

# Tests expérimentaux

Quatre expérimentations ont été réalisées. La première était un pré-test de compréhension/production et avait pour objectif de vérifier l'hypothèse générale sous-jacente aux plans expérimentaux (3.2.1) : les propriétés verbales d'une construction comitative influent sur les enchaînements pronominaux dont peuvent faire l'objet ses arguments. Le pré-test valide notre hypothèse générale et montre, en plus, que le facteur « genre du référent en tête de phrase » détermine des patterns de réponse distincts. Le plan d'expérience des tests « on line » suivants est donc scindé en deux modules, l'un où le genre du référent en tête est féminin (FFirst) et l'autre où le genre du référent en tête est masculin (MFirst). Le matériel de ces tests est élaboré en complétant les items du pré-test par une phrase qui reprend soit le référent masculin (*Il*), soit les deux référents dispersés de la construction comitative (*Ils*) ; une troisième phrase suit la phrase cible pour éviter que les processus de résolution du pronom, susceptibles de se prolonger jusqu'en fin de phrase cible, ne coïncident avec la fin du texte. La phrase qui supporte les conditions expérimentales a été désignée par le terme de *phrase amorce*. La phrase qui suit dont les temps de traitement sont enregistrés, est désignée comme *la phrase cible*. Les items de test, au nombre de 24, sont mélangés de façon aléatoire avec 24 distracteurs par le logiciel E-prime version 2 professionnelle (Psychology Software Tools, Inc.). Ce test ne donne pas de résultats concluants. Les items de test sont donc modifiés 1) en ajoutant une phrase d'introduction avant la phrase amorce, 2) en segmentant les phrases en trois parties de façon à enregistrer les temps de lecture dans le segment qui contient le pronom, puis dans le segment proche et enfin dans le dernier segment. Le deuxième test valide la plupart de nos hypothèses. Un dernier test reprend les mêmes items en transformant la construction comitative en une construction intransitive équivalente, c'est-à-dire en composant un objet référentiel complexe avec les référents dispersés de la construction comitative.

## 4.1 Points méthodologiques

### 4.1.1 Procédures de test et conditions de passation

Deux procédures de test ont été utilisées dans ce travail : la technique de complétude de phrases pour le pré-test et la technique de l'APS (Auto Présentation Segmentée) pour les tests en temps réel

(« on line »). Le pré-test a été réalisé par petits groupes d'étudiants à l'université Sorbonne-Nouvelle-Paris 3 : les étudiants étaient informés de l'expérience dans le cadre des enseignements qu'ils suivaient. Les tests se déroulaient en dehors des cours<sup>42</sup> : seuls les volontaires y participaient. Les tests en temps réel avec la technique de l'APS ont été réalisés avec le logiciel E-prime version 2 professionnelle (*Psychology Software Tools, Inc.*) : il y en a eu trois. Dans cette méthodologie, les items sont affichés sur un écran d'ordinateur, phrase par phrase ou segment par segment (quelquefois mot par mot). Les participants au test avaient reçu la consigne de progresser dans la lecture de l'item à leur rythme en appuyant sur la barre d'espace. Une question de compréhension était posée à la fin de chaque items. Cette question amenait une réponse binaire en oui ou non à laquelle le participant répondait en appuyant sur le chiffre un (1) du clavier ou la chiffre zéro (0). Les items de test étaient précédés de deux items d'entraînement pour familiariser le participant au protocole de lecture. Les temps de lecture de chaque segment étaient enregistrés ainsi que les temps de réponse à la question et la réponse elle-même. Dans nos expérience en temps réel, les participants passaient le test un par un, dans une petite salle mise à notre disposition par l'UMR 7172 de l'université Sorbonne Nouvelle Paris 3<sup>43</sup> dirigée par M. Michel Collot puis par Alain Schaffner. Ils étaient recrutés directement par affiches ou par une présentation personnalisée à la fin de leurs cours. A la fin du test, ils recevaient une entrée au Forum des images aux Halles. Les participants étaient interrogés 1) sur les langues qu'ils parlaient et en cas de bilinguisme, sur la langue qu'ils considéraient comme dominante, 2) sur des problèmes de dyslexie éventuels qu'ils auraient rencontrés dans leur scolarité, 3) sur le cursus suivi à l'université Sorbonne Nouvelle. Comme précisé à chaque participant, les réponses à ces questions n'étaient pas nominatives.

## 4.1.2 Analyses statistiques et modèles utilisés

Les analyses statistiques ont été réalisées avec le logiciel R<sup>44</sup>. Elles ont portées sur le nombre de reprises au singulier et au pluriel dans le pré-test et sur les temps de lecture pour les tests en temps réel. Nous exposons ci-dessous les étapes suivies lors des analyses statistiques des temps de lecture.

---

<sup>42</sup> A trois exceptions près : M. Charolles, S. Fleury et C. Puech ont accepté d'écourter leurs enseignements de quelques dizaines de minutes pour que leurs étudiants puissent y participer.

<sup>43</sup> THALIM : Théorie et Histoire des Arts et des Littératures de la Modernité (UMR 7172)

<sup>44</sup> R Development Core Team. (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing. [Computer program (version 3.3.0)]. Accessible sur : <http://www.r-project.org>.

La première étape consistait à trier les temps de lecture. Les taux de bonnes réponses de chaque participant étaient examinés et seules les données des participants ayant au moins deux tiers de réponses correctes étaient conservées. De plus, pour chaque participant, les temps supérieurs de deux écart-type ou inférieurs de deux écart-type par rapport à sa moyenne étaient éliminés.

Les analyses utilisaient les fonctions *glmer* (loi binomiale pour le nombre de reprises plurielles dans le pré-test) et *lmer* de R (loi normale pour les temps de réponse). On cherchait d'abord des covariations éventuelles entre les temps de lecture et 1) le nombre de caractères du texte (ou segment) correspondant à ce temps de lecture, 2) la fréquence du verbe de la phrase en cours de lecture et 3) la position ordinale de la phrase dans le test en cours. Lorsque l'une de ces variables covariait avec les temps de lecture analysés (ce qui arrivait systématiquement pour la position ordinale de l'item), on faisait une régression des temps de lecture sur la ou les variables covariantes (Clifton et Ferreira, 2012) avec le modèle linéaire mixte de R : les résidus issus de cette régression faisaient l'objet d'une analyse linéaire mixte sur les facteurs du plan d'expérience, avec les variables aléatoires sujets et items (termes d'erreur). Les termes d'erreur pouvaient comporter des effets aléatoires moyens dus aux sujets et aux items ou des effets aléatoires moyens augmentés d'effets aléatoires dus à des incertitudes de modélisation (calculs sur la pente sur la droite de régression). Dans toutes nos analyses, il n'y avait pas de différence entre les résultats obtenus avec les termes d'erreur qui tenaient compte des incertitudes de la modélisation et ceux obtenus avec les termes d'erreurs moyens. Le modèle le plus simple a donc été utilisé (Baayen, 2008). Le résultat de l'analyse avec le modèle linéaire mixte donnant des effets de significativité peu fiables du fait des incertitudes sur les calculs des degrés de liberté, il fallait confirmer l'effet d'un facteur apparaissant comme significatif après cette première analyse. Une deuxième analyse avec le modèle linéaire mixte était effectuée sans ce facteur : puis une comparaison (Anova) entre les résultats de la première et de la deuxième analyse était réalisée : si une différence était détectée (chisquare <0.05), on en déduisait que le facteur supprimé de la deuxième analyse était bien significatif.

Dans chaque test, les données ont été analysées pour vérifier que le facteur « groupe expérimental » n'était pas significatif.

### 4.1.3 Elaboration du matériel de test

Les verbes symétriques ne sont pas marqués syntaxiquement et peuvent donc être contrastés avec des verbes mixtes dans les contextes phrastiques identiques comme le montrent les exemples (210a-d) et (211a-b).

Les verbes de nos expérimentations ont été sélectionnés à partir de la liste de verbes symétriques établie par A. Borillo (1971)<sup>45</sup>. Ils devaient répondre à certains critères : avoir une construction transitive indirecte en *avec*, mais aussi ne pas être pronominaux (*se disputer*, *se concerter* ont été écartés, par exemple). Le dictionnaire électronique des verbes français de Dubois et Dubois-Charlier (1997), indique que les verbes sélectionnés partagent un schéma verbal de base que nous présentons en annexe 2. La catégorie linguistique des arguments utilisés dans les items de test était celle des noms propres : des prénoms de genre féminin et masculin. Cette catégorie a été choisie parce qu'elle n'interagit pas avec le sémantisme du verbe, permettant au schéma verbal et aux rôles thématiques correspondants d'être testés sans biais. Les prénoms ont été choisis de façon à ce qu'ils désignent un genre sans ambiguïté ; ils ne devaient pas faire plus de deux syllabes pour homogénéiser la longueur des phrases testées.

Les différentes conditions expérimentales étaient contrastées dans une structure phrastique qui 1) évitait les collusions entre prédicat et syntagme prépositionnel en insérant un adverbe ou un circonstanciel entre eux, 2) se terminait toujours par un adverbe ou un circonstanciel de façon à éviter une trop grande proximité entre le référent régi et la reprise pronominale qui suivait. La phrase qui supportait les conditions expérimentales a été désignée par le terme de *phrase amorce*. La phrase qui suivait et dont le traitement était testé était désignée comme *la phrase cible*. Phrases amorces et phrases cibles étaient calibrées : le nombre de caractères des phrases amorces était compris entre 67 et 72 et le nombre de caractère des phrases cibles était compris entre 52 et 60.

Les verbes non symétriques choisis étaient soit des verbes d'action (comme *grimper*, *bosser*, *chercher*, *ricaner*, *contrôler*, *militer*, etc...) sélectionnés pour leur capacité à regrouper les protagonistes dans une action collective, soit des verbes de perception (*geler*, *pleurer*, *angoisser*, *trembler*, *se débrouiller*, *s'ennuyer*, etc...), sélectionnés pour leur caractère individualisant qui ne s'applique qu'à l'individu du fait de particularités qui lui appartiennent en propre (Clifton et Frazier, 2012). Les verbes d'action, associés à des informations de lieu et de temps qui participent à

---

<sup>45</sup> Je lui renouvelle mes remerciements.

l'élaboration d'une base d'association commune (Eschenbach et al., 1989) suscitent plutôt une interprétation collective de la prédication : la préposition *avec* est alors équivalente à une structure coordonnante. Dans certaines prédications, cependant, seul le contexte permettait de trancher entre interprétation collective et distributive<sup>46</sup>. Par exemple le verbe *transpirer* est à priori un distributif mais l'exemple (213) induit une base d'association commune et une interprétation collective. De même l'exemple (214). En fait, le principal critère de classement des prédicats en collectifs ou distributifs, était le sens de la préposition *avec* (cf. 2.1.2) : chaque fois qu'elle introduisait une nuance de causalité ou de non engagement du référent régi, le prédicat était catégorisé comme distributif. Dans une phrase amorce comme (215), Miguel peut être son professeur ou être un étudiant qui progresse également en portugais à l'heure du déjeuner, auquel cas les rôles thématiques de l'un et de l'autre diffèrent : il y a ambiguïté et de ce fait le prédicat *progresser en Portugais avec quelqu'un à l'heure du déjeuner* est classé comme un distributif.

(213) *Clément transpire sous un soleil torride avec Elise en plein Sahara.*

(214) *Louise tremble sous une pluie diluvienne avec Rémi faute d'abri*

(215) *Sandra progresse en Portugais avec Miguel à l'heure du déjeuner.*

A chaque nouvelle expérimentation, le matériel de test a été amélioré dans la mesure du possible : on a éliminé des items les verbes symétriques transitifs comme *troquer*, *échanger* qui ont été utilisés dans le pré-test.

## 4.2 Pré-test

Ce pré-test est conçu pour vérifier l'hypothèse générale d'un effet des propriétés prédictives du noyau verbal sur le regroupement ultérieur d'arguments dispersés. Les constructions testées sont des constructions comitatives. Les résultats des expériences antérieures rapportées en 3.1.1 permettent d'affirmer que les prédicats mixtes favorisent non pas la formation d'une pluralité à partir des référents sujet et régi, mais la continuité thématique qui consiste à reprendre le sujet (Gordon et al., 1993 ; Garrod et al., 1994). Toutefois, les propriétés de collectivité ou de distributivité des prédicats n'ont pas été contrôlées dans ces expériences. Concernant les prédicats

---

<sup>46</sup> Généralement, l'intuition n'est pas une bonne base de sélection du matériel expérimental. Néanmoins, les résultats des tests permettent de valider les propriétés collectives des prédicats retenus dans cette catégorie.

symétriques, Bianco & Schnedecker (2000) ont montré que les taux de reprises plurielles étaient équivalents dans les constructions intransitives et transitives indirectes de ces prédicats.

## 4.2.1 Plan expérimental

### A- Matériel et Facteurs expérimentaux

Le pré-test est conçu pour vérifier l'hypothèse générale d'un effet des propriétés prédicatives du noyau verbal sur le regroupement ultérieur d'arguments dispersés. Les constructions testées sont des constructions comitatives. Les résultats des études expérimentales rapportées en 3.1.1 permettent d'affirmer que les prédicats mixtes favorisent non pas la formation d'une pluralité à partir des référents sujet et régi, mais la continuité thématique qui consiste à reprendre le sujet (Gordon et al., 1993 ; Garrod et al., 1994). Toutefois, les propriétés de collectivité ou de distributivité des prédicats n'ont pas été contrôlées dans ces expériences. Les prédicats symétriques devraient, se prêter plus que les prédicats mixtes à un regroupement des référents dispersés (Bianco & Schnedecker, 2000)

Le pré-test est un test de compréhension-production. Notre propos est de contraster les taux de reprises plurielles après des prédicats symétriques et mixtes dans des constructions syntaxiques discontinues où le genre du sujet syntaxique est manipulé de manière à moduler le focus discursif. Par ailleurs, les verbes mixtes pouvant recevoir des interprétations collectives dans les constructions comitatives, comme l'a montré l'étude du syntagme comitatif en 2.1.2, la moitié de nos items mixtes sont à interprétation collective, l'autre moitié à interprétation distributive. Sont donc comparées les reprises du sujet, de l'argument régi et du pluriel qu'ils forment, après des constructions comitatives dont les prédicats sont symétriques et distributifs ou collectifs. Selon l'interprétation du prédicat mixte, le prédicat symétrique est comparé à un prédicat distributif ou collectif.

Le plan expérimental comporte donc deux facteurs croisés, qu'un facteur partage en deux groupes (schéma split-plot). Les facteurs croisés sont le facteur « Genre du prénom en tête de prédication » (c'est lui qui détermine si le pronom de reprise *Il(s)*, est en position proéminente ou pas), et le facteur « type de prédicat » ; le facteur qui partage le plan expérimental croisé en deux, est le facteur « Groupe Contraste ».

Dans le pré-test, vingt-huit prédictions symétriques sont comparées à vingt-huit prédictions mixtes, selon quatre conditions expérimentales dans chaque groupe de contraste.

Un exemple d'item de test du groupe contraste Symétrique/Distributif est présenté dans les quatre conditions expérimentales en (216a-d) : le genre du sujet est codé F pour féminin et M pour masculin, la classe de prédicat est codée S pour Symétrique et N pour Non symétrique.

**Condition 1** : Le genre du sujet syntaxique est féminin, celui du référent régi est masculin et le noyau prédicatif est symétrique : condition prénom Féminin en premier et prédicat Symétrique, soit FFirst, S, S/D.

(216a) *Audrey cohabite à Belle Ile avec Didier en pleine mauvaise saison.*

**Condition 2** : Le genre du sujet syntaxique est masculin, celui du référent régi est féminin et le noyau prédicatif est symétrique : condition prénom Masculin en premier et prédicat Symétrique, soit MFirst, S, S/D.

(216b) *Didier cohabite à Belle Ile avec Audrey en pleine mauvaise saison.*

**Condition 3** : Le genre du sujet syntaxique est féminin, celui du référent régi est masculin et le noyau prédicatif est mixte : condition prénom Féminin en premier et prédicat Non symétrique, soit FFirst, N, S/D.

(216c) *Audrey gèle à Belle Ile avec Didier en pleine mauvaise saison.*

**Condition 4** : Le genre du sujet syntaxique est masculin, celui du référent régi est féminin et le noyau prédicatif est mixte : condition prénom Masculin en premier et prédicat Non symétrique, soit MFirst, N, S/D.

(216d) *Didier gèle à Belle Ile avec Audrey en pleine mauvaise saison.*

Un exemple d'item de test du groupe contraste Symétrique/Collectif est présenté dans les quatre conditions expérimentales en (217a-d) : les modalités des facteurs « genre du sujet » « classe du prédicat » sont codées de la même façon.

**Condition 1** : Le genre du sujet syntaxique est féminin, celui du référent régi est masculin et le noyau prédicatif est symétrique : condition prénom Féminin en premier et prédicat Symétrique, soit FFirst, S, S/C,

(217a) *Brenda sympathise dans le métro avec Samy en un dixième de seconde.*

**Condition 2** : Le genre du sujet syntaxique est masculin, celui du référent régi est féminin et le noyau prédicatif est symétrique : condition prénom Masculin en premier et prédicat symétrique, soit MFirst, S, S/C,

(217b) *Samy sympathise dans le métro avec Brenda en un dixième de seconde.*

**Condition 3** : Le genre du sujet syntaxique est féminin, celui du référent régi est masculin et le noyau prédicatif est mixte : condition prénom Féminin en premier et prédicat Non symétrique, soit FFirst, N, S/C,

(217c) *Brenda grimpe dans le métro avec Samy en un dixième de seconde.*

**Condition 4** : Le genre du sujet syntaxique est masculin, celui du référent régi est féminin et le noyau prédicatif est mixte (condition prénom Masculin en premier et prédicat Non symétrique, soit MFirst, N, S/C.

(217d) *Samy grimpe dans le métro avec Brenda en un dixième de seconde.*

## B- Procédure

Cette première expérimentation est un test de complétude de phrases contraint : les participants ont pour consigne de trouver une suite plausible à chaque phrase en démarrant avec un pronom qui leur est donné oralement, et qui est toujours phonétiquement « il » (mais qui peut être interprété comme « ils » du fait de l'homophonie des deux pronoms).

Les quatre versions de chacune des 28 phrases sont réparties sur quatre groupes (carnets de test A, B, C, D) en contrebalançant les items de façon à ce qu'il y ait le même nombre d'items par condition dans chaque groupe (sept items) et qu'un même item n'apparaisse dans un groupe que dans une version (voir le plan expérimental par condition en annexe 3).<sup>47</sup> Entre les items testés dans chaque groupe, viennent s'intercaler dix items à une seule modalité (type de verbe symétrique ou

---

<sup>47</sup> Le nombre d'items par contraste n'est pas équilibré : dans le carnet A, par exemple, il y a 3 items de la condition « genre Féminin/verbe Symétrique » dans le contraste S/NC et 4 dans le contraste S/C ; il y a 4 items de la condition « genre Féminin/verbe Non symétrique » dans le contraste S/NC et 3 dans le contraste S/C ; il y a 3 items de la condition « genre Masculin/verbe Symétrique » dans le contraste S/NC et 4 dans le contraste S/C ; il y a 4 items de la condition « genre Masculin/verbe Non symétrique » dans le contraste S/N et 3 dans le contraste S/C.

pas) dont le sujet syntaxique est une paire de référents coordonnés et dix autres items également à une seule modalité (type de verbe symétrique ou pas), où les deux référents dispersés sont de genre masculin. Contrairement aux items expérimentaux, les deux versions de ces items supplémentaires sont présents dans chaque groupe, avec cependant des noms propres différents dans leur version symétrique et leur version non symétrique. Entre ces 48 items, viennent aussi s'intercaler dix textes distracteurs ou fillers. L'ordre de présentation des 58 items est un ordre pseudo-aléatoire, le même dans chaque groupe (voir matériel complet et l'ordre d'occurrence des items en annexe 3).

Une brève présentation du test et de la tâche à accomplir était donnée oralement aux participants : il était notamment précisé que chacune des phrases inscrites sur les pages du carnet seraient lues oralement par l'expérimentateur qui les ferait suivre d'un pronom : il leur était demandé d'enchaîner sur une suite plausible en reprenant le pronom donné oralement. Ce pronom qui ne figurait pas dans les carnets était toujours le pronom personnel masculin *Il(s)*.

## C- Hypothèses

On s'attend à ce que le genre du pronom sujet en tête de phrase amorce influence les reprises pronominales et notamment les taux de reprises plurielles.

Le pronom de reprise *Il* est du même genre que le sujet en (216b), (216d), (217b) et (217d) et de même genre que l'argument régi en (216a), (216c), (217a) et (217c).

