

**Serge Tassé, Joachim Reinwein et Denis Foucambert**

LE COMPORTEMENT OCULAIRE DE LECTEURS ADULTES LORS DU TRAITEMENT  
DE CORÉFÉRENCES DE TYPE ANAPHORIQUE ET CATAPHORIQUE DANS UN TEXTE  
ILLUSTRÉ : ÉTUDE EXPLORATOIRE

**Résumé.** Cette étude exploratoire porte sur la comparaison de constructions avec pronoms anaphoriques et cataphoriques dans des textes illustrés et leur impact sur les processus cognitifs du lecteur expérimenté. Elle utilise la technologie FaceLab v4.6 (2008), jumelée au logiciel GazeTracker v8 (2008), pour observer et enregistrer les fixations oculaires de 19 lecteurs adultes. Le matériel expérimental, qui consiste en douze textes narratifs comportant cinq phrases chacun, fait varier l'ordre du pronom et de son coréférent [anaphore (*Paul ... il*), cataphore (*Il ... Paul*)] et le genre grammatical (identique, différent) d'un nom-leurre interposé entre les deux éléments coréférentiels. Plusieurs zones globales (2) et sous-zones (5) sont délimitées dans le texte et dans l'illustration. Les fixations localisées dans les zones globales du texte et de l'illustration indiquent une interaction des deux variables. Les lecteurs accordent significativement plus de temps à l'illustration en présence d'une construction cataphorique avec leurre de genre grammatical différent, comparativement aux trois autres conditions expérimentales. Les résultats de l'étude semblent soutenir l'idée que le recours à l'image s'explique, au moins en partie, sur la base de variables textuelles dûment identifiées.

**Mots-Clefs :** compréhension, image, lecture, pronom, oculométrie

## **Introduction**

L'image est souvent utilisée pour accompagner des informations textuelles. Son impact est généralement positif en compréhension en lecture (Reinwein, 1998-...). On l'évalue essentiellement à l'aide de mesures en différé, comme des questionnaires. Or, la manière dont l'image est précisément explorée relativement aux informations textuelles écrites et ce, au moment même où le lecteur traite ces informations, demeure pour l'instant inconnue. Quelques études montrent, cependant, de façon globale, que le temps consacré au texte est plus considérable que celui consacré à l'image et qu'il n'y a pas de réels va-et-vient entre le texte et l'image (Hegarty, 1992; Rayner, Rotello, Stewart, Keir & Duffy, 2001, entre autres).

Par ailleurs, dans la perspective de l'incrémentation des informations, l'exploration visuelle de l'image lors de l'écoute de phrases a permis de valider certains mécanismes anticipatoires liés à la structure argumentative ou thématique de la phrase (Kamide, Altmann & Haywood, 2003; Knoeferle & Crocker, 2006; Knoeferle, Crocker, Scheepers & Pickering, 2005).

Pour comprendre les processus du lecteur de textes illustrés, il nous apparaît d'abord important de cerner avec le plus de précision possible les variables linguistiques qui peuvent moduler la prise des informations picturales et leur intégration. Il se pourrait que l'effet de l'image sur la compréhension de texte ne dépende pas uniquement de sa fonction par rapport au texte dans son

**Revue pour la recherche en éducation**, actes de colloque, Acfas 2010, p. 1-18  
@2011, Ottawa, Canada

ensemble - piste qui a donné lieu à différentes taxonomies pictoverbales (p. ex. : Carney & Levin, 2002; Goldsmith, 1984; Hupka, 1989; Levin, Anglin & Carney, 1987) -, mais également de la construction linguistique spécifique à laquelle se rapporte l'image.

On sait, de façon générale, que le parcours visuel qu'un lecteur effectue sur un texte illustré se caractérise par la priorité accordée au texte relativement à l'image, en termes de nombre et de durée des fixations oculaires (Hannus & Hyönä, 1999; Hegarty, 1992; Rayner et al., 2001; Underwood, Jebbett, & Roberts, 2004). Aussi, selon Underwood (2005, p. 163), le lecteur adopterait un comportement visuel prototypique : « Eye-fixation studies indicate that viewers characteristically look briefly at the picture, then read the text, and then, if necessary, they look more carefully at the picture ».

Au-delà des effets contextuels globaux, le recours à l'image et à certaines de ces constituantes par le lecteur pourrait être modulé par des facteurs linguistiques. Dans le but d'examiner le rôle éventuel de l'image, nous avons choisi de la confronter au phénomène linguistique de la coréférence. Le traitement du pronom, dans son rapport anaphorique (*Paul ... il*) comparé à celui dans son rapport cataphorique (*Il ... Paul*) représente un sujet intéressant pour évaluer si l'image peut contribuer à la cohérence textuelle. Plus précisément, le fait de présenter une image sur laquelle figure le référent du pronom permet-elle de résoudre (aussi) facilement le pronom en situation cataphorique – étant donné que le référent suit le pronom - qu'en situation anaphorique? Examinons plus étroitement quelques aspects de la coréférence.

