ème édition.
- Hawkins, J., 1994. *A Performance Theory of Order and Constituency*. Cambridge University Press, Cambridge.
- Noailly, M., 1999. *L'adjectif en français*. Ophrys, Paris.
- Nølke, H., 1996. Où placer l'adjectif épithète ? Focalisation et modularité. *Langue française* 111, 38–57.
- Thuilier, J., 2012. *Contraintes préférentielles et ordre des mots en français*. Ph.D. thesis, Université Paris 7.

Références bibliographiques V

- Thuilier, J., 2013. Syntaxe du français parlé vs. écrit : le cas de la position de l'adjectif épithète. TIPA. Travaux interdisciplinaires sur la parole et le langage 29. [Http ://tipa.revues.org/1066](http://tipa.revues.org/1066).
- Thuilier, J., Fox, G., Crabbé, B., 2010. Fréquence, longueur et préférences lexicales dans le choix de la position de l'adjectif épithète en français. In : Neveu, F., Toke, V. M., Klingler, T., Durand, J., Mondada, L., Prévost, S. (Eds.), Actes du 2ème Congrès Mondial de Linguistique Française 2010 (CMLF 2010). Nouvelle-Orléans.
- Thuilier, J., Fox, G., Crabbé, B., 2012. Prédire la position des adjectifs épithètes en français. *Linguisticae Investigationes* 35, 28–75.
- Wasow, T., 1997. Remarks on grammatical weight. *Language Variation and Change* 9, 81–105.
- Wasow, T., 2002. *Postverbal Behavior*. CSLI publications.

Références bibliographiques VI

- Waugh, L. R., 1977. *A semantic Analysis of Word Order : Position of the Adjective in French*. E. J. Brill, Leiden.
- Wilmet, M., 1981. La place de l'épithète qualificative en français contemporain : étude grammaticale et stylistique. *Revue de linguistique romane* 45, 17–73.

Homonymes alternants dans le corpus

(24) ancien

- a. il y en a trois l'un dans l'autre hein il y a l'**ancien** cimetière si tu veux c'est le central (CORAL)
- b. et puis j' avais acheté beaucoup de livres déjà quand j'avais les **anciens** appareils le Elgi (CORAL)

(25) propre

- a. on s'est très vite rendu compte qu'il y avait aucune volonté nationale pour maintenir le pays ensemble c'est-à-dire que chacun des chacune des communautés faisait passer son intérêt **propre** avant l'intérêt national (CORAL)
- b. on les attache sur nos **propres** maillots ou soit des maillots qui ont été donnés pour pour les foulées de Poitiers (CORAL)

Homonymes alternants dans le corpus

(26) pure

- a. c' est la folie de la folie **pure** et simple (CORAL)
- b. ça a été un système coopératif de **pure** autogestion qui a complètement fonctionné (CORAL)

(27) simple

- a. pour que ce travail ne soit pas une **simple** mesure supplémentaire ils rendront compte de leurs investigations aux élus du conseil général tous les six mois (CORAL)
- b. c' est la folie de la folie pure et **simple** (CORAL)

(28) seul

- a. Pour les voitures particulières **seules**, la situation est un peu meilleure (FTB)
- b. [...] la création d' une franchise générale pour les deux premières ruptures de travail pourrait ne pas être limitée aux **seules** très petites entreprises. (FTB)

Proportions d'inversions par adjectif

ANT	correct	inversion	POST	correct	inversion
considérable	55.6%	44.4%	considérable	100%	0
éventuel	94.7	5.3%	éventuel	33.3 %	66.7%
irrésistible	85.7	14.3%	irrésistible	88.2%	11.8%
mauvais	100%	0	mauvais	20%	80%
meilleur	100%	0	meilleur	61.1%	38.9%
nouveau	100%	0	nouveau	52.2%	47.8%
possible	100%	0	possible	80%	20%
rapide	100%	0	rapide	58.8%	41.2%
redoutable	87.9%	12.1%	redoutable	56.3%	43.7%
regrettable	100%	0	regrettable	62.5%	37.5%
traditionnel	90.5%	9.5%	traditionnel	100%	0
véritable	100%	0	véritable	52.6%	47.4%