Le référent masculin en tête de prédication (conditions (216b), (216d), (217b) et (217d)), est proéminent : on s'attend à des taux de reprises de ce référent importants, lorsqu'il est argument d'un prédicat distributif (Hielscher et Müsseler, 1990) et moins importants lorsqu'il est argument d'un prédicat symétrique ou collectif. Le taux de reprises du référent proéminent devrait toutefois être plus important pour les prédictions symétriques que pour les prédictions collectives, étant donné l'individuation qu'opèrent les prédicats symétriques entre ses arguments (Patson & Ferreira, 2009) et les caractéristiques formelles de ces prédicats étudiées en 1.3.

Le référent féminin en tête de prédication (conditions (216a), (216c), (217a) et (217c)), au contraire, nécessite d'extraire le référent du syntagme prépositionnel pour enchaîner sur une reprise de ce référent : on s'attend donc à ce que les taux de reprises plurielles soient plus importants conformément aux résultats Hielscher et Müsseler (1990), surtout lorsque la prédication est collective ou symétrique. La question est de savoir si les taux de reprises plurielles seront plus importants après les prédicats collectifs relativement aux prédicats symétriques.

Nous nous référons à la petite étude de corpus conduite sur le verbe *bavarder* (2.3.1), pour prédire que les taux de reprises plurielles seront plus importants après les prédicats collectifs qu'après les prédicats symétriques. Cette étude, en effet a montré que les arguments pluriels du verbe *bavarder* ne donnaient lieu à leur reprise intégrale que dans 22% des cas : dans 40,5% des cas, les reprises de l'argument pluriel étaient partielles. En comparaison, le verbe *parler*, qui est mixte, présentait le même taux de reprises partielles mais un taux de reprises intégrales plus élevé : 30%. Ce pattern d'enchaînement donne à penser que les prédicats à interprétation collective favoriseront plus le regroupement d'arguments que les symétriques. De plus, la section 1.3 sur les caractéristiques formelles des prédicats a montré que les arguments d'un prédicat symétrique ont un statut complexe. Et Patson et Ferreira (2009) ont trouvé expérimentalement que les membres d'un pluriel coordonné entrent directement en composition avec un prédicat symétrique : ils sont donc plus visibles avec cette catégorie de prédicats qu'avec des prédicats collectifs. Le regroupement des référents dispersés suite à des constructions comitatives symétriques et collectives ne s'opèrerait pas sur les mêmes bases : relation d'équivalence entre arguments pour les prédicats symétriques et rôle coordonnant de la préposition *avec* pour les prédicats à interprétation collective.

## D- Codage

Les reprises produites par les participants ont été codées « P » lorsqu'elles concernent les deux référents de la phrase amorce, y compris lorsqu'un prédicat de médiation est utilisé comme en (218) ; ou lorsqu'ils sont repris dans un complément comme en (219).

(218) *Il semblerait qu'ils soient différents des autres.*

(219) *Il flotte autour d'eux une atmosphère enjouée.*

Les reprises produites par les participants ont été codées « S » lorsqu'elles ne concernent qu'un référent de la phrase amorce, y compris dans les énoncés qui indiquent une situation commune comme en (220).

(220) *Il partage la maison avec elle.*

Les reprises produites par les participants sont codées « A » lorsqu'elles reprennent le référent féminin sous une forme indirecte comme dans *Il est impossible de la faire taire* ; ou lorsque le participant n'a pas réussi à imaginer une suite à un item.

## E- Conditions de passation

Quatre-vingt-cinq étudiants de Paris-Sorbonne, qui recevaient une entrée au forum des images des Halles en échange de leur participation, ont passé ce test. Ils étaient répartis en quatre groupes : vingt-cinq dans le groupe 1, vingt dans le groupe 2, vingt-et-un dans le groupe 3 et dix-neuf dans le groupe 4. Un carnet (de 6,75 cm x 21 cm) était distribué à chaque participant.

Les passations s'effectuaient par petits groupes d'une dizaine de personnes en général, avec les exceptions qui ont été signalées en 4.1.1.

### **4.2.2 Analyses et résultats**

Le test du Khi<sup>2</sup> permet de calculer les taux de réponses au pluriel et au singulier en fonction des conditions ainsi que la significativité des différences trouvées (calcul des résidus).

Le modèle linéaire mixte généralisé (*glmer* suivant la loi binomiale du logiciel R) est utilisé pour analyser l'impact des facteurs expérimentaux et leurs possibles interactions sur le nombre de reprises plurielles, les trois facteurs expérimentaux étant le facteur « Genre » du prénom en tête de phrase amorce (Féminin ou Masculin), le facteur « Verbe » de la phrase amorce (Symétrique ou Non symétrique) et le facteur « groupe de Contraste » (Symétrique/Collectif ou Symétrique/Distributif).

- **ANALYSE PRINCIPALE**

50,29% de l'ensemble des items test donnent lieu à des reprises plurielles et 43,15% à des reprises au singulier ; 6,55% ne correspondent ni à des reprises pronominales au masculin pluriel, ni à des reprises pronominales au masculin singulier (elles ont été codées « A »).

#### **Facteur «Genre»**

Les analyses indiquent un effet significatif du facteur «Genre» (Chisquare=99.49, df=2, p<0.001) : Tableau 8.

Conditions	NP de tête Féminin	NP de tête masculin
Pluriel	58,24	42,35
Singulier	33,45	52,86
Autre	8,32	4,79

Tableau 8 : Pourcentages des reprises calculés sur le total des réponses de chaque condition<sup>48</sup> dans le pré-test

Comme attendu, les reprises plurielles sont significativement (résidus du Chisquare à 3,86) plus nombreuses lorsque le référent en tête de phrase amorce est féminin plutôt que masculin : 58 % du total des réponses dans la condition « NP féminin en tête » sont au pluriel contre 42% du total des réponses dans la condition « NP masculin ». Inversement, les reprises au singulier sont significativement (résidus du Chisquare de 5,10) plus nombreuses lorsque le référent en tête de phrase amorce est masculin plutôt que féminin : 53 % du total des réponses dans la condition « NP masculin en tête » sont des reprises au singulier contre 33,5% du total des réponses dans la condition « NP féminin en tête ».

Les taux de reprises plurielles obtenus après les constructions comitatives dans le pré-test sont dans la « fourchette » des taux obtenus dans les études antérieures (Tableau 9).

% de reprises plurielles	Constructions plurielles	Constructions comitatives
Sanford et Lockhart (1990)	55	30
Moxey et al. (2004)	40	40
Moxey et al. (2011) : personnages non congruents	60	49
Hielscher et Müsseler (1990)	92	47 (en moyenne) 28 : personnage Féminin en tête 66 : personnage Masculin en tête
Bianco & Schnedecker (2000)	73 (en moyenne)	52 (en moyenne)
Argenti (2017) : pré-test	-	50 (en moyenne) 58 : personnage Féminin en tête 42 : personnage Masculin en tête

Tableau 9 : Pourcentages de reprises plurielles obtenus dans des tâches de complétion de phrases dans différentes études contrastant les constructions plurielle et comitative

<sup>48</sup> Nota bene : les taux rapportés le seront toujours par rapport au total des réponses dans la condition examinée dans la suite de l'analyse.

La comparaison avec les résultats d’Hielscher et Müsseler (1990) montre que les taux de reprises plurielles après des constructions comitatives où les référents de même genre que le pronom de reprise est en tête de prédication, sont plus élevés dans le pré-test (42%) que dans le test de ces auteurs (28%) : en revanche, les taux de reprises plurielles lorsque le référent en tête de prédication est de genre différent de celui du pronom de reprise, sont moins élevés dans le pré-test (58%) que dans le test de ces auteurs (66%). Cette différence est compréhensible étant donné que nos items de test comprennent des prédicats symétriques qui sont susceptibles de favoriser les reprises plurielles, malgré la présence d’un référent saillant en tête de prédication ; mais aussi susceptibles de favoriser les reprises du référent régi peu saillant (cas des NP féminin en tête).

Ces résultats suggèrent donc que 1) la formation d’un pluriel à partir des référents dispersés des constructions discontinues élaborées autour des prédicats sélectionnés dont une moitié est symétrique, un quart est à interprétation collective et un autre quart à interprétation distributive, concurrence la reprise du NP sujet en tête de prédication et 2) l’extraction d’un référent régi par le comitatif est une opération plus coûteuse que la formation d’un pluriel mais tout à fait accessible. Les analyses complémentaires vont apporter des précisions sur la façon dont ces taux globaux de réponses se répartissent dans les groupes de contraste.

### **Interaction des facteurs «Verbe» et «Contraste»**

On trouve une interaction significative entre les facteurs «Verbe» et «Contraste» (Chisquare=47.47, df=2,  $p < 0.001$ ) : Tableau 10 et Figure 1.

Conditions %	«Verbe» NM Groupe S/C	«Verbe» NM Groupe S/D	«Verbe» M Groupe S/C	«Verbe» M Groupe S/D
Pluriel	60,50	39,50	49,58	51,60
Singulier	32,44	54,12	43,53	42,52
Autre	7,06	6,39	6,89	5,88

Tableau 10 : Pourcentages de réponses calculés sur le total des réponses de chaque condition dans le pré-test

Dans le groupe de contraste Symétrique/Collectif, il y a 60,5% de reprises plurielles lorsque le verbe n’est pas symétrique et 49,5% lorsque le verbe est symétrique. Dans le groupe de contraste Symétrique/Distributif, il y a 39,5% de reprises plurielles lorsque le verbe n’est pas symétrique et 51,6% lorsque le verbe est symétrique.

Ces résultats montrent que 1) les **taux de reprises plurielles qui suivent les phrases amorces symétriques**, sont significativement différents des taux de reprises plurielles qui suivent les phrases amorces mixtes dans les deux groupes de contraste, mais 2) dans le groupe de contraste S/C, ils sont significativement moindres (résidus du Chisquare de 3,51), alors que dans le groupe de contraste S/D, ils sont significativement plus élevés (résidus du Chisquare de 3,71).

Ces résultats indiquent que les taux de reprises plurielles sont constants d'un groupe de contraste à l'autre après les verbes symétriques. Lorsque le genre du référent en tête de phrase amorce est masculin, les taux de reprises plurielles dans les deux groupes de contraste sont légèrement différents : 39% dans le groupe S/C et 45,5% dans le groupe S/D, mais cette différence n'est pas significative.

Pour résumer, les taux de reprises plurielles vont en décroissant selon que les phrases amorces sont interprétées collectivement (60,5%), symétriquement (49,6%/51,6%) ou distributivement (39,5%) :

Les reprises au singulier reflètent en miroir les résultats sur le nombre (grammatical) des reprises : les taux de reprises au singulier qui suivent les phrases amorces symétriques sont en effet significativement différents des taux de reprises au singulier qui suivent les phrases amorces mixtes dans les deux groupes de contraste. Toutefois, dans le groupe de contraste S/C, ils sont significativement plus élevés (43,5% contre 32,5%, résidus du Chisquare 4,07), alors que dans le groupe de contraste S/D, ils sont significativement moindres (42,5% contre 54%, résidus du Chisquare 3,98).

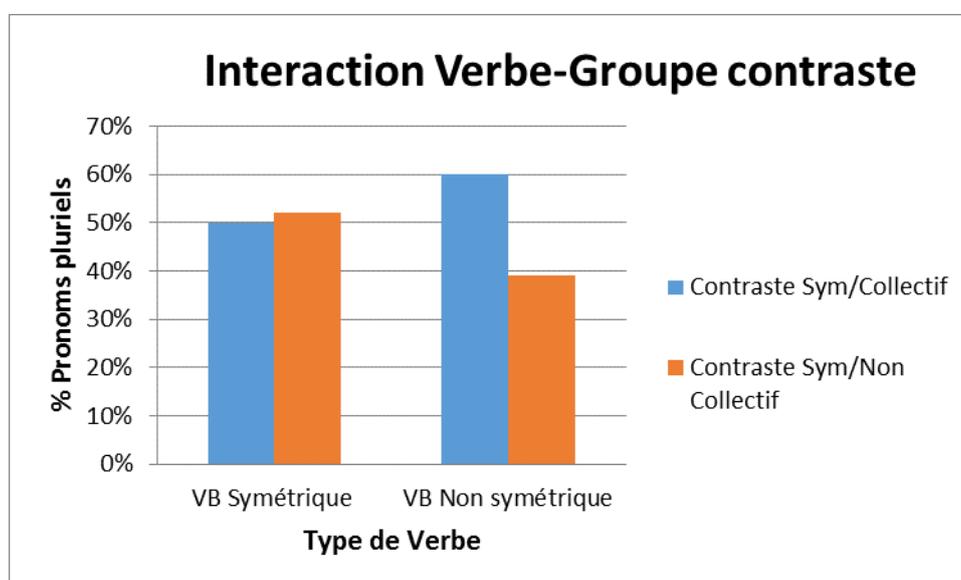


Figure 1 : Pourcentages des reprises plurielles dans le groupe contraste S/C versus celles du groupe contraste S/NC en fonction du type de verbe de la phrase amorce

Ces résultats montrent que le regroupement ultérieur de référents dispersés dans une construction discontinue dépend de plusieurs facteurs, dont la classe collective, distributive ou symétrique de cette prédication. Et plus précisément, ils montrent 1) que les lectures collectives de prédicats mixtes prédisposent plus à la formation d'un pluriel que ne le font les prédications symétriques et 2) que les prédications symétriques prédisposent plus à la formation d'un pluriel que les prédications distributives.

#### • ANALYSES COMPLEMENTAIRES

Le facteur «Genre» étant significatif, des analyses complémentaires ont été réalisées sur les items test qui comportent un NP de genre féminin sujet en tête d'une part, et sur les items de test qui comportent un NP de genre masculin sujet en tête, d'autre part.

#### **ANALYSE COMPLEMENTAIRE : CONDITION « NP féminin en tête »**

##### **Facteur « Contraste »**

Avec le sous-ensemble des phrases amorces commençant par un NP féminin, on trouve un effet principal du groupe «Contraste» (Chisquare=5.67, df=1, p<0.018). Les phrases amorces donnent lieu à plus de reprises plurielles dans le groupe de contraste Symétrique/Collectif (65%) que dans le groupe de contraste Symétrique/Distributif (52%) : la différence est significative (résidus de 2,07). Inversement pour les reprises au singulier : elles sont significativement (résidus de 2,34) plus nombreuses dans le groupe de contraste S/D (39%) que dans le groupe de contraste S/C (28%).

##### **Interaction facteurs «Verbe» et «Contraste»**

Dans le sous-ensemble avec un NP féminin en tête, l'interaction significative (Chisquare=19.66, df=1, p<0.001) des facteurs « Contraste » et «Verbe» montrent les différences significatives suivantes (Tableau 11):

- 1) plus de reprises plurielles après les phrases amorces à lecture collective (69%) qu'après les phrases amorces symétriques (60%) : résidus de 2,56,
- 2) moins de reprises plurielles après les phrases amorces distributives (46%) qu'après les phrases amorces symétriques (57,5%) : résidus de 2,71,
- 3) autant de reprise plurielles après les phrases amorces symétriques dans le groupe de contraste S/C (60%) que dans le groupe de contraste S/D (57,5%),

4) moins de reprises au singulier (23%) après les phrases amorces à lecture collective qu'après les phrases amorces symétriques (32,5%),

5) plus de reprises au singulier après les phrases amorces distributives (44,5%) qu'après les phrases amorces symétriques (33,5%) : résidus de 2,71,

6) autant de reprise au singulier après les phrases amorces symétriques dans le groupe de contraste S/C (32,5%) que dans le groupe de contraste S/D (33,5%),

7) autant de reprises au singulier (44,5%) qu'au pluriel (46%) après les phrases amorces distributives.

Conditions,%	«Verbe» NM Groupe S/C	«Verbe» NM Groupe S/D	«Verbe» M Groupe S/C	«Verbe» M Groupe S/D
Pluriel	69,24	46,05	60,17	57,48
Singulier	23,19	44,37	32,61	33,61
Autre	7,06	10,08	7,73	8,40

Tableau 11 : Pourcentages calculés sur le total des réponses de chaque condition avec le NP féminin en tête

## **ANALYSE COMPLEMENTAIRE : CONDITION « NP masculin en tête »**

### **Facteur « Contraste »**

Avec le sous-ensemble des phrases amorces commençant par un NP masculin, il n'y a pas d'effet significatif (Chisquare=2.4, df=1, p=0.12) du groupe «Contraste» (taux de reprises au pluriel : 45,5% dans le groupe contraste S/C et 39,5% dans le groupe contraste S/D; taux de reprises au singulier : 48% dans le groupe contraste S/C et 58% dans le groupe contraste S/D).

### **Interaction facteurs «Verbe» et «Contraste»**

Dans le sous-ensemble avec un NP masculin en tête, l'interaction significative (Chisquare=27.93, df=1, p<0.001) des facteurs «Contraste» et «Verbe» montre qu'il y a significativement (Tableau 12) :

1) plus de reprises plurielles (52%) après les phrases amorces à lecture collective qu'après les phrases amorces symétriques (39%) : résidus de 2,43,

2) moins de reprises plurielles après les phrases amorces distributives (33%) qu'après les phrases amorces symétriques (45,5%) : résidus de -2,44,

3) autant de reprises plurielles (39%) après les phrases amorces symétriques dans le groupe de contraste S/C que dans le groupe de contraste S/D (45,5%) : différence non significative,

4) moins de reprises au singulier après les phrases amorces à lecture collective (41,5%) qu'après les phrases amorces symétriques (54,5%) : résidus de -2,71,

5) plus de reprises au singulier après les phrases amorces distributives (64%) qu'après les phrases amorces symétriques (51,5%) : résidus de 2,71,

6) autant de reprises au singulier après les phrases amorces symétriques dans le groupe de contraste S/C (54,5%) que dans le groupe de contraste S/D (51,5%).

Conditions %	«Verbe» NM- Groupe S/C	«Verbe» NM- Groupe S/D	«Verbe» M- Groupe S/C	«Verbe» M- Groupe S/D
Pluriel	51,76	32,94	38,99	45,71
Singulier	41,68	63,87	54,45	51,43
Autre	7,06	2,69	6,05	3,36

Tableau 12 : Pourcentages calculés sur le total des réponses de chaque condition avec le NP masculin en tête

Les profils des reprises au pluriel et au singulier lorsque le NP de tête est féminin et lorsqu'il est masculin sont similaires : les pourcentages dépendent, eux, du NP de tête.

Pour résumer ces résultats :

Le facteur « Contraste » n'est significatif que dans la condition NP féminin en tête. Mais l'interaction entre les facteurs «Contraste» et «Verbe», similaire à celle trouvée dans l'analyse principale, est significative quel que soit le genre du NP en tête de phrase.

Les taux de reprises plurielles après les phrases amorces dépendent du type de prédicat et du genre du référent sujet en tête de ces phrases. En particulier :

Lorsque le NP de tête est féminin et que :

- 1) le prédicat est à lecture collective, plus des 2/3 des reprises sont des reprises plurielles,
- 2) le prédicat est distributif, les taux des reprises au pluriel et au singulier sont équivalents,
- 3) le prédicat est symétrique, les taux de reprises plurielles sont majoritaires mais restent inférieurs à ceux des prédicats à lecture collective et supérieurs à ceux des prédicats distributifs.

Lorsque le NP de tête est masculin et que :

- 1) le prédicat est à lecture collective, le taux de reprises plurielles chute de 20 points par rapport à ce qu'il est lorsque le NP féminin est en tête, mais reste supérieur au taux de reprises au singulier,
- 2) le prédicat est distributif, près des 2/3 des reprises sont au singulier,
- 3) le prédicat est symétrique, le taux de reprises au pluriel chute d'un peu plus de 10 points par rapport à ce qu'il est lorsque le NP féminin est en tête et devient moindre que le taux des reprises au singulier : les reprises au singulier sont toutefois moins nombreuses après les prédicats symétriques qu'après les prédicats distributifs.

### 4.2.3 Discussion

Dans ce test de production contraint, on s'attendait à ce que la reprise du référent masculin en tête soit facilitée par rapport à la formation d'un pluriel, ce qui est le cas : il est repris avec un taux de 53% pour l'ensemble des prédicats et un taux de 64% dans le sous-ensemble des prédicats distributifs.

La proéminence discursive dont bénéficie le NP sujet placé en tête de prédication (maintes fois trouvée dans les tests psycholinguistiques que redouble la possibilité d'un parallélisme syntaxique dans le choix qui est laissé aux participants des tests de faire du pronom de reprise le sujet grammatical de l'énoncé suivant et la moindre accessibilité du référent régi par le comitatif expliquent pourquoi le genre du NP de tête conditionne les taux globaux de reprises plurielles.