### **Le pronom comme élément de reprise**

L'anaphore est un moyen linguistique essentiel pour assurer la continuité discursive et, par ce fait même, la cohérence textuelle. Une expression anaphorique, comme le pronom, se substitue à une expression linguistique qui la précède ou, occasionnellement, qui la suit (*cataphore*) et contribue à établir des liens entre les différentes informations séquentiellement présentées. À l'intérieur de la phrase, l'emploi des anaphores est généralement considéré comme étant régi par des règles linguistiques, même si les analyses psycholinguistiques ne confirment pas toujours les analyses basées sur l'intuition du locuteur natif (Gordon & Hendrick, 1997; Kaiser, Runner, Sussman & Tanenhaus, 2009). Du point de vue psycholinguistique, le traitement de l'anaphore par le lecteur est un processus à charge cognitive fort variable, lequel est modulé par des facteurs tels que la proximité des éléments coréférentiels, la fonction grammaticale des éléments coréférentiels, la présence de plusieurs éléments potentiellement coréférentiels dans le contexte ou encore l'ordre séquentiel des éléments coréférentiels. Sur le plan discursif, il semble que la capacité du lecteur à reconstruire des liens coréférentiels résulte de l'interaction de plusieurs facteurs et ne peut facilement s'expliquer par l'exécution d'un simple algorithme comparable à celui opérant à l'intérieur de la phrase.

La compréhension d'expressions anaphoriques, et du pronom défini en particulier, a fait l'objet d'études en direct au moyen de la technique d'autoprésentation par segment, de l'observation des mouvements oculaires ou de tâches de décision lexicale avec amorce (Garnham, 2001). Le traitement cognitif du pronom anaphorique en lecture silencieuse est d'intérêt central pour

**Revue pour la recherche en éducation**, actes de colloque, Acfas 2010, p. 1-18  
@2011, Ottawa, Canada

*l'immediacy assumption* de Just et Carpenter (1980), selon laquelle un lecteur essaierait d'interpréter chaque mot au moment où il le rencontre, ou pour leur *eye-mind assumption* selon laquelle un lecteur fixerait un mot aussi longtemps que nécessaire pour le traiter. Depuis, ces postulats ont été remis en question (p. ex. : Ehrlich & Rayner, 1983) ou se sont avérés incompatibles avec certains faits empiriques rapportés (van Gompel & Liversedge, 2003). Par ailleurs, il n'est pas facile d'expliquer, selon cette conception, le traitement supposément instantané d'un pronom cataphorique puisque le coréférent lexical du pronom *suit* ce dernier (voir aussi Filik & Sanford, 2008; Kazanina, Lau, Lieberman, Yoshida & Phillips, 2007). Si l'on admet l'hypothèse générale d'un traitement incrémentiel sur une base mot-à-mot (Kamide et al., 2003), le pronom, dans une cataphore, pourrait ne pas être traité suivant le même mécanisme que dans une anaphore.

Pour Gordon et Hendrick (1997), l'identification instantanée d'un pronom cataphorique pourrait poser problème et être suspendue délibérément en attendant la lecture du nom coréférentiel suivant. Filik et Sanford (2008) ont toutefois montré que le sujet ne suspend pas ses processus, mais que le processeur continue à vouloir incrémenter l'information, comme ils l'ont montré avec le pronom singulier, engendrant rapidement une difficulté lorsque la résolution du pronom essuie un échec. Par ailleurs, selon Kennison, Fernandez & Bowers (2009), l'identification, en temps réel, des deux éléments linguistiques coréférentiels dans des constructions anaphoriques et cataphoriques semble reposer sur un mécanisme anticipatoire fondé sur la nature discursive des coréférents, la cataphore étant plus difficile à traiter que l'anaphore.