Dans ce test de production contraint, on s'attendait aussi à ce que les phrases amorces où le NP de genre féminin est en tête, prédisposent plus ou moins à des reprises au pluriel selon la classe de leur prédicat. Ces phrases amorces placent les participants devant le choix, soit de constituer un pluriel avec le référent sujet et le référent régi, soit d'extraire le référent régi, ce qui n'a pas le même coût cognitif pour les prédications amorces symétrique, collectives et distributives. Les résultats vérifient nos hypothèses. Ils montrent que :

- qu'après les phrases amorces à lecture collective, le choix portait préférentiellement sur la constitution d'un pluriel, et cela quel que soit le genre du NP sujet de ces phrases amorces ((217c) et (217d)),
- qu'après les phrases amorces distributives, le choix portait préférentiellement sur la reprise du NP sujet quand il était masculin (216d) et était équitablement réparti entre la formation d'un pluriel et l'extraction du référent régi si le NP sujet était féminin (216c),
- qu'après des phrases amorces symétriques, le choix portait préférentiellement sur la reprise du NP sujet quand il était masculin ((217b) et (216b)) préférentiellement sur la constitution d'un pluriel, si le NP sujet de ces phrases amorces était féminin ((216a) et (217a)).

Les phrases amorces à verbes symétriques, pourtant catalogués comme des verbes pluriels (Kemmer, 1993 ; Moxey et al., 2011 ; Kaup et al. 2002), favorisent moins les reprises plurielles que les phrases amorces à lecture collective dans les constructions discontinues à référents dispersés : ces référents semblent plus individualisés, moins cohésifs, dans les prédications symétriques que dans les prédications à lecture collective, comme le suggèrent Patson et Ferreira (2009).

Dans les prédications à lecture collective, les référents semblent former un groupe qui se comporte comme une seule entité plurielle au sens de Link (1983) et Landmann (1996). Cela confirme le statut de structure coordonnante du syntagme prépositionnel lorsqu'il se compose avec un prédicat collectif.

Dans les prédications à lecture distributive, les référents restent individualisés du fait de rôles thématiques différents ce qui a pour conséquence de favoriser les reprises d'un seul référent.

Nos résultats sont en accord avec le statut polysémique du syntagme comitatif *avec SN* : les divers rôles thématiques dévolus au référent qu'il régit dont il a été question en 2.1.2 ont des effets sur les enchaînements discursifs suivants. L'utilisation de verbes de sensation corporelle (*geler, pleurer, angoisser, trembler* etc...) dans les items tests ont eu pour conséquence de placer le référent sujet et le référent régi dans des rôles différents : le regroupement des référents dispersés en était rendu plus difficile, du fait de l'absence de lien sémantique entre les deux référents, ce qui entraînait moins de reprises plurielles. Ce résultat est conforme à ceux de Moxey et al. (2004) qui montraient qu'alors que les référents partagent la même identité ontologique, leur regroupement dans une entité plurielle est moins fréquent lorsqu'ils ont des rôles distincts dans la situation décrite.

Les différences de traitements qu'entraîne le genre du NP en tête de phrases amorces relèvent de paramètres syntactico-sémantiques et discursifs distincts. Pour en tenir compte, l'expérience de compréhension qui suit se compose de deux modules : dans le premier, le NP féminin sera en tête et le plan d'expérience comportera les facteurs « Verbe », « Contraste » et « Nombre de la reprise » (au sens grammatical du terme qui distingue le singulier du pluriel) de la phrase suivante (phrase cible). Dans le second, le NP masculin sera en tête tout en suivant un plan expérimental identique (facteurs, « Verbe », « Contraste » et « Nombre de la reprise » de la phrase cible).

## **4.3 Test 1**

### **4.3.1 Plan expérimental**

#### **A- Matériel et Facteurs expérimentaux**

On fait suivre les phrases amorces du pré-test d'une phrase cible (dont le nombre de caractères est compris entre 52 et 61), qui commence par une reprise pronominale soit au masculin singulier, soit au pluriel, c'est-à-dire qui reprend soit le référent masculin des items, soit les deux référents.

Une phrase tampon après la phrase cible permet d'introduire une distance entre cette dernière et la question de compréhension qui suit chaque texte : elle permet de ne pas faire coïncider les traitements de résolution pronominale et la fin de l'item.

Les trois phrases, amorce, cible et tampon, sont suivies d'une affirmation dont les participants ont à évaluer la véracité par rapport au texte qu'ils viennent de lire. Pour alterner les réponses positive et négative, cette interrogation implicite portait toujours sur le NP féminin et sur l'évènement ou l'action qui s'était produit-e dans la phrase cible : de cette manière, quand la phrase cible ne reprenait que le référent masculin, la réponse à la question était fausse.

Les phrases sont présentées l'une après l'autre selon la méthodologie APS (Auto Présentation Segmentée : voir 4.2.2).

Pour tenir compte des résultats du pré-test sur l'effet de proéminence du NP en tête de phrase amorce et l'influence de son genre sur les reprises anaphoriques suivantes, nous réalisons deux modules indépendants pour ce test : l'un où c'est le NP féminin qui est en tête de phrase amorce et

l'autre où c'est le NP masculin. Dans chacun des modules, les facteurs fixes sont les facteurs « Nombre » de la reprise de la phrase cible (Singulier ou Pluriel), catégorie du « Verbe » de la phrase amorce (syMétrique ou Non syMétrique), et groupe de « Contraste » des items (Symétrique/Collectif ou Symétrique/Distributif<sup>49</sup>). Dans le module FFirst, le référent féminin est sujet grammatical de la phrase amorce et le référent masculin est régi par la préposition *avec*. Dans le deuxième module, MFirst, c'est l'inverse.

Un exemple d'item test du groupe de contraste S/D commençant par un NP féminin est présenté en (221a-d), ainsi qu'un exemple d'item de test du contraste S/C qui commence par un NP féminin en (222a-d).

(221a) *Audrey cohabite/déprime à Courchevel avec Robin pendant les vacances scolaires.*

(221b) *Ils/Il fréquente(nt) tous les lieux branchés du centre-ville.*

(221c) *Les célébrités ne manquent pas dans le coin.*

(221d) ?? *Audrey fréquente les lieux branchés de Courchevel.*

(222a) *Brenda sympathise (grimpe) dans le métro avec Samy en un dixième de seconde.*

(222b) *Ils/Il attire(nt) les regards scrutateurs des ados présents.*

(222c) *Les jeunes ont des capacités d'observation acérées.*

(222d) ?? *Louise attire les regards scrutateurs des ados présents.*

Le module Mfirst suit le même plan expérimental que FFirst. Les mêmes items sont repris en permutant les prénoms féminin et masculin (dans les textes mais pas dans les questions).

Le test comporte 24 textes (l'ensemble du matériel de test est en annexe 4), déclinés en quatre versions selon les conditions. Celles-ci sont réparties sur quatre groupes en contrebalançant les items de façon à ce qu'il y ait le même nombre d'items par condition dans chaque groupe (six) et qu'un même texte n'apparaisse dans un groupe que dans une version (voir le plan expérimental en annexe 4). Les items expérimentaux sont complétés par 24 textes distracteurs qui sont les mêmes dans les quatre groupes (voir les items de test et les distracteurs en annexe 4).

## B- Hypothèses

Etant donnés les résultats du pré-test, les hypothèses sont les suivantes :

---

<sup>49</sup> Sur les figures, le groupe contraste Symétrique/Collectif est codé S/C et le groupe contraste Symétrique/Distributif, codé S/NC ; les prédicats symétriques sont codés M et les prédicats mixtes sont codés NM.

Dans le premier module, l'hypothèse est que les phrases cibles reprenant les deux référents 1) seront lues plus rapidement après les phrases amorces collectives qu'après les phrases amorces symétriques et 2) seront lues plus rapidement après les phrases amorces symétriques qu'après les phrases amorces distributives. Dans le deuxième module, l'hypothèse est que les phrases cibles reprenant le référent masculin 1) seront lues plus rapidement après les phrases amorces distributives qu'après les phrases amorces symétriques et 2) seront lues plus rapidement après les phrases amorces symétriques qu'après les phrases amorces collectives.

Et plus précisément :

La lecture d'un pronom pluriel après une phrase amorce collective devrait être rapide étant donné que le syntagme comitatif est équivalent à une coordination avec ces prédicats. Celle d'un pronom masculin devrait être moins rapide pour la même raison car à devrait nécessiter un coût d'extraction des membres d'un collectif.

La lecture d'un pronom pluriel après une phrase amorce symétrique devrait être rapide ainsi que la reprise d'un pronom masculin lorsque le NP en tête est féminin. Lorsqu'il est masculin, la lecture d'un pronom de reprise pluriel devrait être également rapide et celle d'une reprise au singulier encore plus rapide que la reprise plurielle.

La lecture d'un pronom de reprise pluriel après une phrase amorce distributive devrait être ralentie par rapport aux autres prédicats et celle d'un pronom masculin encore plus ralentie, lorsque le NP en tête est féminin. Lorsqu'il est masculin, la lecture d'un pronom masculin devrait être très rapide et celle d'un pronom pluriel au contraire ralentie.

## C- Méthodologie APS

Les expériences sont réalisées avec le logiciel E-prime version 2 professionnelle (*Psychology Software Tools, Inc.*).

Chacune des phrases des textes apparaît au centre de l'écran, sur une ligne. Les participants au test reçoivent la consigne de les lire à leur rythme et d'appuyer sur la barre d'espace lorsqu'ils ont fini. La phrase suivante prend alors place au centre de l'écran remplaçant celle qui vient d'être lue. Et ainsi de suite. La consigne précise également qu'une question est posée en fin de texte sous forme d'affirmation précédée de deux points d'interrogation : les participants sont informés que leur tâche consiste à appuyer sur la touche verte (le chiffre 1 du clavier sur lequel on a collé une pastille verte)

s'ils considèrent que l'affirmation est fidèle au texte qu'ils viennent de lire et sur la touche rouge (le chiffre 0 du clavier sur lequel on a collé une pastille rouge) dans le cas contraire.

En début de session, deux exemples permettent aux participants de se familiariser avec la procédure. L'ordre d'apparition des textes et des distracteurs est aléatoire permettant d'éviter les biais dus à des enchaînements intertextes figés.

Les temps de lecture de chaque phrase sont enregistrés ainsi que les réponses aux questions.

## D- Conditions de passation

Pour les tests APS, les étudiants étaient recrutés sur la base du volontariat à l'université de Paris-Sorbonne. Ils recevaient une entrée au forum des images en échange de leur participation.

Ils ont passé le test un par un dans un petit local de la bibliothèque de l'UMR 7172 -THALIM (Théorie et Histoire des Arts et de la Littérature de de la Modernité », grâce à la bienveillance de ses directeurs successifs, M. Michel Collot et M. Alain Schaffner.

Vingt-huit étudiants ont passé le test FFirst et 28 autres le test MFirst. Chaque test durait en moyenne 25 mn.

## 4.3.2 Analyses et résultats

### 4.3.2.1 Module NP féminin en tête : FFirst

La méthodologie APS souffre des moments d'inattention qui peuvent surgir du fait de la procédure de lecture adoptée et qui ont des causes multiples. Les temps bruts de lecture sont examinés participant par participant, et leur moyenne et écart-type calculés, pour éliminer les temps qui seraient supérieurs ou inférieurs de deux écart-types à la moyenne du participant. L'étape suivante, toujours pour les mêmes raisons, consiste à supprimer les temps de lecture qui correspondent à des textes auxquels le participant n'a pas donné la bonne réponse. Enfin, malgré ce premier « nettoyage », certains participants présentent des temps de lecture tellement différents de ceux des autres que leur contribution ne peut être retenue : une fonction du logiciel R (scale qui calcule le z-score du groupe et signale les outsiders) permet de les détecter. La lenteur excessive des temps de lecture peut provenir du fait, par exemple, que certains étudiants, pour être à même de donner une

réponse juste, essaient de mémoriser les phrases des textes en les relisant plusieurs fois, ce qui biaise leur lecture ; certains autres participants sont écartés des analyses parce que leur taux de mauvaises réponses est supérieur à 33%.

Dans FFirst, certains participants ont été écartés des analyses : S1 a fait plus de 33% d'erreurs, S12 n'a aucune bonne réponse dans la condition « verbe non symétrique et reprise du référent régi », S9 est de langue maternelle arabe (problème d'intégration des pluriels : 5863 ms pour les phrases cibles au pluriel pour 3356 ms pour celles qui sont au singulier) et S20 connaissait les caractéristiques des textes pour avoir été informée par une amie. Restent 24 participants.

La moyenne des temps de lecture bruts de la phrase cible (P2RT) est de 3520 ms (écart type ou standard deviation= 381<sup>50</sup>). La moyenne des temps de lecture « nettoyés » de la phrase cible (P2RT) est de 3250 ms (sd=253). La moyenne des temps de lecture de la phrase cible (P2RT) après élimination des mauvaises réponses et des participants dont les temps sont hors cadre est de 3147 ms (sd=257).

Les analyses conduites avec le modèle linéaire mixte du logiciel R (fonction « lmer » de la librairie lme4, suivant la loi normale), portent sur les temps de lecture – une fois « nettoyés » – des 24 participants restants.

### **Facteur « Nombre »**

On trouve un effet principal du facteur « Nombre » sur les temps de lecture de la phrase cible (Chisquare=8.24, df= 1,  $p < 0.005$ ) : les phrases cibles démarrant par un pronom pluriel sont lues plus rapidement (2978 ms, sd=244) que les phrases cibles démarrant par un pronom singulier (3301 ms, sd=264).

Ce résultat réplique en le transposant le résultat trouvé dans le pré-test lorsque le NP féminin est en tête de phrase amorce, à savoir que globalement le taux de reprises plurielles est significativement plus élevé que celui des reprises au singulier : transposée dans le test de compréhension FFirst, la prédisposition globale à former une entité plurielle dans cette condition se traduit par une plus grande vitesse de lecture de la phrase cible, ce qui est tout à fait cohérent.

---

<sup>50</sup> Dans les analyses nous utiliserons l'abréviation *sd* du terme anglais *standard deviation* pour sa maniabilité.

Par ailleurs, on trouve que les temps de réponse à la question finale sont beaucoup plus lents (Chisquare=47, df= 1,  $p < 0.001$ ) lorsqu'il y a une reprise plurielle (3970 ms, sd=435) que lorsqu'il y a une reprise du référent masculin (2934 ms, sd=322) dans la phrase cible.

Aucun des autres effets attendus n'est obtenu.

#### 4.3.2.2 Module NP masculin en tête : MFirst

Les étapes de l'analyse sont les mêmes que dans le test FFirst.

Cinq participants ont été écartés des analyses : S23 qui n'a pas fait la différence entre les reprises des deux référents et les reprises du seul référent masculin en tête (aucune réponse n'est bonne quand la phrase cible est au singulier), S14 et S16 détectés comme différents de la cohorte des participants par la fonction « scale », S16 qui a précisé après le test qu'il lisait toutes les phrases deux fois, S9 qui est de langue maternelle espagnole. Restent 24 participants.

La moyenne des temps de lecture bruts de la phrase cible (P2RT) est de 3178 ms (sd=368). La moyenne des temps de lecture nettoyés de la phrase cible (P2RT) est de 2989 ms (sd=273). La moyenne des temps de lecture de la phrase cible (P2RT) après élimination des mauvaises réponses et des participants dont les temps sont hors cadre, est de 2708 ms (sd=227).

Les analyses conduites avec le modèle linéaire mixte (lmer) du logiciel R sur les participants sélectionnés, n'indiquent aucun effet, ni sur la phrase cible, ni sur la phrase amorce.

Le facteur « Nombre » est significatif sur les temps de réponse : ils sont plus courts (2622 ms) (Chisquare=54, df= 1,  $p < 0.001$ ) pour les reprises au singulier que pour les reprises plurielles (3500 ms).

#### 4.3.2.3 Analyses des données regroupées

Le fait de grouper les données implique d'introduire dans les analyses un nouveau facteur, le facteur « TETE » à deux modalités : NP féminin ou NP masculin en tête de phrase.

Les étapes de l'analyse sont les mêmes que précédemment.

Les participants écartés dans les tests FFirst et MFirst, sont également écartés de cette analyse.

La moyenne des temps de lecture bruts de la phrase cible (P2RT) est de 3349 ms (sd=266). La moyenne des temps de lecture nettoyés de la phrase cible (P2RT) est de 3114 ms (sd=185). La moyenne des temps de lecture de la phrase cible (P2RT) après élimination des mauvaises réponses et des participants dont les temps sont hors cadre est de 2930 ms (sd=174).

### **Facteur « Nombre »**

Les analyses conduites avec le modèle linéaire mixte du logiciel R sur les participants sélectionnés sur les facteurs « Verbe », « Nombre », « Contraste » et « TETE » montrent un effet principal du facteur « Nombre » sur les temps de lecture des phrases cibles (Chisquare=3.86, df= 1,  $p < 0.05$ ) : les phrases cibles démarrant par un pronom pluriel sont lues plus rapidement (2844 ms, sd=171) que les phrases cibles démarrant par un pronom singulier (3010 ms, sd=177).

On retrouve, transposé à la tâche de compréhension, l'effet global du facteur « Nombre » trouvé dans le pré-test.

### **Interaction significative des facteurs « TETE » et « NB »**

L'analyse montre aussi une interaction significative entre les facteurs « TETE » et « NB » (Chisquare=5.82, df= 1,  $p < 0.02$ ) : Tableau 13 et Figure 2.

Tps de lecture de la phrase cible (écart-type)	Reprises au pluriel	Reprises au singulier
NP Féminin en tête	2978 (173)	3301 (186)
NP Masculin en tête	2709 (166)	2707 (155)

Tableau 13 : Temps de lecture des phrases cibles en fonction du facteur « Nombre » selon le genre du NP en tête de phrase amorce dans le premier test – regroupement des données de FFirst et MFirst

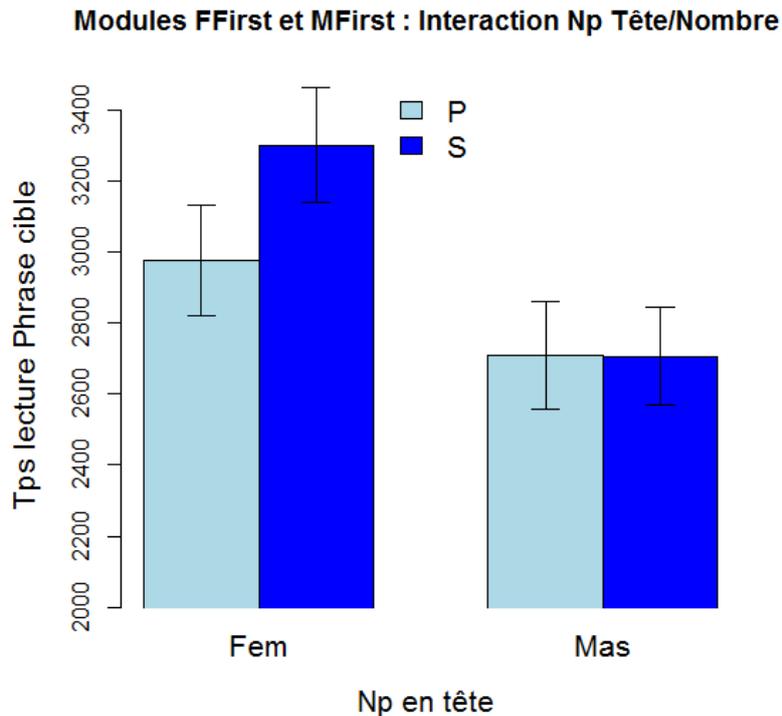


Figure 2 : Temps de lecture de la phrase cible en fonction du NP en tête et de la reprise (données regroupées, test 1)

L'interaction est cohérente avec les résultats trouvés dans les modules FFirst et MFirst. Lorsque le NP féminin est en tête de phrase amorce, les temps de lecture de la phrase cible dépendent de la reprise : ils sont plus longs (Chisquare=8.47, df= 1, p=0.004) pour les reprises du référent masculin que pour les reprises des deux référents.

De plus, les phrases cibles qui reprennent le référent masculin sont lues plus rapidement (Chisquare=4.23, df= 1, p=0.04) lorsque le NP masculin est en tête (2707 ms, sd=155) que lorsque c'est le NP féminin (3301 ms, sd=186). Il n'y a pas d'autres différences : les phrases cibles au singulier et au pluriel sont lues à des vitesses équivalentes lorsque le NP masculin est en tête. Ce résultat confirme le statut proéminent du NP de tête relativement à celui d'un référent régi par la préposition *avec*.