### **L'image comme support référentiel**

À une exception près (Glenburg & Kruley, 1992), on s'est servi jusqu'ici de textes non illustrés pour étudier le traitement d'expressions anaphoriques par le lecteur. Glenburg et Kruley se sont intéressés au rôle potentiellement facilitateur de l'image lors de la résolution d'anaphores nominales. D'après eux, l'image devrait diminuer la charge cognitive de la mémoire de travail, en particulier, lorsque les deux syntagmes nominaux coréférentiels sont éloignés l'un de l'autre dans le texte. Dans une telle situation, l'image pourrait servir de «mémoire externe» rendant l'antécédent directement accessible et évitant ainsi une recherche cognitivement coûteuse dans la mémoire à long terme. Comme il avait déjà été démontré que le processus de résolution d'une anaphore était sensible à la distance entre les éléments coréférentiels (p. ex. Clark & Sengul, 1979; Ehrlich & Rayner, 1983; Kintsch, 1988), Glenburg et Kruley s'attendaient à un effet d'interaction Image X Distance, selon lequel la présence de l'illustration profiterait particulièrement au traitement des coréférents éloignés, comparativement aux autres conditions. Le résultat n'a pas confirmé cette hypothèse puisque l'interaction réellement obtenue pointait dans la direction opposée : pour ce qui était du temps de lecture, le traitement des coréférents éloignés en version illustrée prenait significativement plus de temps que les trois autres versions. À cet égard, il nous apparaît que la méthodologie employée est questionnable. Cependant, nous n'élaborerons pas cet aspect ici.

Malgré l'échec apparent de Glenburg et Kruley, leur idée de départ s'inscrit bien dans les modèles dits *mental* (Johnson-Laird, 1980) ou *situationnel* (Kintsch, 1988). On présume que la **Revue pour la recherche en éducation**, actes de colloque, Acfas 2010, p. 1-18 @2011, Ottawa, Canada

construction d'un *modèle* lors de la compréhension de textes facilite la représentation de la situation décrite dans un texte et partant, des liens coréférentiels sous-jacents. Même si le concept de *modèle* a été développé sans référence au support pictural, dans les recherches récentes sur la compréhension de textes illustrés, celui-ci est devenu un concept central pour expliquer l'effet (souvent) facilitateur de l'image sur la compréhension du texte qu'elle accompagne (Gyselinck, 1996; Gyselinck & Tardieu, 1999). Dans un *modèle*, l'image pourrait jouer le rôle d'un élément facilitateur lors du traitement de constructions linguistiques complexes (absence de marqueurs morphosyntaxiques, éloignement entre éléments coréférentiels, etc.). Précédemment, Bock & Milz (1977) avaient poursuivi une piste semblable, en examinant le rôle de désambiguïsation de l'image sur le rappel de phrases avec un pronom référentiellement ambigu. D'après eux, l'effet facilitateur de l'image était un argument en faveur d'un traitement unifié du texte et de l'image.

Dans la présente étude, nous nous intéressons à la comparaison de constructions anaphoriques et cataphoriques dans des textes illustrés et leur impact sur les processus cognitifs du lecteur. Nous avons étendu l'usage du terme « cataphore », limité habituellement au domaine de la phrase. La coréférence anaphorique a attiré davantage l'attention des chercheurs que la coréférence cataphorique, la première étant un phénomène linguistique plus répandu (Filik & Sanford, 2008). Pour cela, il est préférable d'avoir recours à une méthodologie d'observation en temps réel permettant de suivre le parcours visuel du lecteur : quand et à quelle fin, le lecteur a-t-il recours à l'image? Rayner et ses différents collaborateurs ont abondamment utilisé la technique de l'enregistrement des mouvements oculaires pour explorer différents aspects en lecture [cf. Rayner (1998, 2009)]. Avant d'explorer cette situation dans une perspective développementale, nous avons priorisé l'observation de lecteurs expérimentés. Roy-Charland, Saint-Aubin & Evans (2007) ont montré que l'attention accordée au texte dans un environnement picto-verbal est fonction du niveau d'habileté des enfants. Nous utiliserons un matériel (texte-dessin) simple, mais notre intérêt est avant tout théorique. Nous cherchons à cerner l'impact de l'image relativement à un phénomène linguistique particulier. Autrement dit, l'image, comme contexte pictural, peut-elle fournir des informations qui faciliteraient le traitement de la coréférence?

Un des facteurs qui influe sur la facilité à résoudre le pronom est la présence de marques morphologiques distinctives (Arnold, Eisenband, Brown-Schmidt & Trueswell, 2000; Vonk, 1984) C'est le cas dans l'exemple (1) : le pronom féminin « elle » ne peut se référer à « Paul ». Dans l'exemple (2) par contre, la coréférence du pronom n'est pas aussi claire (Vanessa ou Mélodie).

(1) **Vanessa** entra doucement dans la classe. **Paul** fit taire ses élèves. **Elle** salua les élèves et fit quelques pas de danse.

(2) **Vanessa** entra doucement dans la classe. **Mélodie** fit taire ses élèves. **Elle** salua les élèves et fit quelques pas de danse.