Les temps de réponse à la question sont considérablement plus longs (Chisquare=69, df= 1, p<0.001) lorsque la phrase cible reprend les deux référents (4102 ms, sd=420) plutôt que le seul référent masculin (2984 ms, sd=293). La question portant sur le référent féminin exclusivement, les participants ont pu après quelques items associer pronom au singulier et réponse négative. Dans le test suivant, la question porte sur les deux référents pour corriger ce biais.

### 4.3.3 Discussion

Quand on regroupe les données des deux modules, on trouve que le facteur « Nombre » a un effet significatif sur les temps de lecture : les phrases cibles sont lues plus vite quand elles reprennent les deux référents plutôt qu'un seul.

Par ailleurs, une interaction entre les facteurs « Nombre » et « Genre du référent en tête » montre que lorsque ce référent est féminin, les phrases cibles au pluriel sont lues plus rapidement que les phrases cibles au singulier, alors que lorsqu'il est masculin, les phrases cibles au pluriel sont lues à la même vitesse que les phrases cibles au singulier : ces résultats sont conformes à ceux de Hielscher et Müsseler (1990) qui ont testé les constructions syntaxiques discontinues avec plusieurs syntagmes prépositionnels (*and, as well as, neither/nor of, with, instead of, without*) : lorsque le syntagme prépositionnel est comitatif, ils trouvent que les reprises plurielles sont lues à la même vitesse que les reprises d'un seul référent si ce référent est sujet syntaxique en tête de phrase amorce et qu'en revanche, les reprises plurielles sont lues plus rapidement que les reprises d'un seul référent lorsque ce dernier est régi par la préposition.

Ces résultats suggèrent que l'élaboration d'une entité plurielle à partir de référents dispersés est un processus plus rapide à effectuer que l'extraction d'un référent régi par la préposition *avec*. Le fait que les phrases cibles reprenant le référent masculin sujet en tête de phrase amorce, ne soient pas lues plus rapidement que les phrases cibles pluriel, malgré la grande accessibilité du référent, renforce encore cette hypothèse d'accès très rapide à l'entité plurielle formée par les référents dispersés des constructions discontinues en *avec*. Compte tenu des propriétés plurielles de nos items de test qui comportent pour moitié des prédicats symétriques et pour un quart des prédicats à interprétation collective, ce résultat est cohérent.

Malgré l'intérêt de ces résultats, ils ne vérifient pas les hypothèses que nous avons proposées à partir des résultats du pré-test : on ne trouve aucun effet de la nature des prédicats amorces sur la résolution des anaphores pronominales des phrases cibles.

La facilité de formation d'une entité plurielle à partir de référents dispersés, relevée ci-dessus est peut-être le signe de traitements précoces rendant difficiles de différencier les effets des prédicats des phrases amorces sur la phrase cible. Il est donc décidé de segmenter les phrases amorces et cibles dans le prochain test.

Par ailleurs, comme l'indiquent les temps de lecture des phrases amorces qui sont presque deux fois supérieurs à ceux des phrases cibles malgré un nombre de caractères comparable, la première phrase des textes reçoit une attention particulière de la part des lecteurs : c'est en les lisant qu'ils posent les bases de la représentation ou scénario sur lequel développer la suite du texte. Une possible amélioration du matériel de test consiste donc à séparer la phase de prise de connaissance du scénario qui sert de toile de fond au texte, de la phrase amorce qui prend en compte les facteurs que l'on veut tester. Il est donc décidé d'insérer une phrase d'introduction au texte avant la phrase amorce dans le prochain test.

Enfin, pour éviter l'association des réponses à la question avec les reprises pronominales au singulier, les questions sont modifiées : elles porteront sur le pluriel coordonné formé par les deux référents du texte, référent masculin en tête de coordination. De cette façon, seront contrastés les pluriels coordonnés dont le référent de tête est de même genre que le NP en tête de phrase amorce (MFirst-I) et ceux qui sont de genre différent (FFirst-I).

## **4.4 Test 2**

### **4.4.1 Plan expérimental**

#### **A- Matériel et Facteurs expérimentaux**

Deux modules de ce test sont réalisés, comme dans le premier test en temps réel (« on line ») : un module où le NP en tête de phrase amorce est féminin (FFirst-I) et un module où le NP en tête de phrase amorce est masculin (MFirst-I).

Conformément au bilan après le premier test « on line », une phrase d'introduction a été ajoutée à tous les textes et une segmentation en trois parties des phrases amorces et cibles et en deux parties des phrases tampons, a été réalisée. De plus, les questions en fin d'items sont modifiées. Au lieu de faire porter l'interrogation sur l'engagement du référent féminin dans l'évènement décrit par la phrase cible, on la fait porter sur l'engagement des deux référents. Les réponses attendues sont positives pour un pronom pluriel dans la phrase cible et négatives sinon. Le référent masculin est toujours le référent en tête de la coordination dans les questions. De cette façon, seront contrastés les pluriels coordonnés dont le référent de tête est de même genre que le NP en tête de phrase

amorce (MFirst-I) et ceux qui sont de genre différent (FFirst-I). La phrase d'introduction et l'interrogation en fin de texte ne sont pas segmentées.

Les facteurs expérimentaux sont les mêmes que précédemment : « Verbe », « Nombre », « Contraste ». Un exemple d'item du groupe contraste S/D, où le NP féminin est en tête est donné en (223a-j), et un autre exemple d'item du groupe contraste S/C où le NP féminin est en tête également est donné en (224a-j).

(223a) *La montagne n'est pas que le royaume de la glisse au grand air.*

(223b) *Audrey cohabite/déprime à Courchevel*

(223c) *avec Robin*

(223d) *dès l'ouverture de la station.*

(223e) *Ils/il fréquente(nt)*

(223f) *tous les lieux branchés*

(223g) *du centre-ville.*

(223h) *Les célébrités*

(223i) *ne manquent pas dans le coin.*

(223j) ?? *Robin et Audrey rencontrent des célébrités à Courchevel.*

(224a) *La dépersonnalisation n'est pas une fatalité dans les transports.*

(224b) *Louise sympathise/grimpe dans le métro*

(224c) *avec François*

(224d) *en un dixième de seconde.*

(224e) *Ils/Il irradie(nt)*

(224f) *une évidente et irrépressible*

(224g) *joie de vivre.*

(224h) *L'expression du visage*

(224i) *peut refléter une force intérieure.*

(224j) ?? *François et Louise respirent visiblement la joie de vivre.*

Le test comporte 24 textes (voir les items expérimentaux en annexe 5), déclinés en quatre versions selon les conditions<sup>51</sup>. Les quatre versions de chacun des 24 textes sont réparties sur quatre groupes

---

<sup>51</sup> NB : Certains contenus textuels du test 1 sont modifiés pour les rendre plus naturels. Par ailleurs, les questions qui portaient sur le référent féminin dans le test 1 portent maintenant sur la paire de référents coordonnés pour éviter les biais dus à une focalisation sur le prénom féminin en cours de lecture.

en contrebalançant les items de façon à ce qu'il y ait le même nombre d'items par conditions dans chaque groupe (6) et qu'un même texte n'apparaisse dans un groupe que dans une version (voir le plan expérimental en annexe 5). Les items expérimentaux sont complétés par 24 textes distracteurs (voir les distracteurs utilisés en annexe 5) qui sont les mêmes dans les quatre groupes.

Les items et le plan expérimental restent les mêmes mais les prénoms féminin et masculin sont permutés dans les textes sans l'être dans les questions.

## B- Procédure

La méthodologie de présentation des textes est l'APS (Auto Présentation segmentée), comme précédemment : la phrase d'introduction s'affiche au centre de l'écran dans son entier, puis le premier segment, suivi du deuxième segment, suivi lui-même du troisième segment de la phrase amorce, puis le premier segment suivi du deuxième segment, suivi lui-même du troisième segment de la phrase cible, puis le premier segment suivi du deuxième segment de la phrase tampon, puis la question qui s'affiche en entier.

## C- Hypothèses

On fait les mêmes hypothèses que dans le test 1 : Dans le module FFirst-I, l'hypothèse est que les phrases cible reprenant les deux référents 1) seront lues plus rapidement après les phrases amorces collectives qu'après les phrases amorces symétriques et 2) seront lues plus rapidement après les phrases amorces symétriques qu'après les phrases amorces distributives. Dans le module MFirst-I, l'hypothèse est que les phrases cibles reprenant le référent masculin 1) seront lues plus rapidement après les phrases amorces distributives qu'après les phrases amorces symétriques et 2) seront lues plus rapidement après les phrases amorces symétriques qu'après les phrases amorces collectives.

## D- Conditions de passation

Trente-neuf participants ont passé la version FFirst-I et 27 la version MFirst-I dans des conditions identiques à celles du test 1.

## 4.4.2 Analyses et résultats

### 4.4.2.1 Module NP féminin en tête : FFirst-I

Les temps de lecture des phrases cibles sont calculés en additionnant les temps de lecture de chaque segment des phrases cibles.

Le « nettoyage » des temps de lecture est effectué sur le temps de lecture global des phrases cibles, conformément à la méthodologie d'analyse adoptée dans le premier test.

Certains participants sont écartés des analyses : les participants S8, S9, S12, S16, S24 ont fait plus de 33% d'erreurs et ne sont pas retenus. De plus, les temps de lecture de chaque segment sont traités par la fonction « scale » de R qui calcule le z-score du groupe sur une variable dépendante et signale les outsiders - participants S3, S4, S6, S10, S22, S25, S31.

La moyenne des temps de lecture bruts de la phrase cible (P2RT) est de 3466 ms (sd=259). La moyenne des temps de lecture nettoyés de la phrase cible (P2RT) est de 3329 ms (sd=228). La moyenne des temps de lecture de la phrase cible (P2RT) après élimination des mauvaises réponses et des participants dont les temps sont hors cadre est de 3109 ms (sd=214).

Les analyses conduites avec le modèle linéaire mixte du logiciel R (fonction « lmer » de la librairie lme4) suivant la loi normale, portent sur les temps de lecture « nettoyés » des 27 participants restants.

#### **A- Premier segment des phrases cibles**

Il n'y a aucun effet sur ce premier segment.

#### **B- Deuxième segment des phrases cibles**

##### **Facteur « Nombre »**

On trouve un effet principal du facteur « Nombre » sur les temps de lecture du deuxième segment des phrases cibles et une interaction entre les facteurs « Verbe » et « Contraste ».

Le deuxième segment des phrases cibles est lu plus rapidement (Chisquare=9.05, df=1, p= 0.002) pour les reprises des deux référents dispersés (957 ms, sd=101) que pour les reprises du référent régi par le syntagme comitatif (1076 ms, sd=117).

### Interaction entre les facteurs « Verbe » et « Contraste »

Le temps de lecture du deuxième segment des phrases cibles présentent une interaction à la limite de la significativité entre les facteurs « Verbe » et « Contraste » (Chisquare=5.13, df=1, p= 0.023) : Tableau 14 et Figure 3.

FFirst-I : Tps de lecture du 2 <sup>ième</sup> segment des phrases cibles (ms)	Verbe Symétrique	Verbe non symétrique
Contraste S/C	998 (105)	955 (97)
Contraste S/D	1016 (101)	1079 (130)

**Tableau 14** : Temps de lecture du 2<sup>ième</sup> segment de la phrase cible en fonction du groupe contraste et du type de phrase amorcée dans le module FFirst-I

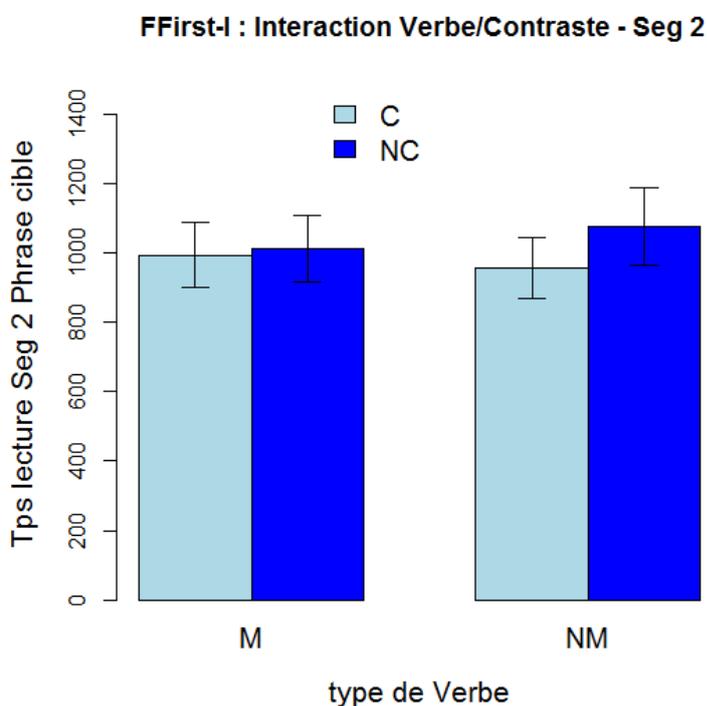


Figure 3 : Temps de lecture des deuxièmes segments de la phrase cible en fonction groupe de contraste et de la phrase amorcée dans le module FFirst-I

Les temps de lecture des 2<sup>èmes</sup> segments des phrases cibles sont comparables après des phrases amorces symétriques et collectives. En revanche, ils sont plus lents (Chisquare= 5.16, df=1, p=0.023) dans le groupe contraste S/D : autrement dit, ils sont plus longs après les phrases amorces distributives relativement aux autres conditions.

## **C-Troisième segment des phrases cibles**

Il n'y a aucun effet sur le troisième segment.

## **D- Phrase cible**

- **ANALYSE PRINCIPALE**

### **Facteur « Nombre »**

On trouve un effet principal du facteur « Nombre » : les phrases cibles sont lues plus rapidement (Chisquare=9.17, df=1, p=0.01) pour les reprises des deux référents (2963 ms, sd=204) des phrases amorces que pour les reprises du référent régi par le syntagme comitatif (3235 ms, sd=230).

### **Triple interaction entre les facteurs « Verbe », « Nombre » et « Contraste »**

De plus, les temps de lecture de la phrase cible dans son ensemble présentent une triple interaction entre les facteurs « Verbe », « Nombre » et « Contraste » (Chisquare=5.4, df=1, p=0.02).

- **ANALYSES COMPLEMENTAIRES**

### **ANALYSE COMPLEMENTAIRE : CONDITION GROUPE CONTRASTE S/C**

Des analyses complémentaires montrent que dans le **groupe de contraste S/C**, la vitesse de lecture des phrases cibles **ne dépend pas** des aspects suivants : la phrase amorce (Chisquare=0.57 df=1, p=0.45), le nombre de la reprise (Chisquare=2.70, df=1, p=0.10) ou l'interaction entre les facteurs (chisquare=2.36, df=1, p=0.12) : les temps de lecture des phrases cibles qui reprennent les deux référents (2935 ms, sd=196) sont similaires à ceux des phrases cibles qui reprennent le seul référent régi (3097 ms, sd=203)

### **ANALYSE COMPLEMENTAIRE : CONDITION GROUPE CONTRASTE S/D**

#### **Facteur « Nombre »**

En revanche, dans le groupe contraste S/D, les analyses complémentaires montrent que le facteur « Nombre » est significatif : les phrases cibles sont lues plus rapidement (Chisquare= 5.01, df=1, p= 0.023) en cas de reprise des deux référents (3044 ms, sd=204) qu'en cas de reprise du seul référent régi (3339 ms, sd=251).

### Interaction entre les facteurs « Verbe » et « Nombre »

De plus, on trouve une interaction significative entre les facteurs « Verbe » et « Nombre » sur les temps de lecture de la phrase cible (Chisquare=5.33, df=1, p=0.02) : Tableau 15 et Figure 4.

FFirst-I : Tps de lecture des phrases cibles (ms)	Reprise des 2 référents	Reprise du référent régi
Verbe Symétrique	3158 (191)	3080 (190)
Verbe distributif	2937 (214)	3575 (291)

Tableau 15 : Temps de lecture de la phrase cible en fonction du « Nombre » de la reprise et du type de phrase amorce dans le module FFirst-I

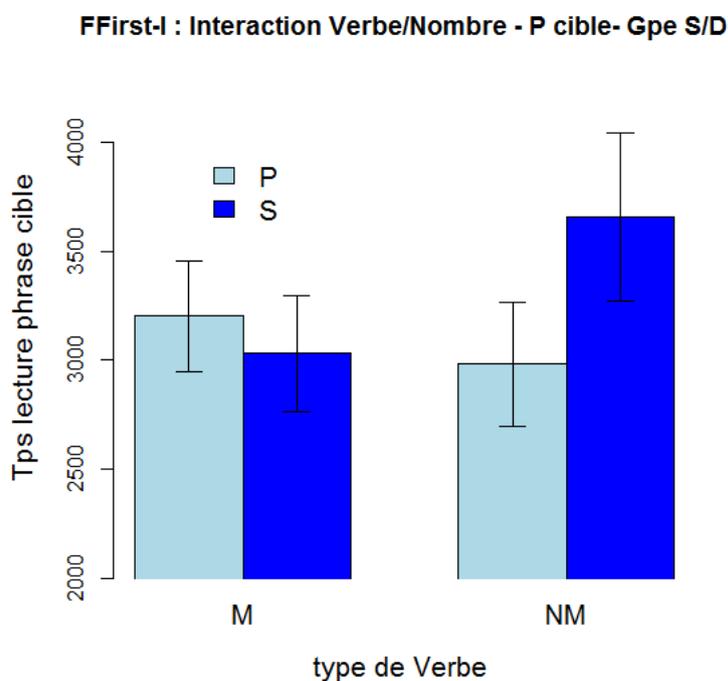


Figure 4 : Temps de lecture des phrases cibles en fonction de la reprise et de la phrase amorce dans le groupe de contraste S/D - module FFirst-I

Les temps de lecture des phrases cibles au singulier après les prédicats distributifs (3575 ms) sont plus élevés (Chisquare=3.5, df=1, p=0.06) que ceux qui suivent les prédicats symétriques (3080 ms). Les temps de lecture des phrases cibles après les prédicats distributifs sont aussi plus élevés (Chisquare=9.64, df=1, p=0.002) pour les reprises du référent régi (3575 ms) que pour celles des deux référents (2937 ms).

Nos résultats ne sont conformes à ceux de Hielscher et Musseler (1990) que pour les prédicats distributifs : les phrases cibles qui reprennent les deux référents sont lues plus rapidement que les phrases cibles qui reprennent le référent régi après les phrases amorces distributives du test. En revanche, pour les prédicats amorces symétriques et collectifs, le « Nombre » de la reprise n'a pas d'influence sur les temps de lecture. Ces résultats suggèrent que les prédicats symétriques et collectifs de ces constructions discontinues en *avec* relient le référent du syntagme nominal sujet et le référent du syntagme comitatif de façon plus étroite que les prédicats distributifs. Les prédicats collectifs créent un rôle commun entre eux. Les prédicats symétriques créent une relation d'équivalence entre eux. Ces « bases d'association communes » entre le référent sujet et le référent régi permettent de comprendre que le référent régi reste accessible avec ces prédicats.

### **E- Temps de réponse**

Il n'y a aucun effet des facteurs expérimentaux sur les taux de bonnes réponses, ni sur les temps de réponse (4133 ms en moyenne) à la question finale.

#### **4.4.2.2 Module NP masculin en tête : MFirst-I**

Les temps de lecture des phrases cibles sont calculés en additionnant les temps de lecture de chaque segment des phrases cibles.

Le « nettoyage » des temps de lecture est effectué sur le temps de lecture global des phrases cibles, conformément à la méthodologie d'analyse adoptée dans le test 1.

Quatre participants ont été écartés des analyses. S20 a fait plus de 33% d'erreurs et n'est pas retenu. De plus, les temps de lecture de chaque segment sont traités par la fonction « scale » de R qui calcule le z-score du groupe sur une variable dépendante et signale les outsiders : à la suite de quoi les participants S3, S18, S24 n'ont pas été retenus. Restent 23 participants.

La moyenne des temps de lecture bruts de la phrase cible (P2RT) est de 3746 ms (sd=284). La moyenne des temps de lecture nettoyés de la phrase cible (P2RT) est de 3652 ms (sd=269). La moyenne des temps de lecture de la phrase cible (P2RT) après élimination des mauvaises réponses et des participants dont les temps sont hors cadre est de 3419 ms (sd=231).

Les analyses conduites avec le modèle linéaire mixte du logiciel R (fonction « lmer » de la librairie lme4 suivant la loi normale), portent sur les temps de lecture « nettoyés » des 23 participants restants.

## A- Premier segment de la phrase cible

- ANALYSE PRINCIPALE

### Interaction des facteurs « Verbe » et « Nombre »

Les premiers segments des phrases cibles présentent une interaction des facteurs « Verbe » et « Nombre » (Chisquare= 3.77, df=1, p= 0.052) : Tableau 16 et Figure 5.

MFirst-I : Tps de lecture du 1 <sup>er</sup> segment des phrases cibles	Reprise des 2 référents	Reprise du référent masculin
Verbe symétrique	1041 (99)	935 (73)
Verbe Non symétrique	1013 (95)	1042 (120)

Tableau 16 : temps de lecture des phrases cibles en fonction du « Nombre » des reprises et de la phrase amorce – module MFirst-I

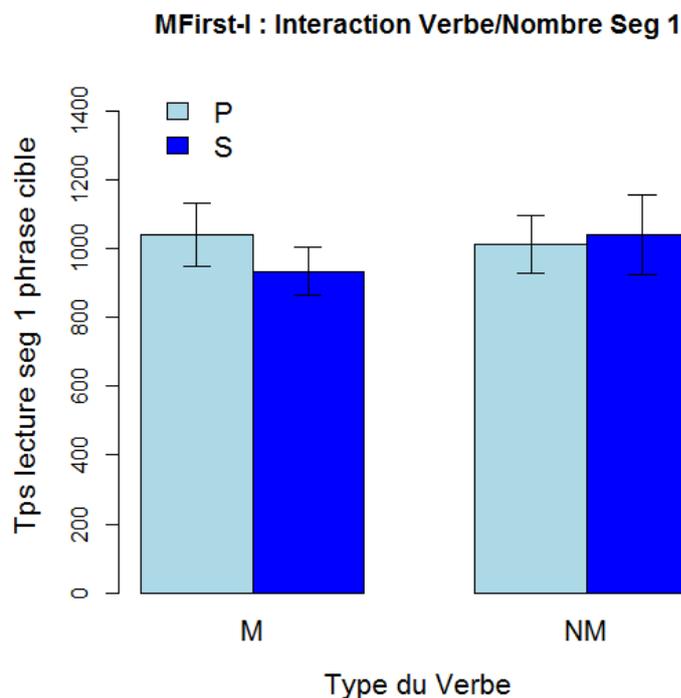


Figure 5 : Temps de lecture du premier segment des phrases cibles en fonction de la reprise et de la phrase amorce - module MFirst-I

Les temps de lecture des premiers segments des phrases cibles reprenant le référent masculin sont plus rapides après les phrases amorces symétriques qu’après les phrases amorces non symétriques ; ils sont également plus rapides que les temps de lecture des premiers segments reprenant les deux référents, quel que soit la phrase amorce.

### Triple interaction des facteurs « Verbe », « Nombre » et « Contraste »

De plus, les temps de lecture des 1<sup>ers</sup> segments des phrases cibles présentent une triple interaction entre les facteurs « Verbe », « Nombre » et « Contraste » (Chisquare=6.72, df=1, p< 0.01).

- **ANALYSES COMPLEMENTAIRES**

#### **ANALYSE COMPLEMENTAIRE : CONDITION GROUPE CONTRASTE S/C**

Des analyses complémentaires sur le groupe de contraste S/C montrent une interaction significative (Chisquare=9.58, df=1, p< 0.01) des facteurs « Verbe » et « Nombre » sur les lectures de ce 1<sup>er</sup> segment des phrases cibles : Tableau 17 et Figure 6.

Tps de lecture du 1 <sup>er</sup> segment de la phrase cible (écart-type)	Reprise des 2 référents	Reprise du référent en tête de phrase amorce
Verbe Symétrique	1066 (90)	885 (67)
Verbe collectif	1003 (81)	1159 (136)

Tableau 17 : temps de lecture des premiers segments des phrases cibles en fonction du « Nombre » des reprises et de la phrase amorce, dans le groupe de contraste Symétrique/Collectif – MFirst-I

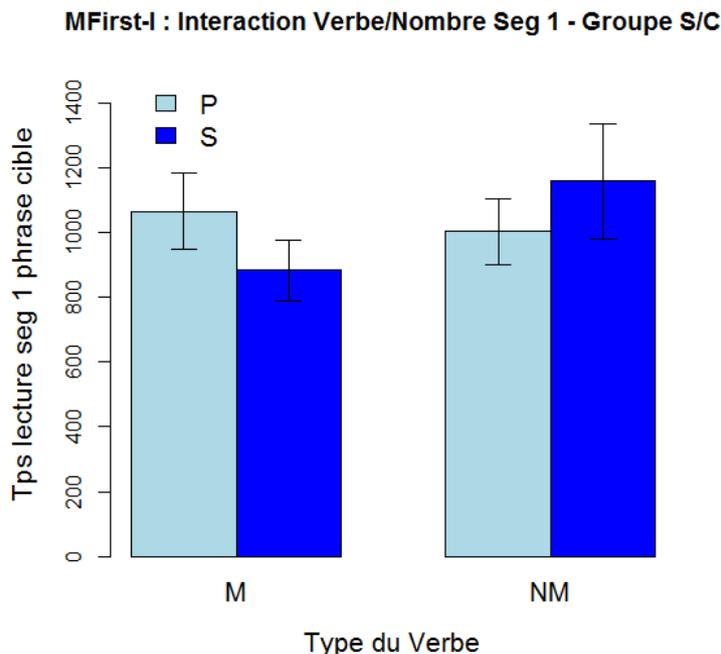


Figure 6 : Temps de lecture du premier segment des phrases cibles en fonction de la reprise et de la phrase amorce –dans le groupe de contraste S/C - module MFirst-I

Après les phrases amorces symétriques, les temps de lecture des premiers segments ne dépendent pas du nombre de la reprise (Chisquare=2.49, df=1, p=0.11).

Après les phrases amorces collectives, les temps de lecture des reprises plurielles (1003 ms) sont plus courts (Chisquare=7.47, df=1, p=0.006) que ceux des reprises au singulier (1159 ms).

Les temps de lecture des segments qui reprennent le référent proéminent sont aussi plus courts (Chisquare=9.05, df=1, p=0.003) en cas de prédicats amorces symétriques (885 ms) qu'en cas de de prédicats amorces collectives (1159 ms).

Ce résultat se comprend, si l'on tient compte du fait que la nature collective du prédicat prédispose la création d'une entité qui regroupe les deux référents et qui rend l'extraction d'un des membres de ce collectif plus coûteux, fut-il, saillant. Si cette analyse est juste, les différences de temps de lecture de ce premier segment selon qu'il suit un prédicat amorce symétrique ou collectif est un argument supplémentaire pour ne pas assimiler les prédicats symétriques à des prédicats collectifs.

#### **ANALYSE COMPLEMENTAIRE : CONDITION GROUPE CONTRASTE S/D**

Des analyses complémentaires sur le groupe de contraste S/D ne montrent aucun autre effet significatif. On peut en déduire que l'interaction entre les facteurs « Verbe » et « Nombre » trouvée sur le premier segment pour l'ensemble des temps de lecture est due au groupe de contraste S/C.

#### **B- Deuxième segment de la phrase cible**

##### **Interaction des facteurs « Contraste » et « Verbe »**

Les deuxièmes segments des phrases cibles présentent une interaction des facteurs « Contraste » et « Verbe » (Chisquare= 4.83, df=1, p<0.03) : Tableau 18 et Figure 7.

MFirst-I : Tps de lecture du 2 <sup>ème</sup> segment des phrases cibles	Verbe symétrique	Verbe Non symétrique
Contraste S/Collectif	1128 (117)	1135 (136)
Contraste S/Distributif	1129 (116)	1264 (144)

Tableau 18 : temps de lecture du deuxième segment des phrases cibles en fonction de la phrase amorce et du groupe contraste – module MFirst

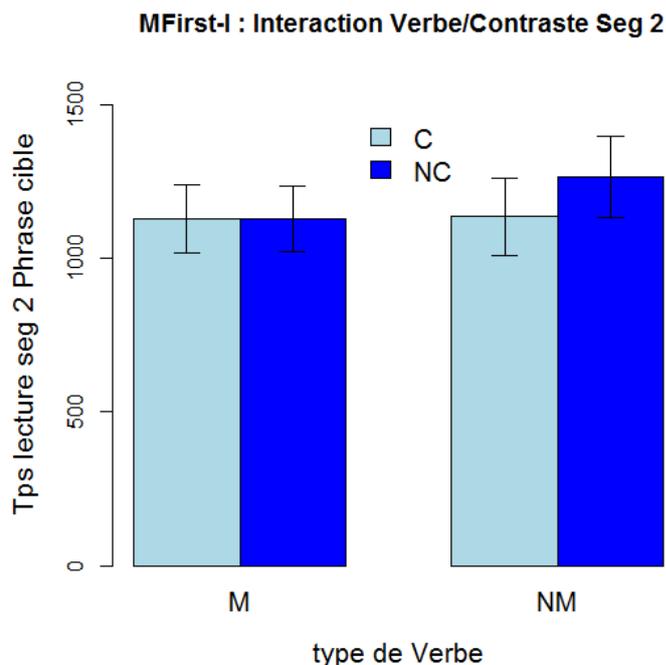


Figure 7 : Temps de lecture du deuxième segment des phrases cibles en fonction de la phrase amorce et du groupe contraste - module MFirst-I

Les temps de lecture des deuxièmes segments des phrases cibles sont équivalents après les prédicats amorces symétriques dans les deux groupes, résultat déjà obtenu dans le prétest, et dans la version FFirst de ce test où le syntagme nominal sujet est de genre féminin et le référent régi par la préposition de genre masculin.

En revanche, après les prédicats amorces non symétriques, les temps de lecture du deuxième segment des phrases cibles dépendent du groupe de contraste (Chisquare= 4.74, df=1, p= 0.029) : ils sont plus longs pour les distributifs (1264 ms) que pour les collectifs (1135 ms). Or, le référent repris en cas de phrase cible au singulier est proéminent dans la phrase amorce et devrait donner lieu à des traitements rapides après les prédicats distributifs.

### **Interaction contraste et « Nombre »**

On trouve aussi une interaction contraste et « Nombre » à la limite de la significativité (Chisquare= 3.76, df=1, p<0.052) : Tableau 19 et Figure 8.

MFirst-I : Tps de lecture du 2 <sup>ième</sup> segment des phrases cibles	Reprises au Pluriel	Reprise au Singulier
Contraste S/Collectif	1182 (137)	1077 (114)
Contraste S/Distributif	1136 (128)	1260 (133)

Tableau 19 : temps de lecture des deuxièmes segments des phrases cibles en fonction de la reprise et du groupe contraste – Module MFirst-I

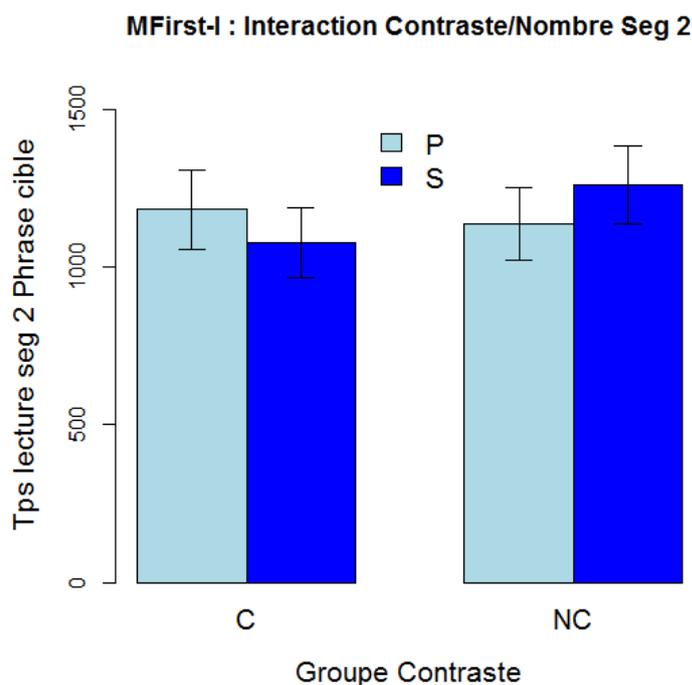


Figure 8 : temps de lecture des deuxièmes segments des phrases cibles en fonction de la reprise et du groupe contraste – Module MFirst-I

Les temps de lecture des deuxièmes segments des phrases cibles reprenant les deux référents sont équivalents dans les groupes de contraste, d'une part. Les reprises du référent proéminent dans le groupe de contraste S/C ne se différencient pas des reprise plurielles, d'autre part. En revanche, les reprises du référent proéminent sont plus longues (Chisquare= 4.45, df=1,  $p < 0.034$ ) à traiter dans le groupe contraste S/D (1260 ms), c'est-à-dire le groupe où se trouvent les distributifs, que dans le groupe de contraste S/C (1077 ms). Ce résultat est tout à fait inattendu.

## **C- Troisième segment de la phrase cible**

### **Facteur Groupe Contraste**

Les troisièmes segments des phrases cibles sont lues plus rapidement (Chisquare= 4.18, df=1, p= 0.04) dans le groupe contraste S/C (1173 ms, sd=113) que dans le groupe contraste S/D (1317 ms, sd=115).

On peut attribuer cette différence entre groupe contraste sur ce troisième segment à la prolongation des traitements entamés dans le deuxième segment après les phrases amorces distributives.

## **D- Phrase cible**

### **Facteur « Verbe »**

On trouve un effet principal du facteur « Verbe » : les phrases cibles sont lues plus rapidement (Chisquare= 4.49, df=1, p< 0.04) lorsque la phrase amorce est symétrique (3371 ms, sd=212) que lorsqu'elle est non symétrique (3465 ms, sd=231), quels que soient la reprise et le groupe contraste.

Ce résultat est cohérent avec ceux des segments précédents : ralentissement des temps de lecture du premier segment de la phrase cible après des phrases amorces collectives, en cas de reprise au singulier, ralentissement des temps de lecture du deuxième segment après des phrases amorces distributives, toutes reprises confondues.

## **E- Temps de réponse**

### **Facteur « Nombre »**

Le facteur « Nombre » est significatif sur les temps de réponse ; ils sont plus longs (Chisquare=9.9, df=1, p=0.007) lorsque les deux référents (4728 ms) sont repris dans la phrase cible que lorsque le seul référent masculin est repris (4037 ms).

### **Interaction entre les facteurs « Contraste » et « Nombre »**

Par ailleurs, il y a une interaction (Chisquare=5.84, df=1, p=0.016) entre les facteurs contraste et « Nombre » : Tableau 20 et Figure 9.

MFirst-I : Tps de réponse à la question finale	Reprises au Pluriel	Reprise au Singulier
Contraste S/Collectif	5047 (680)	3928 (370)
Contraste S/Distributif	4435 (511)	4146 (528)

Tableau 20 : temps de réponse à la question finale en fonction des facteurs « Contraste » et « Nombre » - Module MFirst-I

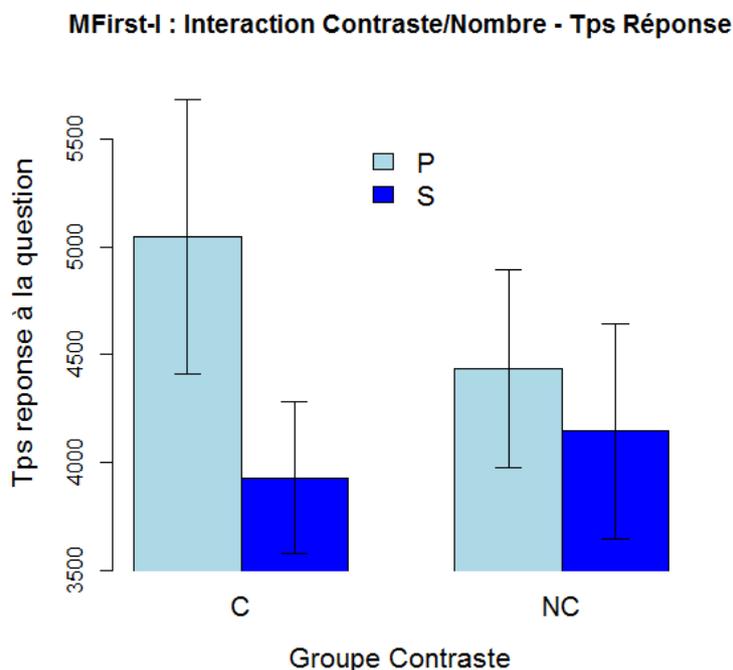


Figure 9 : temps de réponse à la question finale en fonction des facteurs « Contraste » et « Nombre » - Module MFirst-I

Les temps de réponse sont significativement ( $\chi^2=12.7$ ,  $df=1$ ,  $p<0.001$ ) plus lents lorsque la reprise dans la phrase cible concernait les deux référents (5048 ms,  $sd=639$ ) plutôt que le seul référent masculin (3928 ms,  $sd=349$ ) de la phrase amorce, dans le groupe S/C, mais pas dans le groupe S/D.

Les questions, qui étaient des affirmations dont le participant devait vérifier la véracité par rapport au texte lu, commençaient par un pluriel coordonné : c'était le NP masculin qui était en tête de coordination, c'est-à-dire l'argument en tête de phrase amorce. Tout se passe comme si le premier collectif élaboré dans le groupe S/C par les prédicats à interprétation collective devait être réévalué lorsqu'il est nommément retrouvé dans la question : un processus déjà rencontré pour la répétition de NP dans les enchaînements anaphoriques sur un argument saillant (Gordon et al., 1993). Le facteur « Verbe » n'a pas d'effet, ni dans le groupe S/C, ni dans le groupe S/D : il n'y a pas de différences entre les autres conditions sur les temps de réponse à la question.

### 4.4.2.3 Analyses des données regroupées

Le fait de grouper les données implique d'introduire dans les analyses un nouveau facteur, le facteur « TETE » à 2 modalités : NP féminin ou NP masculin en tête de phrase.

Les étapes de l'analyse sont les mêmes précédemment.

Les participants écartés dans les tests FFirst-I et MFirst-I, sont également écartés de cette analyse.

La moyenne des temps de lecture bruts de la phrase cible (P2RT) est de 3580 ms (sd=193). La moyenne des temps de lecture nettoyés de la phrase cible (P2RT) est de 3462 ms (sd=175). La moyenne des temps de lecture de la phrase cible (P2RT) après élimination des mauvaises réponses et des participants dont les temps sont hors cadre est de 3252 ms (sd=158).

#### **A- Premier segment de la phrase cible**

##### **Interaction entre les facteurs « Nombre » et « Verbe »**

L'interaction entre les facteurs « Nombre » et « Verbe » est significative (Chisquare= 6.27, df=1,  $p= 0.01$ ) : Tableau 21 et Figure 10.

F-M-First-I : Tps de lecture du 1 <sup>er</sup> segment des phrases cibles	Reprises au pluriel	Reprises au singulier
Prédicat symétrique	932 (61)	879 (47)
Prédicat Non symétrique	917 (60)	982 (73)

Tableau 21 : Temps de lecture des premiers segments des phrases cibles en fonction de la reprise et de la phrase amorce – regroupement des données de FFirst-I et MFirst-I

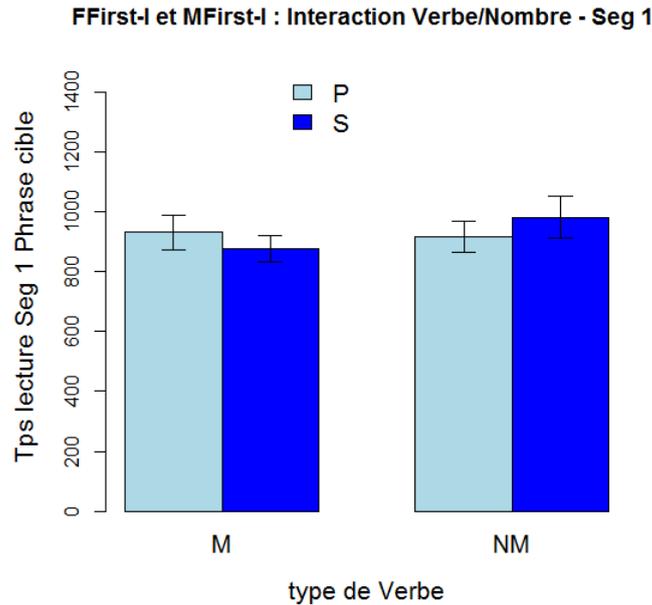


Figure 10 : Temps de lecture des premiers segments des phrases cibles en fonction de la reprise et de la phrase amorce – regroupement des données de FFirst-I et MFirst-I

Les premiers segments des phrases cibles sont lus plus lentement (Chisquare= 7.9, df=1, p= 0.005) lorsque la reprise concerne le seul référent masculin après les phrases amorces non symétriques (982 ms) relativement aux autres conditions (reprises plurielles après les phrases amorces symétriques et non symétriques et reprises au singulier après les phrases amorces symétriques). Ce résultat correspond à l'effet trouvé dans MFirst-I dans le groupe S/C.