Quel rôle l'image peut-elle jouer dans la compréhension de liens coréférentiels ambigus? Supposons que l'image ci-dessous (figure 1) soit jumelée avec (2). En observant l'image, un lecteur n'a aucune difficulté à identifier « Mélodie » comme étant l'enseignante (c'est-à-dire le

**Revue pour la recherche en éducation**, actes de colloque, Acfas 2010, p. 1-18  
@2011, Ottawa, Canada

personnage à gauche) et « Vanessa » comme celle qui « salua les élèves et fit quelques pas de danse ». Pour les besoins de la discussion, remplaçons dans l'image l'enseignante par un enseignant et jumelons-la avec (1). Dans le cas du segment textuel (1), la coréférence n'est pas ambiguë (i.e. « Vanessa – elle ») de sorte que l'image n'a plus de fonction de désambiguïsation.

FIGURE 1 : Exemple d'image



Comparons (1) et (2) avec leur construction cataphorique analogue, c'est-à-dire (3) et (4) :

(3) **Elle** entra doucement dans la classe. **Paul** fit taire ses élèves. **Vanessa** salua les élèves et fit quelques pas de danse.

(4) **Elle** entra doucement dans la classe. **Mélo die** fit taire ses élèves. **Vanessa** salua les élèves et fit quelques pas de danse.

Les segments textuels (3) et (4) peuvent paraître moins naturels. Ceci dit, on peut présumer que grâce à la présence de l'image, le segment (4) sera interprété de la même manière que le segment (2). Et en remplaçant de nouveau dans l'image l'enseignante par un enseignant, la même chose vaudra pour les segments textuels (3) et (1).

Ce bref regard sur l'image que porte un lecteur au début de sa lecture, comme le soutient Underwood (2005), permet-il une prise d'information suffisante, laquelle garantirait la résolution du pronom quelle que soit sa position et, auquel cas, on ne devrait pas obtenir de différence (en termes de temps de lecture et de nombre de fixations) entre les conditions anaphorique et cataphorique sur l'image ou dans le texte? Ou bien, une construction cataphorique étant plus difficile à traiter entraînera-t-elle un recours à l'image plus considérable? De plus, une construction référentiellement ambiguë entraînera-t-elle une attention accrue à l'image afin d'identifier, dans le cas qui nous concerne, le bon référent? Enfin, y aura-t-il interaction entre les deux phénomènes linguistiques (type de coréférence et ambiguïté ou non ambiguïté morphologique)? C'est donc à ces questions que tentera de répondre notre étude.

On doit rappeler que celle-ci est exploratoire. Elle représente le début d'une démarche de recherche; elle comporte assurément certaines limites. Cependant, nous croyons bon de diffuser à la communauté, dans le cadre d'un colloque, l'état de son avancement. Son objectif est double. Premièrement, au risque de nous répéter, il s'agit de mieux comprendre comment un lecteur utilise et articule dans le décours de ses processus de traitement des informations linguistiques et picturales. Deuxièmement, comme l'observation en direct nécessite une technologie appropriée, nous avons eu l'occasion d'utiliser un nouvel outil pour l'enregistrement des mouvements oculaires. Dans une perspective méthodologique, il importe de déterminer les paramètres adéquats, d'une part, pour créer des environnements contrôlés et, d'autre part, pour opérationnaliser la cueillette des observations et leur extraction pour les fins de traitement statistique.

## **Méthodologie**

### ***Participants***

Vingt sujets adultes provenant de la communauté de l'Université du Québec à Montréal ont participé à l'expérimentation. Sur un total de quarante volontaires, nous avons retenu ceux qui avaient une vision normale, donc sans lunettes ni lentilles cornéennes, et qui avaient rencontré les exigences de précision lors de la phase de calibration de l'outil.

### ***Instruments***

Les mouvements oculaires des sujets ont été enregistrés à l'aide de la technologie FaceLab v4.6 (2008). La captation des informations visuelles s'effectue avec deux caméras à l'infrarouge positionnées devant le sujet. Afin d'établir la séquence des fixations, la prise d'information est effectuée à toutes les 16 ms (échantillonnage à 60 Hz). Une phase préliminaire à l'enregistrement proprement dit est nécessaire. Elle consiste à modéliser la tête du sujet et à calibrer son regard. Cet appareillage offre une position de lecture plus naturelle, sans les contraintes d'un casque ou d'un appui-menton. À cette technologie est combiné le logiciel GazeTracker v8 (2008), lequel permet la présentation des stimuli, la synchronisation du regard avec les éléments affichés et la saisie des données oculaires. Pour l'expérimentation, nous avons utilisé un écran plat (de dimension 376,3 mm sur 301,1 mm), de résolution 1280 sur 1024 pixels. Les sujets en étaient éloignés d'environ 80 cm.