## **B- Deuxième segment de la phrase cible**

### **Facteur « Nombre »**

Le deuxième segment des phrases cibles est plus rapide à lire (Chisquare= 4.6, df=1, p= 0.03) pour les reprises des deux référents (1052 ms, sd = 81) que pour les reprises du seul référent masculin (1114 ms, sd=84). Cet effet a été trouvé dans le module FFirst-I : il n'était pas présent dans MFirst-I. Ce résultat est dû à la lenteur des traitements liés à l'extraction du référent régi dans le module « NP féminin en tête ».

### **Interaction entre les facteurs « TETE » et « Nombre »**

L'interaction entre les facteurs « TETE » et « Nombre » montre que les reprises au pluriel (961 ms) sont traitées plus rapidement (Chisquare= 12.44, df=1, p<0.001) que les reprises au singulier (1065 ms) lorsque le NP féminin est en tête. Ce résultat duplique la significativité du facteur

« Nombre » dans FFirst-I sur ce deuxième segment. De plus, l'interaction montre que les reprises plurielles avec un NP féminin en tête (961 ms) sont aussi plus rapides (Chisquare= 6.12, df=1, p= 0.013) à traiter que les reprises plurielles avec un NP masculin en tête (1059 ms) : Tableau 22 et Figure 11.

M-F-First-I : Tps de lecture du 2 <sup>ème</sup> segment des phrases cibles	Reprises au pluriel	Reprises au singulier
NP féminin en tête	961 (71)	1065 (82)
NP masculin en tête	1159 (90)	1170 (85)

Tableau 22 : Temps de lecture des deuxièmes segments des phrases cibles en fonction du « Nombre » des reprises et du NP en tête – regroupement des données de FFirst-I et MFirst-I

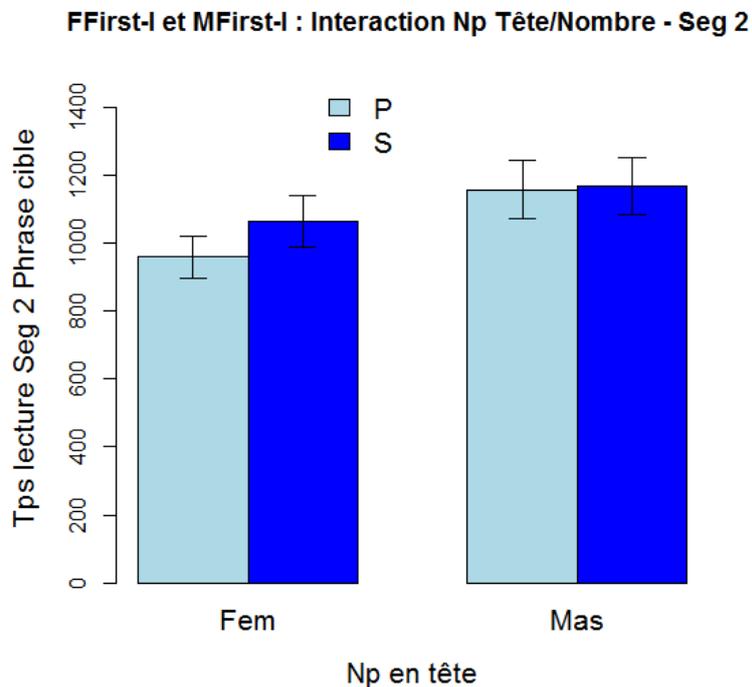


Figure 11 : Temps de lecture des deuxièmes segments des phrases cibles en fonction du « Nombre » des reprises et du NP en tête - regroupement des données de FFirst-I et MFirst-I

Ce dernier résultat montre que la formation d'une entité plurielle avec les référents dispersés des constructions comitatives est plus facile lorsque le NP féminin est en tête que lorsque le NP masculin est en tête. La prééminence de ce référent modifie les résolutions des pronoms singuliers et pluriels, tous prédicats confondus.

## **C- Troisième segment de la phrase cible**

### **Facteur contraste**

On trouve un effet principal du facteur contraste dans ce troisième segment. Les traitements dans le groupe contraste S/C (1180 ms, sd=81) sont plus rapides (Chisquare= 9.8, df=1, p= 0.001) que ceux du groupe contraste S/D (1309 ms, sd=93). Cet effet était présent dans MFrst-I. Il indique que 1) les prédicats non symétriques des deux groupes de contraste ont des propriétés plurielles réellement différentes et 2) que les traitements anaphoriques peuvent se prolonger bien au-delà du segment pronominal.

## **D- Phrase cible**

### **Facteur « Verbe »**

On trouve un effet principal du facteur « Verbe » (Chisquare= 6.26, df=1, p= 0.012) : les traitements pronominaux après des phrases amorces symétriques (3195 ms, sd=147) sont plus rapides qu'après des phrases amorces non symétriques (3309 ms, sd= 169) sur la phrase cible.

### **Facteur « Nombre »**

On trouve aussi un effet principal du facteur « Nombre » (Chisquare= 4.6, df=1, p= 0.03) : les reprises des deux référents sont résolues (3203 ms, sd=156) plus rapidement que les reprises du seul référent masculin (3308 ms, sd= 161) sur la phrase cible.

### **Facteur « contraste »**

Enfin, on observe un effet principal du facteur « contraste » (Chisquare= 5.53, df=1, p= 0.018) : les traitements pronominaux du groupe de contraste S/C (3155 ms, sd=154) sont plus rapides que ceux du groupe de contraste S/D (3350 ms, sd= 162) sur la phrase cible.

### **Interaction entre les facteurs « TETE » et « Contraste »**

L'interaction (Chisquare= 4.01, df=1, p= 0.045) entre « TETE » et groupe de contraste montre que dans le module où le NP masculin est en tête, les temps de lecture des phrases cibles sont équivalents dans les deux groupes de contraste, alors que dans le module où le NP féminin est en tête, les temps de lecture sont plus courts (Chisquare= 9.6, df=1, p= 0.002) dans le groupe de contraste S/C (2971 ms) que dans le groupe de contraste S/D (3254 ms) : Tableau 23 et Figure 12.

F-M-First-I : Tps de lecture des phrases cibles	Groupe S/C	Groupe S/D
NP féminin en tête	2971 (145)	3254 (168)
NP masculin en tête	3379 (160)	3457 (154)

Tableau 23: Temps de lecture des phrases cibles en fonction du NP de tête et du groupe contraste - regroupement des données de FFirst-I et MFirst-I

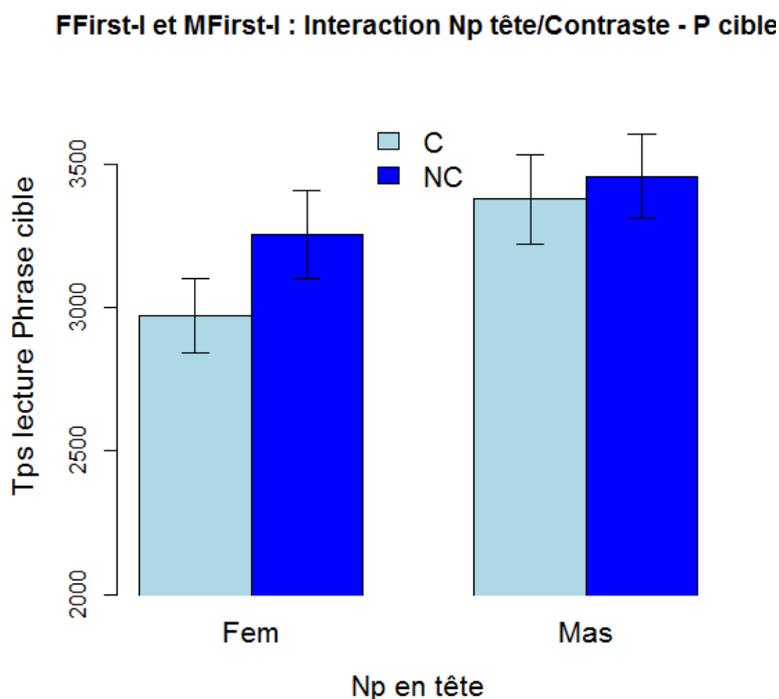


Figure 12 : Temps de lecture des phrases cibles en fonction du NP de tête et du groupe contraste - regroupement des données de FFirst-I et MFirst-I

### **Interaction entre les facteurs « TETE » et « Nombre »**

L'interaction (Chisquare= 4.33, df=1, p= 0.037) entre « TETE » et « Nombre » de la reprise montre qu'avec un NP masculin en tête, les temps de lecture des phrases cibles sont équivalents pour les reprises au pluriel et au singulier, alors qu'avec un NP féminin en tête, les reprises au pluriel sont traitées plus rapidement (Chisquare= 9.02, df=1, p= 0.003) que les reprises au singulier : Tableau 24 et Figure 13.

F-M-First-I : Tps de lecture de des phrases cibles	Reprises au pluriel	Reprises au singulier
NP féminin en tête	3010 (149)	3220 (165)
NP masculin en tête	3428 (159)	3408 (155)

Tableau 24 : Temps de lecture des phrases cibles en fonction du NP de tête et de la reprise - regroupement des données de FFirst-I et MFirst-I

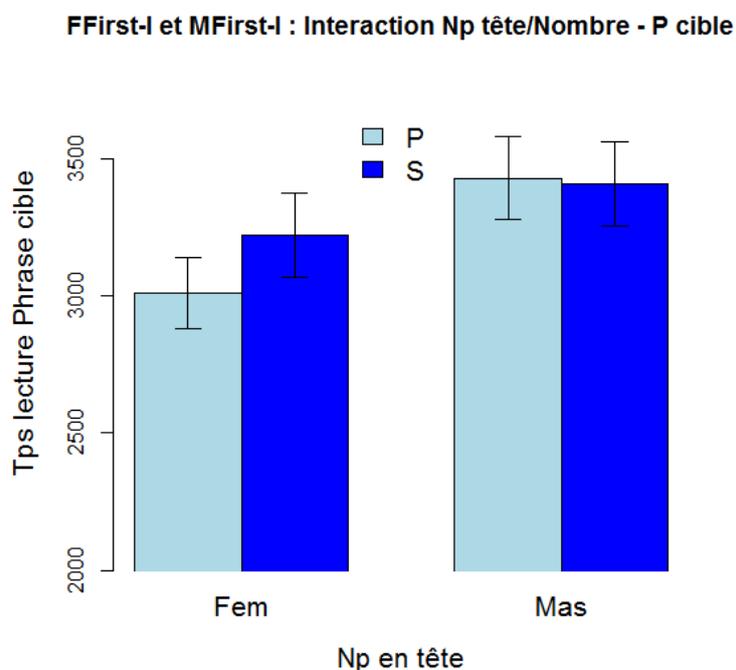


Figure 13 : Temps de lecture des deuxièmes segments des phrases cibles en fonction du NP de tête et de la reprise - regroupement des données de FFirst-I et MFirst-I

De plus, les reprises au pluriel tendent à être traitées plus vite dans le module où le NP féminin est en tête (Chisquare= 3.5, df=1, p= 0.06) que dans le module où le NP masculin est en tête.

Ces deux dernières interactions vont dans le même sens : différences entre les facteurs « Contraste » et « Nombre » dans le module où le NP féminin est en tête, mais pas d'effets de ces facteurs dans le module où le NP masculin est en tête. Le seul facteur qui a un effet dans les deux modules du test est le facteur « Verbe ».

Ces résultats valident l'hypothèse d'un effet des propriétés verbales de symétrie sur les enchaînements discursifs.

## **E- Temps de réponse**

### **Facteur « Nombre »**

Le facteur « Nombre » a un effet significatif sur les temps de réponse ; ils sont plus longs (Chisquare=7.48, df=1, p=0.006) lorsque les deux référents (5177 ms, sd= 642) sont repris dans la phrase cible que lorsque le seul référent est repris (4365 ms, sd=460).

### **Facteur « Verbe »**

Le facteur « Verbe » a un effet significatif sur les temps de réponse ; ils sont plus courts (Chisquare=4.38, df=1, p=0.036) en cas de prédicats amorces symétriques (4490 ms, sd=512) qu'en cas de prédicats amorces non symétriques (5026 ms, sd= 611).

## **4.4.3 Discussion**

**Lorsque le NP en tête de phrase amorce est féminin**, on trouve un effet principal du facteur « Nombre » : les phrases cibles qui reprennent les deux référents sont lues plus rapidement que les phrases cibles qui reprennent le référent régi. Ce résultat est conforme aux résultats du pré-test. La triple interaction trouvée entre les facteurs « Contraste », « Verbe » et « Nombre » donne des précisions sur les conditions qui favorisent les reprises plurielles.

Dans le groupe contraste S/C les facteurs « Nombre » et « Verbe » ne sont pas significatifs, ce qui veut dire que dans le groupe des symétriques/collectifs la formation d'une entité plurielle est aussi facile que l'extraction du référent régi. Ce résultat permet de vérifier plusieurs hypothèses 1) que la préposition comitative *avec* a bien une fonction coordonnante avec les prédicats collectifs dans les constructions comitatives, 2) que l'appariement des arguments des prédicats symétriques permet leur reprise en tant que collectif, 3) que la relation de symétrie créée entre les arguments des prédicats symétriques est bien une relation d'équivalence, puisqu'elle rend accessible le référent régi. Par ailleurs, après les prédicats collectifs, les temps de formation de la pluralité à partir des arguments dispersés et ceux nécessaires pour extraire l'argument régi ne diffèrent pas : la fonction coordonnante de la préposition *avec* ne va pas jusqu'à générer un coût d'extraction de l'antécédent du groupe formé par les arguments dispersés, lorsque le référent repris est le référent régi (ce n'est pas le cas lorsque le référent repris est proéminent : voir ci-dessous).

Dans le groupe de contraste S/D, on trouve que globalement les reprises plurielles sont traitées plus rapidement que les reprises du référent régi. L'interaction des facteurs « Verbe » et « Nombre » montre toutefois que ce sont les reprises au singulier derrière les prédicats distributifs qui sont lentes à traiter. Ce résultat vérifie l'hypothèse que nous avons faite sur les difficultés d'extraction de l'argument régi dans les constructions comitatives distributives. Lorsque le NP en tête de phrase amorce est féminin, la significativité globale du facteur « Nombre » sur la phrase cible est due par conséquent aux difficultés d'extraction des référents régis dans les constructions comitatives distributives.

Nos résultats ne dupliquent les résultats de Hielscher & Musseler (1990) que pour les prédicats distributifs : les phrases cibles qui reprennent les deux référents sont lues plus rapidement que celles qui reprennent le référent régi. En revanche, pour les prédicats amorces symétriques et collectifs, le « Nombre » de la reprise n'a pas d'influence sur les temps de lecture.

**Lorsque le NP en tête de phrase amorce est masculin**, en revanche, les résultats obtenus ne rendent pas compte de la grande accessibilité du référent masculin : la formation d'une entité plurielle et la résolution des reprises au singulier ont des durées de traitement équivalentes, excepté après les phrases amorces symétriques

Dans le premier segment, les reprises du référent proéminent ne sont lues plus rapidement que derrière les prédicats amorces symétriques, comme le montre l'interaction entre les facteurs « Nombre » et « Verbe ». La triple interaction entre les facteurs « Contraste », « Nombre » et « Verbe », donne les précisions suivantes : les reprises du référent proéminent causent un ralentissement par rapport aux reprises au pluriel derrière les prédicats amorces collectifs. Compte tenu de la fonction coordonnante de la préposition *avec* dans les constructions comitatives collectives, ce résultat correspondrait au « conjunction cost » des pluriels coordonnés lorsqu'un de ses membres est repris pronominalement. Il faut cependant expliquer pourquoi, ce coût cognitif n'existe pas lorsque le pronom de reprise porte sur le référent régi : le cadre théorique du « Structure Building Framework » de Gernsbacher (1986) serait une piste à approfondir. En comparaison, les prédicats amorces symétriques permettent une lecture rapide des premiers segments reprenant le référent saillant dans le groupe S/C. Dans le groupe contraste Symétrique/Distributif, rien n'est significatif : les premiers segments reprenant les deux référents sont lues à la même vitesse que ceux reprenant le référent proéminent et cela derrière les prédicats symétriques et distributifs : dans ce premier segment, donc les reprises au singulier et au pluriel

derrière des prédicats amorces symétriques et distributifs ne se différencient au niveau des vitesses de traitement.

Dans le deuxième segment, l'interaction des facteurs « Contraste » et « Verbe » indique que les temps de lecture après les prédicats amorces symétriques ne dépendent pas du groupe contraste (résultat identique dans le pré-test) : en revanche, ils sont plus longs après les prédicats distributifs. Par ailleurs, une autre interaction, l'interaction des facteurs « Contraste » et « Nombre » indique les reprises du référent proéminent sont ralenties relativement aux reprises plurielles derrière les prédicats amorces du groupe S/D, mais pas derrière les prédicats amorces du groupe S/C. Ainsi, la reprise d'un référent proéminent derrière un prédicat distributif serait plus longue à traiter qu'une reprise des référents dispersés de constructions comitatives. Ce résultat est inattendu. Un retour sur les résultats du premier test apporte un éclairage :

- Les données regroupées des deux modules du premier test ont montré que les reprises du référent masculin sont traitées plus lentement quand il est régi que quand il est en tête de phrase amorce,
- Elles ont aussi montré que, quand le référent repris est en tête de phrase amorce, il n'y avait aucun effet des facteurs « Contraste », « Nombre » et « Verbe » sur les temps de lecture des phrases cibles.

Ces résultats montrent que dans le premier test, les résolutions des pronoms au singulier sont plus rapides quand l'antécédent est proéminent que quand il est régi par le comitatif ; mais ils montrent aussi que les reprises plurielles sont aussi rapides à résoudre que les reprises du référent saillant lorsqu'il est en tête. L'équivalence des temps de résolution des reprises au pluriel et au singulier lorsque l'antécédent est saillant suggère que la formation d'un pluriel à partir des référents sujet et régi d'une construction comitative n'a pas de coût cognitif, ce qui pourrait s'expliquer par la symétrie syntaxique de cette construction. Mais, on ne peut exclure que ce résultat soit dû à un biais des textes narratifs de l'expérience : tous construits autour de deux protagonistes dont les trois quarts étaient introduits en relation de symétrie ou de co-agentivité. Les reprises au pluriel prolongeaient un pattern narratif où le focus était mis sur le groupe formé par les protagonistes, plutôt que sur les individus en présence : la modification apportée sur l'argument externe des questions en fin d'item a accentué le phénomène en focalisant encore plus l'attention des lecteurs sur le groupe.

Dans le premier test, certes il n'y avait pas de différence entre les résolutions des pronoms au singulier et au pluriel de la phrase cible en cas de proéminence du référent masculin ; cependant, les résolutions des reprises de ce référent étaient plus rapides en cas de proéminence qu'en cas de non proéminence. Dans le second test, les temps de traitements des pronoms singuliers deviennent plus longs, dans le deuxième segment, en cas de proéminence qu'en cas de non proéminence du référent. A côté des aménagements expérimentaux, consistant à segmenter les phrases et à faire précéder la phrase amorce d'une introduction, la modification de l'argument externe de la question est une probable cause des résultats trouvés dans MFirst-I. En effet, dans le module où le référent masculin est en tête de phrase amorce, il occupe tour à tour la fonction de sujet de la phrase amorce, puis celle d'antécédent du pronom dans les phrases cibles au singulier et enfin de premier membre de la coordination argument externe des questions. Si la position de tête des coordinations est leur point d'ancrage (Gernsbacher et al., 1988), le référent masculin devait donc passer d'un statut de focus en tant qu'individu à un statut de représentant d'une coordination dans la question. La répétition des questions à sujet coordonné a pu créer un biais de lecture au niveau des seconds segments. On peut exprimer cette possible explication en disant que dans les items de test, il y a un déplacement du focus vers le groupe que forment les deux référents de la construction comitative : ce déplacement est sans coût cognitif après les phrases amorces collectives et symétriques mais demandent des traitements supplémentaires après des phrase amorces distributives. Cette hypothèse demanderait à être testée. Elle permettrait de comprendre pourquoi les résolutions des pronoms singuliers derrière les phrases amorces collectives sont plus rapides que les résolutions des pronoms singuliers derrière les phrases amorce distributives.

Les effets trouvés dans chacun des segments dans le module où les NP masculin est en tête donnent des indications sur **les phases de traitements des reprises pronominales**. Dans le premier segment, les temps de lecture dépendent des représentations des référents induites par le noyau prédicatif des phrases amorces. Les reprises du référent masculin après un prédicat amorce collectif, ralentissent la lecture relativement aux prédicats amorces symétriques et distributifs. Les collectifs orientent l'attention vers le groupe formé par les référents : ils nécessiteraient un traitement spécifique lorsqu'un seul référent est repris. Les collectifs se distingueraient des symétriques dans ce module parce que le référent repris est proéminent : les symétriques accèdent immédiatement à ce référent contrairement aux collectifs. Lorsque le référent repris n'est pas proéminent, en revanche, il est autant accessible derrière un prédicat collectif que symétrique.