### ***Matériel expérimental***

Nous avons d'abord conçu, suivant un même modèle, douze textes narratifs comportant cinq phrases chacun. La première phrase a pour fonction d'introduire la situation de façon très générale. La seconde met en scène un premier personnage. La troisième phrase présente un second personnage. La quatrième phrase donne une information contextuelle de la situation présentée et la dernière décrit une situation propre au premier personnage. Deux versions ont été constituées à partir de ces textes (cf. le tableau 1), et correspondant aux deux niveaux de la

variable Type de coréférence (anaphore et cataphore). Le premier personnage est soit présenté à l'aide d'un prénom usuel (version anaphore), soit par le pronom correspondant (version cataphore). Par conséquent, à la quatrième phrase, le pronom correspondant au premier personnage introduit la phrase dans la version avec anaphore, alors que dans le cas de la cataphore, il s'agit du prénom correspondant. Le nombre de mots écrans entre les coréférents est de dix-sept mots en moyenne. Six textes ont des personnages qui partagent le même genre, et six autres qui ne le partagent pas.

Pour chaque texte, nous avons créé une illustration. Les illustrations ont été composées à l'aide de différentes banques d'images libres de droit et celles-ci ont été intégrées et modifiées avec le logiciel Paint de Microsoft. Elles ont pour fonction de *représenter* l'information textuelle (Levin et al., 1987) et, le cas échéant, de la désambiguïser. Pour la moitié des situations, le personnage sur lequel se fonde la relation coréférentielle a été placé à gauche du second personnage et pour l'autre moitié, à droite; ceci afin de neutraliser toute expectative due à une présentation uniforme. La même illustration est utilisée pour les deux versions. Nous avons donné un exemple à la figure 1.

TABLEAU 1 : Versions anaphorique et cataphorique de deux textes illustrés

<p>(1a) <u>Anaphore / même genre</u> : C'était la semaine de la culture à l'école. <b>Vanessa*</b> entra doucement dans la classe. <b>Mélodie</b> fit taire ses élèves. Le silence emplît aussitôt la classe. <b>Elle</b> salua les élèves et fit quelques pas de danse.</p> <p>(1b) <u>Cataphore / même genre</u> : C'était la semaine de la culture à l'école. <b>Elle</b> entra doucement dans la classe. <b>Mélodie</b> fit taire ses élèves. Le silence emplît aussitôt la classe. <b>Vanessa</b> salua les élèves et fit quelques pas de danse.</p> <p>(2a) <u>Anaphore / genre différent</u> : On prévoit beaucoup de pluie pour les prochains jours. <b>Paul</b> doit rapidement terminer le travail. <b>Luce</b> craint une infiltration au sous-sol. De gros nuages gris s'amoncellent déjà. <b>Il</b> colmate la fissure avec du béton.</p> <p>(2b) <u>Cataphore / genre différent</u> : On prévoit beaucoup de pluie pour les prochains jours. <b>Il</b> doit rapidement terminer le travail. <b>Luce</b> craint une infiltration au sous-sol. De gros nuages gris s'amoncellent déjà. <b>Paul</b> colmate la fissure avec du béton.</p>
--

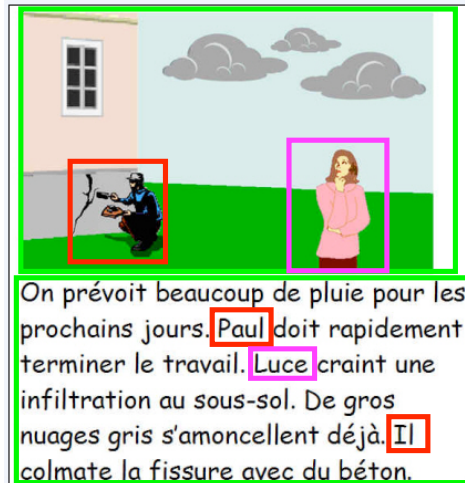
\*Indice typographique **gras** absent dans le matériel expérimental

Deux ensembles expérimentaux (A et B) ont été constitués de manière à ce que chaque sujet rencontre trois fois chaque combinaison (i.e. 4) découlant du croisement des deux variables à l'étude (Type de coréférence : anaphore/cataphore et Genre : même genre/genre différent, cf. Tableau 1). Dix lecteurs étaient assignés à l'ensemble A, et 10 autres à l'ensemble B.

Pour la présentation à l'aide de GazeTracker, nous avons préparé douze écrans; chacun comportant un texte et l'illustration correspondante. L'image est toujours située en haut du texte et occupe environ la moitié de la page. La police de caractères choisie est *Comic Sans MS* de taille 24 points. Nous avons défini plusieurs zones d'intérêt sur chaque écran (figure 2): trois sur l'illustration et quatre sur le texte. Une première zone encadre globalement l'image, deux autres encadrent spécifiquement les deux personnages. Pour le texte, on retrouve une première zone qui encadre tout le texte et trois autres qui entourent séparément les deux prénoms et le pronom. Ces cadres demeurent invisibles pour le lecteur. Les données oculaires de ces zones seront spécifiquement examinées.