Dans le deuxième segment, les processus concerneraient des traitements sémantiques où les participants chercheraient à inférer le lien entre les référents dispersés, aisés en cas de prédicats

amorces symétriques et collectifs, plus compliqués en cas de prédication distributive : cette élaboration absente du premier test serait causée par la forme des questions qui ont pour argument externe le pluriel coordonné formé par les référents dispersés.

Dans le troisième segment, le facteur « Contraste » est significatif indiquant que les traitements entrepris dans le deuxième segment se prolongent.

Les temps de lecture de la phrase cible sont plus courts pour les prédicats symétriques que les prédicats non symétriques ; ils sont également plus courts pour les reprises plurielles que pour les reprises au singulier ; ils sont enfin plus courts dans le groupe contraste S/C que le groupe de contraste S/D. L'interaction entre le groupe de contraste et le NP de tête indique que lorsque le NP de tête est masculin, les traitements des reprises pronominales sont équivalents dans les deux groupes : l'effet principal du facteur contraste s'explique en partie par des temps de lecture de la phrase cible plus courts dans le groupe S/C lorsque le NP féminin est en tête. L'interaction entre les facteurs « Nombre » et « TETE » montre que l'effet principal du facteur « Nombre » est en partie dû au module où le NP féminin est en tête : les prédicats à interprétation collective favorisent les reprises plurielles dans ce module.

L'effet principal du facteur « Verbe » trouvé sur la phrase cible montre que les **prédicats symétriques se caractérisent** par un accès au référent régi rapide, un traitement des reprises du référent saillant rapide et un regroupement des référents dispersés rapide. Ces résultats sont en accord avec les propriétés plurielles de ces prédicats : ils appartiennent leurs arguments en un ensemble qui peut constituer l'antécédent naturel d'une reprise plurielle, tout en les distinguant suffisamment pour que chacun puisse être repris individuellement. Ces résultats valident l'hypothèse d'une équivalence des arguments en relation de symétrie. Ils confortent les résultats de Patson et Ferreira (2009) qui montrent que les prédicats symétriques traitent leurs arguments comme « un collectif dont les éléments sont visibles ».

Par ailleurs, ce deuxième test interrogeait sur la participation des deux protagonistes de la narration à l'évènement décrit dans la phrase cible et dans **la question**. Le pluriel coordonné regroupant les deux référents, avec le référent masculin en tête, était argument externe des questions : elle respectait donc l'ordre d'introduction des référents dans le module MFirst-I mais pas dans le module FFirst-I. Dans le module FFirst-I, les facteurs expérimentaux n'ont pas d'effet sur les temps de réponse à la question ; dans le module MFirst-I, on trouve un effet du facteur « Nombre » et une interaction entre les facteurs « Contraste » et « Nombre ». Ce résultat suggère que l'ordre

d'introduction des référents est un critère de représentation des pluriels coordonnés : l'inversion de la position des référents dans la coordination de la question, dans FFirst-I, annule l'effet des facteurs expérimentaux.

Dans MFirst-I, on trouve un effet du facteur « Nombre » : les phrases cibles au singulier entraînaient des temps de réponse à la question plus courts que les phrases cibles au pluriel. On trouve aussi une interaction entre les facteurs « Contraste » et « Nombre » : les temps de réponse sont plus longs pour les phrases cibles au pluriel dans le groupe S/C, le groupe des collectifs donc. Dans le groupe de contraste S/D, ils sont équivalents pour les reprises au singulier et au pluriel de la phrase cible. Ces résultats sur les temps de réponse à la question, montrent que les référents dispersés d'une construction comitative à interprétation collective qui ont fait l'objet d'une reprise, sont peu disponibles pour participer à la formation d'une nouvelle entité plurielle, moins disponibles, en tout cas, que les référents dispersés d'une construction comitative distributive. Cela signifie que les traitements supplémentaires dans le deuxième segment de la phrase cible dans la condition « reprise du référent proéminent/prédicats amorces distributifs », même s'ils portent sur la construction de liens entre les référents dispersés, préserveraient leur statut d'individus. Au contraire dans le groupe S/C, la représentation du collectif resterait prégnante en cas de reprise plurielle dans la phrase cible : les temps de réponse à la question dénoteraient alors des traitements liés à la décomposition du collectif formé précédemment. Le fait que les phrases cibles au singulier ne causent pas de ralentissement des temps de réponse à la question dans le groupe S/C montrent que les phrases cibles au singulier ont réalisé cette décomposition : les constituants du collectif seraient alors disponibles pour former une autre entité plurielle. La dénomination d'une entité plurielle présente en mémoire nécessiterait donc des traitements supplémentaires par rapport à la dénomination d'une entité plurielle sans antériorité, sauf en ce qui concerne les arguments de prédicats symétriques. Ces résultats sont compatibles avec une persistance en mémoire des entités plurielles en relation de symétrie au cours des items et une persistance également des entités plurielles formant un collectif si le développement narratif n'a pas désolidarisé les membres du groupe (reprise d'un seul référent dans la phrase cible).

Le prochain test portant sur les pluriels coordonnés, on se propose de conserver le pattern des questions. Les pluriels coordonnés introduits dans les phrases amorces rendent leurs constituants solidaires sur le plan des rôles et des fonctions : les biais discursifs observés dans les constructions comitatives n'existent plus. Seule la relation entre un objet référentiel complexe et les propriétés prédictives compte dans les enchaînements pronominaux suivants. La persistance des entités plurielles en mémoire peut être étudiée au niveau des formes de surface et des représentations

sémantiques : les pluriels coordonnés dans les questions seront soit identiques au pluriel coordonné introduit dans les phrases amorces soit verront les positions de leurs membres inversées.

Pour conclure, on retiendra que la plupart de nos hypothèses sont validées par les résultats : la facilité de regroupement des référents dispersés après des prédicats à interprétation collective, la fonction coordonnante du comitatif avec ces prédicats, la difficulté d'accès au référent régi dans les prédicats distributifs, la proximité entre prédicats symétriques et collectifs lorsque le référent repris pronominalement est régi par le comitatif, la différence entre ces mêmes prédicats lorsque le référent repris est proéminent. La rapidité de résolution des pronoms au singulier lorsqu'ils reprennent un référent proéminent ne se vérifie pas, en revanche, ce qui a été interprétée comme un glissement de focus d'entités de niveau individu vers des entités de niveau groupe. L'ensemble de ces résultats montrent que des paramètres discursifs et narratifs interagissent avec les propriétés plurielles des prédicats dans les reprises pronominales, moins avec les prédicats symétriques qu'avec les prédicats collectifs et distributifs, cependant.

## **4.5 Test 3**

### **4.5.1 Plan expérimental**

#### **A- Matériel et Facteurs expérimentaux**

Ce test a pour objectif d'étudier les enchaînements pronominaux après des phrases amorces symétriques et distributives dans des constructions intransitives. Les items du deuxième test sont modifiés en conséquence : les référents sujet et régi sont regroupés pour former un pluriel coordonné, sujet grammatical des constructions. Pour prendre en compte les résultats de Gernsbacher et al. (1988) sur l'avantage du premier nommé, deux modules de test sont réalisés : les objets référentiels complexes commencent par le NP féminin dans l'un (FFirst-I-et) et par le NP masculin dans l'autre (MFirst-I-et). La segmentation en trois des phrases amorces et cibles est conservée, ainsi que la segmentation en deux de la phrase tampon.

L'interrogation en fin d'items porte sur la participation des deux référents à l'évènement décrit par la phrase cible, comme dans le deuxième test. Cependant, le référent en tête de la coordination est le référent féminin dans ce troisième test. Suivant l'interprétation des résultats obtenus dans le deuxième test, le référent masculin en tête de coordination dans la question serait une des raisons possibles des biais de lecture des phrases cibles : en effet, dans le module où le NP masculin est en tête (MFirst-I), le référent masculin occupait tour à tour la fonction de sujet de la phrase amorce, puis celle d'antécédent du pronom dans les phrases cibles au singulier. Lors de la question, il devait être coordonné avec le référent féminin : si la position de tête des coordinations est leur point d'ancrage (Gernsbacher et al., 1988), il devait donc passer d'un statut de focus en tant qu'individu à un statut de représentant d'une coordination. Plus généralement, le référent masculin dans ces items de test est surreprésenté du fait que les antécédents des pronoms de reprise alternent entre ce référent et la pluralité formée par les référents masculin et féminin. Prendre le référent féminin comme premier nommé de la coordination dans les questions permet de contraster le module où la coordination sujet des phrases amorces et celle sujet de la question ont le même premier nommé (FFirst-I-et) et le module où les premiers nommés des deux coordinations sont différents (MFirst-I-et). Ainsi, on teste d'une part l'importance de cette position et d'autre part les modalités d'interprétation d'un pluriel coordonné selon qu'il est la répétition d'une entité au niveau sémantique comme au niveau de son expression de surface ou qu'il est la répétition au niveau sémantique mis pas au niveau de sa forme de surface. Comme dans le test précédent, les réponses attendues sont positives pour un pronom pluriel dans la phrase cible et négatives sinon. Comme dans le deuxième test, la phrase d'introduction et l'interrogation en fin d'items ne sont pas segmentées.

Les facteurs expérimentaux sont les facteurs « Verbe » et « Nombre » uniquement : en effet, les résultats expérimentaux exposés en 3.1.2 montrent qu'un objet référentiel complexe est appréhendé comme un collectif sauf lorsqu'il est argument d'un prédicat symétrique. La différence entre prédicats distributifs interprétés collectivement et les autres n'a plus lieu d'être. Un exemple d'item du module où le pluriel coordonné commence par le NP féminin (module FFIRST-I-et) est en tête est donné en (225a-j).

(225a) *La montagne n'est pas que le royaume de la glisse au grand air.*

(225b) *Audrey et Robin*

(225c) *cohabitent/dépriment à Courchevel*

(225d) *dès l'ouverture de la station.*

(225e) *Ils/il fréquente(nt)*

(225f) *tous les lieux branchés*

(225g) *du centre-ville.*

(225h) *Les célébrités*

(225i) *ne manquent pas dans le coin.*

(225j) ?? *Robin et Audrey rencontrent des célébrités à Courchevel.*

Chaque module de test comporte 24 textes déclinés en quatre versions selon les conditions<sup>52</sup>. Les 4 versions de chacun des 24 textes sont répartis sur 4 groupes en contrebalançant les items de façon à ce qu'il y ait le même nombre d'items par conditions dans chaque groupe (6) et qu'un même texte n'apparaisse dans un groupe que dans une version (voir le plan expérimental en annexe 6). Les items expérimentaux sont complétés par 24 textes distracteurs (voir les distracteurs utilisés en annexe 5) qui sont les mêmes dans les quatre groupes.

Le module MFirst-I-et suit le même plan expérimental : les items de test de FFirst-I-et sont repris en permutant les prénoms féminin et masculin (dans les textes, pas dans les questions).

## B- Procédure

La méthodologie de présentation des textes est l'APS (Auto Présentation segmentée), comme précédemment : la phrase d'introduction s'affiche au centre de l'écran dans son entier, puis le premier segment suivi du deuxième segment, suivi lui-même du troisième segment de la phrase amorce, puis le premier segment suivi du deuxième segment, suivi lui-même du troisième segment de la phrase cible, puis le premier segment suivi du deuxième segment de la phrase tampon, puis la question qui s'affiche en entier.

## C- Hypothèses

Les hypothèses de test sont qu'après les phrases amorces symétriques, la reprise du référent masculin devrait être traitée plus rapidement qu'après les phrases amorces mixtes. Aucune différence n'est attendue pour les reprises plurielles. Concernant la possible influence du NP en tête de pluriel coordonné, aucune prédiction n'est faite eu égard aux résultats de Gordon et al. (1999) qui n'a pas trouvé de différence dans les reprises du premier ou du deuxième référent d'une paire coordonnée contrairement à Gernsbacher et al. (1988).

---

<sup>52</sup> NB : Certains contenus textuels du deuxième test ont été modifiés pour les rendre plus naturels.

## D- Conditions de passation

Trente-cinq participants ont passé le module FFirst-I-et et trente-cinq autres le module MFirst-I-et dans des conditions identiques à celles des tests 1 et 2.

## 4.5.2 Analyses et résultats

### 4.5.2.1 Module NP féminin en tête : FFirst-I-et

Les temps de lecture des phrases cibles sont calculés en additionnant les temps de lecture de chaque segment des phrases cibles.

Le « nettoyage » des temps de lecture est effectué sur le temps de lecture global des phrases cibles, conformément à la méthodologie d'analyse adoptée dans les tests 1 et 2.

Les participants S11 et S21 qui ont fait plus de 33% d'erreurs ne sont pas retenus. De plus, S18 est de langue maternelle non française. Les temps de lecture de chaque segment sont traités par la fonction « scale » de R qui calcule le z-score du groupe sur une variable dépendante et signale les outsiders : à la suite de quoi les participants S7, S14, S19, S22, S27, S31, S34, S35 sont écartés des analyses. Restent 24 participants.

La moyenne des temps de lecture bruts de la phrase cible (P2RT) est de 3473 ms (sd=254).

La moyenne des temps de lecture nettoyés de la phrase cible (P2RT) est de 3300 ms (sd=258).

La moyenne des temps de lecture de la phrase cible (P2RT) après élimination des mauvaises réponses et des participants dont les temps sont hors cadre est de 3093 ms (sd=214).

Les analyses conduites avec le modèle linéaire mixte du logiciel R (fonction « lmer » de la librairie lme4) suivant une loi normale, portent sur les temps de lecture « nettoyés » des 24 participants restants.

### A- Premier segment des phrases cibles

Il n'y a aucun effet sur ce premier segment.

## **B- Deuxième segment des phrases cibles**

### **Facteur « Verbe »**

On trouve un effet principal du facteur « Verbe » sur les temps de lecture du deuxième segment des phrases cibles.

Le deuxième segment des phrases cibles est lu plus rapidement (Chisquare=6.4, df=1, p= 0.01) après les phrases amorces symétriques (1008 ms, sd=100) qu'après les phrases amorces non symétriques (1151 ms, sd=124).

## **C- Troisième segment des phrases cibles**

Il n'y a aucun effet sur le troisième segment.

## **D- Phrase cible**

### **Facteur « Verbe »**

On trouve un effet principal du facteur « Verbe » sur les temps de lecture de la phrase cible.

Les phrases cibles sont lues plus rapidement (Chisquare=6.17, df=1, p= 0.013) après les phrases amorces symétriques (2977 ms, sd=210) qu'après les phrases amorces non symétriques (3205 ms, sd=216).

## **E- Temps de réponse**

### **Facteur « Nombre »**

On trouve un effet principal du facteur « Nombre » (Chisquare=4.6, df=1, p= 0.032). Les temps de réponse sont ralentis quand la reprise est au pluriel dans la phrase cible : 3506 ms (sd=307) pour les pronoms au pluriel contre 3239 ms (sd=312) pour les pronoms au singulier.

### **4.5.2.2 Module NP masculin en tête : MFirst-I-et**

Les temps de lecture des phrases cibles sont calculés en additionnant les temps de lecture de chaque segment des phrases cibles.

Le « nettoyage » des temps de lecture est effectué sur le temps de lecture global des phrases cibles, conformément à la méthodologie d'analyse adoptée dans les premier et deuxième.

Les participants S2, S23, S25, S31 qui ont fait plus de 33% d'erreurs ne sont pas retenus. De plus, les temps de lecture de chaque segment sont traités par la fonction « scale » de R qui calcule le z-score du groupe sur une variable dépendante et signale les outsiders : à la suite de quoi les participants S4, S14, S22, S18, S27, S28, S35 sont écartés des analyses. Restent 24 participants.

La moyenne des temps de lecture bruts de la phrase cible (P2RT) est de 3450 ms (sd=308). La moyenne des temps de lecture nettoyés de la phrase cible (P2RT) est de 3170 ms (sd=213). La moyenne des temps de lecture de la phrase cible (P2RT) après élimination des mauvaises réponses et des participants dont les temps sont hors cadre est de 2960 ms (sd=195).

Les analyses conduites avec le modèle linéaire mixte du logiciel R (fonction « lmer » de la librairie lme4) suivant une loi normale, portent sur les temps de lecture « nettoyés » des 24 participants restants.

## **A- Premier segment de la phrase cible**

### **Facteur « Nombre »**

On trouve un effet principal du facteur « Nombre » sur les temps de lecture des 1<sup>ers</sup> segments des phrases cibles.

Ces derniers sont lus plus rapidement (Chisquare=5.29, df=1, p= 0.021) pour les reprises du pluriel coordonné sujet de la phrase amorce (840 ms, sd=69) que pour les reprises d'un élément de ce pluriel, en l'occurrence le référent masculin (929 ms, sd=90).

## **B- Deuxième segment de la phrase cible**

### **Facteur « Nombre »**

On trouve à nouveau un effet principal du facteur « Nombre » sur les temps de lecture des 2<sup>èmes</sup> segments des phrases cibles.

Ces derniers sont lus plus rapidement (Chisquare=6.62, df=1, p= 0.01) pour les reprises du pluriel coordonné sujet de la phrase amorce (992 ms, sd=89) que pour les reprises d'un élément de ce pluriel, en l'occurrence le référent masculin (1110 ms, sd=101).

## **C- Troisième segment de la phrase cible**

Il n'y a aucun effet sur le troisième segment de la phrase cible.

## **D- Phrase cible**

### **Facteur « Nombre »**

Les phrases cibles sont lues significativement plus rapidement (Chisquare= 10.14, df=1,  $p > 0.01$ ) pour les reprises du pluriel coordonné sujet de la phrase amorce (2839 ms, sd=181) que pour les reprises d'un élément de ce pluriel, en l'occurrence le référent masculin (3074 ms, sd=206).

## **E- Temps de réponse**

### **Facteur « Nombre »**

On trouve un effet principal du facteur « Nombre » (Chisquare=4.04, df=1,  $p = 0.044$ ). Les temps de réponse sont ralentis quand la reprise était au pluriel dans la phrase cible : 3455 ms (sd=411) pour les pronoms au pluriel contre 3156 ms (sd=384) pour les pronoms au singulier.

### **4.5.2.3 Analyses des données regroupées**

Le fait de grouper les données implique d'introduire dans les analyses un nouveau facteur, le facteur « TETE » à 2 modalités : NP féminin ou NP masculin en tête de phrase.

Les étapes de l'analyse sont les mêmes précédemment.

Les participants écartés dans les tests FFirst-I-et et MFirst-I-et, sont également écartés de cette analyse. Restent 48 participants.

La moyenne des temps de lecture bruts de la phrase cible (P2RT) est de 3462 ms (sd=241). La moyenne des temps de lecture nettoyés de la phrase cible (P2RT) est de 3235 ms (sd=183). La moyenne des temps de lecture de la phrase cible (P2RT) après élimination des mauvaises réponses et des participants dont les temps sont hors cadre est de 3093 ms (sd=160).

## **A- Premier segment de la phrase cible**

### **Facteur « Nombre »**

Le facteur « Nombre » est significatif (Chisquare= 4.85, df=1,  $p = 0.027$ ) : les premiers segments de la phrase cible sont plus rapides à lire quand ils commencent par une reprise du pluriel coordonné (884 ms, sd=62) que quand ils commencent par une reprise du référent masculin (938 ms, sd=67).

## **B- Deuxième segment de la phrase cible**

### **Facteur « Nombre »**

Il y a un effet principal du facteur « Nombre » (Chisquare= 6.95, df=1, p= 0.008) et une interaction significative entre les facteurs « TETE » et « Verbe » (Chisquare= 5, df=1, p= 0.025).

Les deuxièmes segments sont plus rapides à lire quand la phrase cible commence par un pronom pluriel (1050 ms, sd=76) que quand il commence par un pronom singulier (1136 ms, sd=84).

### **Interaction entre les facteurs « TETE » et « Verbe »**

L'interaction entre les facteurs « TETE » et « Verbe » montre que lorsque le NP en tête est féminin, les deuxièmes segments des phrases cibles sont plus rapides (Chisquare= 6.5, df=1, p= 0.01) à lire après une phrase amorce symétrique qu'après une phrase amorce non symétrique : ce n'est pas le cas lorsque le NP de tête est masculin. Les temps de lecture des deuxièmes segments après les phrases amorces symétriques sont toutefois équivalents dans les modules FFirst-I-et et MFirst-I-et : Tableau 25 et Figure 14.