**Revue pour la recherche en éducation**, actes de colloque, Acfas 2010, p. 1-18  
@2011, Ottawa, Canada

FIGURE 2 : Les zones



Nous avons, de plus, formulé une « question de compréhension » de type vrai / faux et celle-ci suivait l'écran du texte illustré. La question portait toujours sur la deuxième phrase. La réponse exigeait du sujet qu'il ait consulté l'image, puisque le personnage était identifié relativement à son occupation. Par exemple, pour le premier texte du tableau 1, l'affirmation était la suivante : *L'enseignante entra doucement dans la classe*. Le sujet devait alors fournir oralement la réponse. L'ajout d'une question de type Vrai-Faux avait pour objectif d'engager le lecteur dans une activité de compréhension. Il y avait autant d'affirmations dont la réponse était vraie que d'affirmations dont la réponse était fausse.

### ***Déroulement***

Pour chaque participant, la rencontre débutait avec une phase de préparation (modélisation de la tête et calibration du regard) qui durait environ quatre minutes. Après quoi, il recevait la consigne qu'il aurait à lire de courts textes accompagnés chacun d'une image. Il aurait, en plus, à répondre oralement à une question de type vrai ou faux. Un exercice avec trois textes similaires à ceux de l'expérimentation lui était d'abord proposé. C'est le sujet qui faisait avancer successivement les écrans (écran-texte illustré et écran-question) à l'aide du bouton gauche de la souris. Deux expérimentateurs étaient sur place dont un consignait par écrit les réponses aux questions. L'expérimentation proprement dite durait en moyenne dix minutes.

### **Résultats**

À cause d'une codification erronée, l'un des douze écrans a dû être exclu de l'analyse, de sorte qu'une des quatre combinaisons découlant du croisement des variables Type de coréférence et Genre reposait sur l'analyse de deux textes illustrés (plutôt que de 3).

En conformité avec bon nombre d'études sur le mouvement des yeux, nous avons défini la fixation oculaire comme étant d'une durée minimale de 80 ms (Rayner, 1998; Rayner & Pollatsek, 1987). Le temps total de lecture n'est cependant pas soumis à la contrainte d'une durée minimale. Bien que les deux mesures soient de bons indicateurs des processus de haut niveau, comme ceux de l'intégration sémantique et de la résolution des ambiguïtés (Libben & Titone, 2009), les données concernant le parcours oculaire, comme les saccades de régressions, la durée de la première fixation, etc. (Rayner, 1998), - qui sont des mesures importantes dans ce genre de recherche - n'ont pu être extraites, pour le moment, de la masse des informations enregistrées, faute d'un algorithme pour les récupérer. La récupération manuelle des données s'avère trop ardue.

Différentes zones de traitement ont été définies, que ce soit pour le texte ou pour l'image. La perte de données, définie comme le pourcentage de fixations situées à l'extérieur de ces zones, était de l'ordre de 8.1% et de 8.3% pour les ensembles A et B respectivement. Le pourcentage de réussite aux questions est relativement élevé (85%). Une ANOVA 2x2 prenant le Type de coréférence et le Genre comme mesures intra-items n'apporte des résultats significatifs que pour la variable Genre  $F(1,18)=7,01$ ,  $p<0,02$  avec 79% de réussite pour un genre identique et 91% pour un genre différent.

### *Les temps de lecture sur les différentes zones du texte et de l'image*

Pour les fins de cet article, nous ne présenterons pas les résultats concernant le nombre de fixations. Cette mesure étant étroitement en lien avec celle du temps de fixation, les résultats obtenus sont effectivement très similaires.

#### *1- Les zones du texte*

Les analyses statistiques correspondent au plan décrit plus haut: les variables dépendantes sont successivement les temps totaux de lecture sur les 4 zones du texte décrites précédemment et elles sont analysées à travers quatre ANOVA (2 x 2) successives prenant le Type de coréférence (anaphore / cataphore) et le Genre (même genre / genre différent) comme facteurs de mesures répétées. Le tableau 2 (partie supérieure) présente les temps totaux passés dans les quatre zones du texte.