M-F-First-I-et : Tps de lecture du 2 <sup>ème</sup> segment des phrases cibles	Verbe symétrique	Verbe Non symétrique
NP féminin en tête	1008 (70)	1151 (88)
NP masculin en tête	1112 (79)	1101 (81)

Tableau 25 : Temps de lecture du deuxième segment des phrases cibles en fonction du NP en tête de coordination et du prédicat amorce - regroupement des données de FFirst-I-et et MFirst-I-et

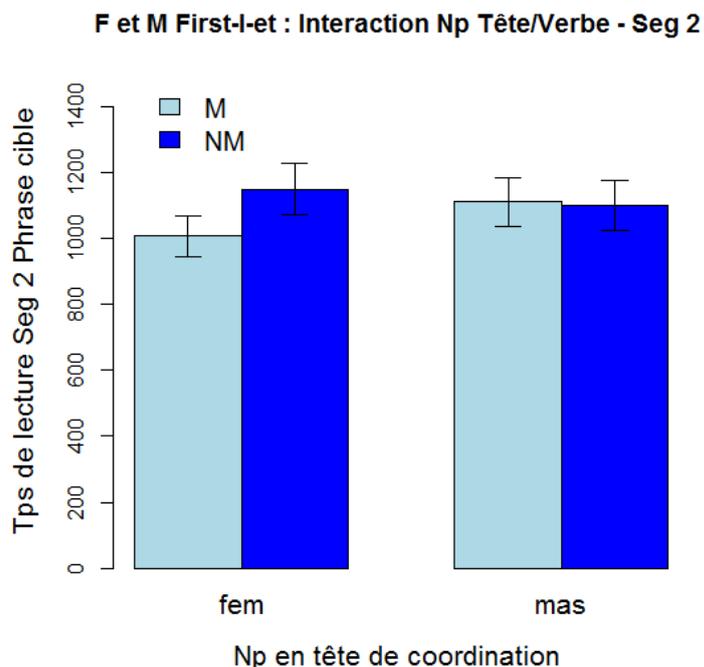


Figure 14 : Temps de lecture du deuxième segment des phrases cibles en fonction du NP en tête de coordination et du prédicat amorce - regroupement des données de FFirst-I-et et MFirst-I-et

### **C- Troisième segment de la phrase cible**

#### **Facteur « Verbe »**

On trouve une tendance à la significativité du facteur « Verbe » dans le troisième segment : les temps de lecture tendent à y être lus plus rapides (Chisquare= 3.37, df=1, p= 0.065) après des phrases amorces symétriques (1050 ms, sd= 74) qu'après les phrases amorce non symétriques (1119 ms, sd=81).

### **D- Phrase cible**

#### **Facteur « Nombre »**

On trouve un effet principal du facteur « Nombre » (Chisquare= 6.6, df=1, p= 0.010) et une interaction entre les facteurs « Verbe » et « TETE » (Chisquare= 4.31, df=1, p= 0.037).

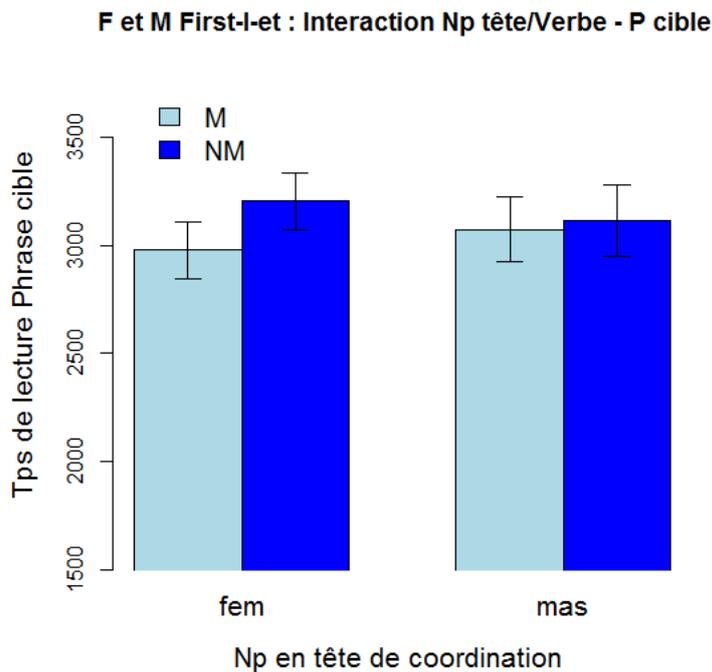
Les reprises plurielles (3018 ms, sd= 157) sont plus rapides à lire que les reprises du référent masculin (3169 ms, sd= 163).

**Interaction entre les facteurs « TETE » et « Verbe »**

L'interaction (Tableau 26 et Figure 15) entre les facteurs « TETE » et « Verbe » montre que la lecture de la phrase cible dépend non seulement de la phrase amorce mais aussi du NP en tête de coordination. Lorsque le NP en tête de la coordination est féminin, les phrases cibles sont lues plus rapidement après les phrases amorces symétriques qu'après les phrases amorces non symétriques : ce n'est pas le cas après un NP masculin en tête de coordination. Les temps de lecture des phrases cibles après les phrases amorces symétriques sont toutefois équivalents dans les deux modules ainsi que les temps de lecture après les phrases amorces non symétriques.

F-M-First-I-et : Tps de lecture phrases cibles	Verbe symétrique	Verbe non symétrique
NP féminin en tête	2977 (148)	3205 (153)
NP masculin en tête	3074 (161)	3112 (179)

Tableau 26 : Temps de lecture des phrases cibles en fonction du NP de tête de coordination et de la phrase amorce - regroupement des données de FFirst-I-et et MFirst-I-et



**Figure 15 :** Temps de lecture des phrases cibles en fonction du NP de tête de coordination et de la phrase amorce - regroupement des données de FFirst-I-et et MFirst-I-et

## **E- Temps de réponse**

### **Facteurs « Nombre » et « Verbe »**

On trouve un effet principal du facteur « Nombre » (Chisquare= 16.9, df=1,  $p < 0.001$ ) et du facteur « Verbe » (Chisquare= 4.69, df=1,  $p = 0.03$ ).

Les temps de réponse sont plus rapides quand la phrase cible est au singulier (3102 ms,  $sd=227$ ) plutôt qu'au pluriel (3465 ms,  $sd=260$ ) et plus rapides quand la phrase amorce est symétrique (3176 ms,  $sd=223$ ) plutôt que non symétrique (3390 ms,  $sd=265$ ).

## **4.5.3 Discussion**

Le test ne vérifie la proéminence de l'entité plurielle par rapport à l'un de ses constituants que dans le module où le NP masculin est en tête de coordination.

Dans le module où le NP masculin est le premier élément du pluriel coordonné, le pronom pluriel est plus vite résolu que le pronom dont l'antécédent est le référent masculin et ceci quel que soit le type de prédicat amorce. On retrouve les résultats d'Albrecht et al. (1998) qui ont mis en évidence la rapidité de traitement des reprises de l'entité coordonnée par rapport aux reprises du référent en tête de coordination. La visibilité de ce référent allouée par la symétrie verbale ne change pas le résultat. Le modèle de Gernsbacher (1989), « the Structure Building Framework », rend compte de processus qui pourraient intervenir sur l'accessibilité du premier terme de la coordination. Selon cette auteure, l'entité en première position dans une unité linguistique ancre cette unité en mémoire de travail (cf. 3.1.1) : elle constitue comme « la marque » de l'unité. Elle acquiert de ce fait une proéminence dans la coordination que le référent en deuxième position n'a pas. Mais surtout, elle porte le trait de « complexité », car elle représente une structure en plus de se représenter. Dans ce modèle, le « conjunction cost » s'explique par des traitements de désassociation de l'entité en tête de coordination d'avec la structure coordonnée. Ces processus de représentation des entités en mémoire sont indépendants des types de prédicats, comme l'avaient trouvé Gernsbacher et al. (1988) : les prédicats symétriques ne modifient pas l'avantage au premier nommé. D'où l'absence d'effet du type de prédicat dans le module où le référent masculin est en tête de coordination.

Dans le module où le NP féminin est le premier élément du pluriel coordonné, les reprises au singulier concernent le référent masculin, deuxième élément de la coordination. Les résultats ne

montrent pas de différence entre les temps de lecture des phrases cibles reprenant les deux référents et celles reprenant le référent masculin. Ce résultat suggère qu'il n'y a pas de coût d'extraction de la coordination pour ce deuxième référent. C'est ce qu'avaient trouvé Garrod et Sanford (1982) en reprenant le deuxième membre de la coordination sujet de la phrase précédente (cf. 3.1.2). Ce n'est pas ce qu'ont trouvé Gordon et al. (1999) : la reprise concernait le référent féminin en deuxième position de la coordination composée d'un référent masculin en tête dans leur test et donnait lieu à un coût d'extraction relativement à la reprise du pluriel coordonné. Les données regroupées des deux modules de notre test fournissent des indications supplémentaires sur l'absence de coût d'extraction lors de la reprise du référent en deuxième position dans la coordination : le deuxième segment mais aussi la phrase cible montrent un effet significatif du facteur « Nombre » ainsi qu'une interaction entre les facteurs « Verbe » et « TETE ». Cette dernière indique que les reprises pronominales sont plus longues à traiter dans la condition « référent en tête de coordination féminin/prédictat distributif » relativement à la condition « référent en tête de coordination féminin/prédictat symétrique » : mais elles ne sont pas plus longues à traiter que les reprises au pluriel et au singulier qui font suite aux phrases amorces qui commencent par le NP masculin. Or, ces dernières présentent un coût d'extraction. Il faut donc en déduire que l'absence de coût d'extraction du référent en deuxième position de la coordination suite aux phrases amorces qui commencent par le NP féminin s'accompagne d'un ralentissement du traitement des reprises au pluriel. Reprenant le modèle de Gernsbacher (1989), nous faisons l'hypothèse que le référent représentant la coordination étant le référent féminin quand la coordination commence par ce référent, il y a une incongruence de genre entre ce représentant et le pronom pluriel de reprise, masculin en français : cette incongruence de genre ralentirait les reprises plurielles. Il faudrait bien sûr tester cette hypothèse. Lorsque la phrase amorce est symétrique, il n'y a pas non plus de coût d'extraction mais les temps de traitement sont globalement plus rapides. Ce résultat est compatible avec les hypothèses de Patson et Ferreira (2009) sur le statut particulier des pluriels coordonnés dans les prédications symétriques : grande accessibilité de la pluralité et de ses constituants.

Les temps de réponses sont ralentis lorsque les phrases cibles sont au pluriel, c'est-à-dire reprennent le pluriel coordonné sujet de la phrase amorce dans les deux modules. Ce pluriel coordonné a la même forme de surface que le sujet de la question dans le premier module, mais pas dans le deuxième module. Ainsi, la répétition de la coordination, ne facilite pas les opérations cognitives liées à la question posée, alors même qu'elle a été relayée dans la phrase cible par le pronom pluriel. Ce résultat suggère que la répétition de la coordination est traitée comme une nouvelle entité qu'il s'agit de constituer, même et y compris si la forme de surface est respectée. Il

rappelle les mécanismes du « Repeated Name Penalty » (cf. 3.1.1) trouvés par Gordon et al. (1993). Les phrases cibles au singulier, en revanche, entraînent des temps de réponse plus courts dans les deux modules : il apparaît donc que la formation du pluriel coordonné dans la question est plus facile lorsque la coordination de la phrase amorce n'est plus aussi prégnante en mémoire et que les référents sont individuellement disponibles : ces résultats confirment ceux des tests précédents. Par ailleurs, l'effet principal du facteur « Verbe » montre que les prédicats symétriques facilitent les réponses, que la reprise soit un singulier ou un pluriel dans la phrase cible. C'est une confirmation du fait que les constituants du pluriel coordonné argument du prédicat amorce ont l'autonomie nécessaire pour se recomposer en un objet référentiel complexe dans la question. C'est surtout le signe que les spécificités des coordinations qui sont arguments externes de prédicats symétriques, dont parlent Patson et al. (2009), se conservent après les phrases cibles jusqu'aux questions. Ainsi, le prédicat d'introduction du pluriel coordonné impacte la représentation du groupe qu'il forme et celles de ses constituants : argument d'un prédicat symétrique, le pluriel coordonné attribue à ses constituants suffisamment de visibilité pour qu'il soit possible de les reprendre individuellement sans éliminer le groupe qu'ils forment des représentations en mémoire de travail.

## 4.6 Discussion générale

Un résultat général important se dégage de nos expérimentations sur les constructions comitatives : les reprises plurielles des référents sujet et régi sont aussi rapides à traiter que celles du référent sujet, proéminent, et plus rapides que celles du référent régi, non proéminent. Les expérimentations apportent des précisions sur les facteurs qui interviennent pour nuancer ce résultat général, à savoir la nature des prédicats. Les prédicats à interprétation collective, les prédicats distributifs et les prédicats symétriques ne lient pas leurs arguments pluriels de la même manière,

Un deuxième résultat important concerne la préposition *avec* qui devient une structure coordonnante dans les constructions comitatives de prédicats à interprétation collective, ce qu'elle n'est pas dans les constructions comitatives de prédicats distributifs. Ce résultat étaye la symétrie syntaxique de la construction constatée dans les études de Lakoff et Peters (1969).

Un troisième résultat important montre que la relation d'équivalence établie par les prédicats symétriques entre leurs arguments influence la formation de pluralités et facilite l'accès au référent régi, non proéminent.

Dans le pré-test, par exemple, les taux de reprises après les constructions comitatives dépendent des prédicats amorces : les prédicats symétriques ont des taux de reprises plurielles moindres que ceux des prédicats collectifs mais supérieurs à ceux des prédicats distributifs ; et des taux de reprises du référent masculin moindres que ceux des prédicats distributifs mais supérieurs à ceux des prédicats collectifs. Cette échelle de reprise s'applique aux cas où le référent entête de phrase amorce est féminin aussi bien qu'aux cas où il est masculin. Simplement les taux sont différents : plus de reprises plurielles lorsque le NP de tête est féminin et plus de reprises au singulier lorsque le NP de tête est masculin. Ce pattern des reprises vérifie la proposition de Patson et Ferreira (2009) sur la « visibilité » de la pluralité formée par les arguments et de chacun des arguments.

Dans le deuxième test en temps réel, les prédicats symétriques ne présentent pas de coût d'extraction du référent régi contrairement aux prédicats distributifs et ils ne présentent pas de coût de traitement des reprises du référent proéminent que présentent les prédicats collectifs dans le premier segment. Le premier résultat vérifie l'hypothèse d'équivalence entre les arguments sujet et régi des prédicats symétriques, posée par Koh et Clifton (2002). La rapidité d'accès au référent proéminent après des prédicats symétriques indiquée par le deuxième résultat vérifie l'autonomie de chaque argument par rapport au tout formé par les deux arguments (Patson & Ferreira, 2009).

Le troisième test en temps réel conforte les résultats de Gernsbacher et al. (1988) : l'ancrage du pluriel coordonné sur le premier référent de la coordination crée un coût d'extraction pour les reprises de ce référent relativement aux reprises du pluriel coordonné. En revanche, les reprises du pluriel coordonné ne sont pas traitées plus rapidement que les reprises du référent en deuxième position dans la coordination. Les temps de traitement globaux dans la condition NP féminin en tête de coordination, cependant, ne diffèrent pas des temps de traitement globaux de la condition NP masculin en tête de coordination : ce qui signifie que les temps de traitement des reprises du pluriel coordonné sont plus longs quand le NP de tête est féminin. Nous avons proposé une explication compatible avec le modèle de Gernsbacher (1989) qu'il faudrait tester. Le fait que les traitements globaux des reprises dans la condition NP féminin en tête soient plus rapides lorsque les prédicats amorces sont symétriques, montre la grande accessibilité à la fois de la coordination et des constituants de cette coordination avec ces prédicats, lorsque les antécédents ne sont pas réquisitionnés pour ancrer une unité linguistique dans la représentation en cours.

Les tests en temps réel suggèrent, en plus, que le pronom pluriel relaie son antécédent jusqu'à la question, ce qui montre que les regroupements opérés par les reprises pronominales plurielles correspondent à des entités représentées en mémoire de travail. Ces dernières restent actives pendant deux ou trois phrases, ce qui influence les temps de réponse aux questions : les pronoms au pluriel dans la phrase cible donnent toujours lieu à des réponses plus lentes que la question porte sur un seul référent ou sur les deux. De plus, la relation entre les arguments des prédicats symétriques semble relayée jusqu'à la question, comme le montre la rapidité de réponse des participants lorsque la phrase amorce est symétrique (dans les données regroupées des deuxième et troisième tests). Fait remarquable, cette relation est relayée, que la reprise de la phrase cible ait concerné les deux référents ou un seul. Apparemment, la structure de l'argument pluriel d'un prédicat symétrique lui permet d'être repris totalement ou partiellement sans qu'elle en soit modifiée : ce résultat va dans le sens de la proposition faite en 2.2.1.2.4, concernant la nature duelle des arguments des prédicats symétriques.

Par ailleurs, l'importance du référent placé en tête de prédication se vérifie dans les trois tests en temps réel : les facteurs n'ont pas les mêmes effets selon que le genre du pronom de reprise est le même ou pas que le genre du référent de tête, y compris quand l'argument externe est une coordination. Le premier test montre que lorsque le référent masculin est en tête et qu'il est repris pronominalement, les temps de lecture de la phrase cible sont plus courts que lorsqu'il n'occupe pas cette première position : de fait, les temps de lecture deviennent équivalents à ceux des phrases cibles qui reprennent les deux référents. L'imbrication de ce facteur avec les facteurs « type de prédicat amorce » et « nombre de la reprise » est cependant, complexe. Dans le deuxième test, par exemple, cet avantage pour le premier nommé disparaît pour les prédicats distributifs, dont le pré-test avait pourtant montré qu'ils présentaient un fort taux de reprise du référent de tête. Il semble que le déplacement de l'attention des participants vers les deux référents par le biais de la question en soit la cause. L'avantage au premier nommé s'avère être un désavantage lorsque des traitements sémantiques concernant les deux référents sont à l'œuvre dans le deuxième segment. Les prédicats symétriques ne présentent pas ce désavantage. Dans l'analyse portant sur les données regroupées des deux modules, les phrases cibles du deuxième test sont lues plus rapidement dans les cas de symétrie verbale. Les résultats de Gernsbacher et al. (1988) sur l'avantage au premier nommé sont de toute première importance pour comprendre les différences de traitements auxquels ont donné lieu les deux modules de chacun des tests que nous avons conduits. Ils expliquent comment les processus discursifs sont influencés par l'ordre d'occurrence au sein des unités linguistiques. Il faudrait d'autres recherches, cependant, pour comprendre comment interagissent l'ordre

d'occurrence des référents, les propriétés des prédicats et les tâches demandées en fin de test qui orientent l'attention des lecteurs. Il faudrait aussi d'autres recherches pour vérifier l'hypothèse évoquée pour expliquer l'allongement des traitements portant sur les reprises de la coordination lorsque le référent de tête est féminin, à savoir une incongruence entre le genre de ce référent et le genre du pronom de reprise pluriel.

Les tests en temps réel sont parcourus de processus parallèles propres aux tâches de compréhension, comme la profondeur des traitements mais confirment l'essentiel de nos hypothèses. On peut considérer, par exemple, que dans le premier test en temps réel, les résolutions pronominales sont basées sur des informations de surface qui permettent de résoudre rapidement les reprises du référent en focus, sans nécessairement élaborer la relation qu'il entretient avec le référent régi. La tâche des participants semble déterminer la profondeur des traitements sémantiques. Dans le deuxième test, la question portant sur les deux référents, un traitement sémantique semble intervenir dans la résolution pronominale : il ralentit la recherche de l'antécédent en focus en cas de prédicat amorce distributif ; en cas de prédicats amorces collectifs ou symétriques, le traitement sémantique est minimal, les deux référents entretenant une relation définie par le noyau verbal. D'autres processus, liés aux traitements parallèles des formes de surface et des représentations mentales interviennent : la répétition de la forme de surface des pluriels coordonnés dans les questions, par exemple, n'est pas associée immédiatement à la représentation mentale déjà constituée, mais passe par une nouvelle interprétation. Ces premiers résultats devront être étayés et précisés. On a évoqué l'importance de l'ancrage en mémoire de la coordination dans le troisième test en temps réel. Mais ce phénomène semble également s'appliquer au référent sujet des constructions comitatives : la littérature psycholinguistique n'évoque la proéminence du sujet grammatical en tête de prédication que comme une conséquence de la fonction occupée par ce référent et de sa position en tête de phrase. Il semble cependant jouer un rôle particulier dans la représentation de la prédication en cours comme le suggèrent les résultats obtenus dans les modules où le référent repris occupait la fonction de sujet et conformément au modèle de Gernsbacher (1989).