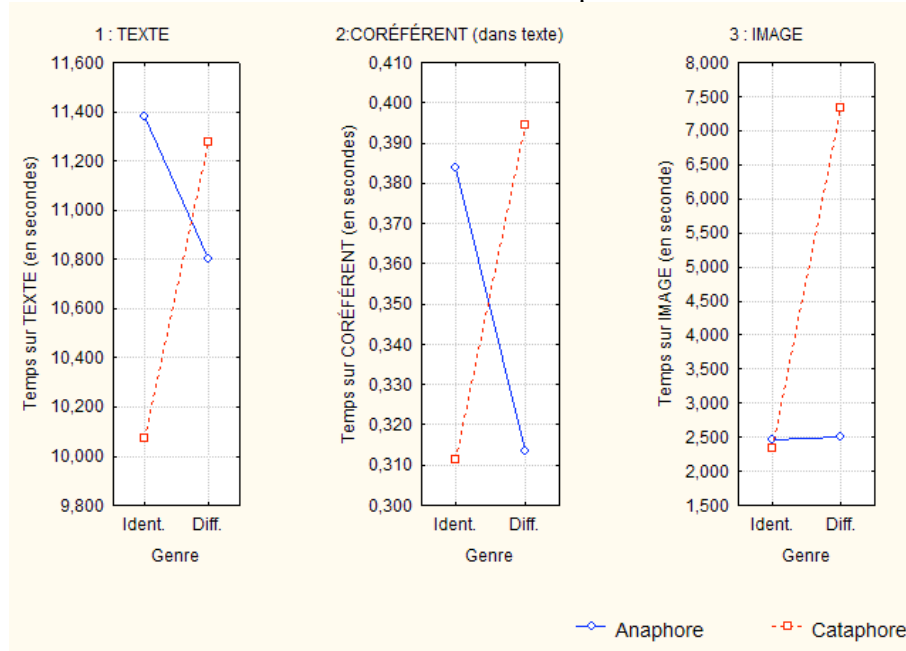
Les résultats sur la zone globale TEXTE n'indiquent ni d'effet du Type de coréférence,  $F(1,18) = 1.78$ ,  $p<0.2$ , ni du Genre,  $F(1,18) = 1.95$ ,  $p<0.18$ . En revanche, on observe une interaction significative de ces deux facteurs,  $F(1,18) = 11.97$ ,  $p<0.003$ . Les tests post-hoc (HSD de Tukey) précisent deux éléments distincts (figure 3; panneau 1) : (a) dans le cas où le genre est identique, les textes anaphoriques sont lus plus lentement que les textes cataphoriques ( $p<0.01$ ) et (b) dans le cas où le genre est différent, les textes cataphoriques sont lus cette fois plus lentement que les textes anaphoriques ( $p<0.02$ ).

TABLEAU 2 : Moyennes et Écarts-types des temps de lecture dans les différentes zones observées en fonction du Type de coréférence et du Genre

	Type : Anaphore				Type : Cataphore			
	Genre Identique		Genre Différent		Genre Identique		Genre Différent	
	M	ET	M	ET	M	ET	M	ET
Zone TEXTE								
TEXTE	11,39	3,90	10,80	3,40	10,07	2,89	11,28	3,27
CORÉFÉRENT	0,38	0,26	0,31	0,23	0,31	0,22	0,39	0,29
PRONOM	0,14	0,16	0,07	0,06	0,09	0,10	0,08	0,11
LEURRE	0,44	0,31	0,44	0,29	0,49	0,19	0,37	0,19
Zone IMAGE								
IMAGE	2.45	0.97	2,49	1,66	2.35	1.04	7.34	4.57
CORÉFÉRENT	0.84	0,48	0.73	0.48	0.67	0.37	0.75	0,08
LEURRE	0.54	0.29	0.49	0.31	0.56	0.32	0.51	0.32

Les résultats sur la zone CORÉFÉRENT n'indiquent ni un effet du Type de coréférence,  $F(1,18) < 1$ , ni du Genre,  $F(1,18) < 1$ . L'interaction Type x Genre est, quant à elle, significative,  $F(1,18) = 8.71$ ,  $p < 0.009$ . Les tests post-hoc (HSD de Tukey) indiquent que les temps de lecture sur la zone CORÉFÉRENT sont significativement plus longs (figure 3; panneau 2) dans les situations de cataphore avec genre différent que dans les situations anaphoriques de genre différent ( $p < 0.05$ ) ou dans les situations cataphoriques de genre identique ( $p < 0.04$ ). Les résultats sur la zone PRONOM ne montrent, quant à eux, aucune différence significative, ni pour le Type de coréférence,  $F(1,18) < 1$ , ni pour le Genre,  $F(1,18) = 2.96$ ,  $p < 0.11$ , ni pour leur interaction,  $F(1,18) = 1.25$ ,  $p < 0.3$ . Enfin, les résultats sur la zone LEURRE ne montrent, eux non plus, aucune différence significative, ni pour le Type,  $F(1,18) < 1$ , ni pour le Genre,  $F(1,18) = 4.33$ ,  $p < 0.06$ , ni pour leur interaction,  $F(1,18) = 2.34$ ,  $p < 0.15$ .

FIGURE 3 : Résultats des temps de lecture



### *Les zones de l'image*

Dans ces analyses, les variables dépendantes sont les temps de lecture sur les 3 zones de l'image décrites supra et elles sont analysées à travers trois ANOVA (2 x 2) successives prenant le Type de coréférence (anaphore / cataphore) et le Genre (même genre / genre différent) comme facteurs de mesures répétées. Le tableau 2 (partie inférieure) présente les temps totaux passés dans les trois zones observées.

Les résultats sur la zone globale IMAGE indiquent un effet du Type de coréférence,  $F(1,18) = 21.86$ ,  $p < 0.001$ , du Genre,  $F(1,18) = 18.09$ ,  $p < 0.001$ , et de l'interaction entre ces deux facteurs,  $F(1,18) = 23.44$ ,  $p < 0.001$ . Les tests post-hoc (HSD de Tukey) indiquent (figure 3; panneau 3) que les temps de lecture de l'image sont plus importants avec un texte cataphorique de genre différent que dans tous les autres cas (tous  $p < 0.001$ ). Les résultats sur la zone CORÉFÉRENT n'indiquent ni un effet du Type de coréférence,  $F(1,18) < 1$ , ni du Genre,  $F(1,18) < 1$ , ni d'interaction Type x Genre,  $F(1,18) = 1,84$ ,  $p < 0.2$ . Enfin, on n'observe pas de différence significative quant aux temps de lecture de la zone de l'image illustrant le leurre, que ce soit en fonction du Type de coréférence, du Genre ou de leur interaction, tous les  $F(1, 18) < 1$ .

### **Discussion des résultats**

**Revue pour la recherche en éducation**, actes de colloque, Acfas 2010, p. 1-18  
@2011, Ottawa, Canada

La résolution du pronom exige l'établissement d'un lien coréférentiel avec une autre entité présente dans le texte. En lui-même, le pronom ne possède pas de contenu lexical et doit être saturé. La présence de marques morphosyntaxiques de genre et de nombre ainsi que sa fonction dans la phrase peuvent être utilisées pour établir cette saturation. Dans la situation anaphorique, le coréférent précède le pronom, tandis que pour la cataphore, il le suit. Dans les situations que nous avons proposées aux sujets, nous avons introduit un second prénom, lequel faisait écran entre le coréférent et le pronom. Dans la moitié des cas, le pronom avait le même genre que le leurre.

Comme on a pu le voir au panneau 1 de la figure 3, le temps de lecture enregistré dans la zone TEXTE, en version anaphorique, est significativement supérieur lorsque le genre grammatical du leurre est le même que celui des mots coréférentiels. Les temps plus longs peuvent signifier que, compte tenu de la proximité du leurre, les lecteurs éprouvent une difficulté dans l'établissement de la coréférence. Ainsi, ils interpréteraient d'abord le leurre comme coréférent possible du pronom, et ensuite seulement - correctement - comme coréférent du prénom plus éloigné; d'où cette augmentation des temps. Compte tenu de la situation décrite dans le texte, rappelons que les images excluent toute interprétation du leurre comme coréférent possible du pronom. Dans le cas où le leurre a un genre grammatical incompatible avec celui du pronom, l'effort serait moindre pour établir immédiatement le lien approprié avec le premier prénom.

On pourrait s'attendre à ce que les mêmes mécanismes pour identifier le lien coréférentiel approprié opèrent en version cataphorique puisque, tout comme en version anaphorique, le prénom-leurre y fait écran entre les deux éléments coréférentiels. Or, les résultats contredisent clairement une telle explication : lorsque le genre grammatical des coréférents potentiels est le même, le temps de lecture (panneau 1 de la figure 3) est significativement inférieur à celui obtenu avec genre grammatical différent. Encore une fois, les images excluent toute interprétation erronée des coréférents. Alors, comment expliquer les temps courts, résultat diamétralement opposé à ce qu'on observe en version anaphorique? On pourrait croire qu'en version cataphorique/genre identique, les lecteurs établissent rapidement un lien coréférentiel – même erroné - sur la base du principe de proximité. En version cataphorique/genre différent, ce principe de proximité ne peut opérer, de sorte que l'établissement du lien coréférentiel doit se faire avec un candidat éloigné, le seul candidat possible compte tenu des contraintes de genre. On pourrait dire que la présence d'un pronom exige sa résolution rapidement (Filik & Sanford, 2008). Comme le pronom est suivi d'un leurre de genre différent, le sujet tenterait d'établir la coréférence avec le nom qui suit. Cependant, le fait que celui-ci s'en distingue en genre entraîne un conflit cognitif. Cela expliquerait l'importance du recours à l'illustration dans cette situation. D'ailleurs, comparativement aux trois autres, c'est la seule situation qui entraîne un allongement significatif du temps de fixation dans la zone IMAGE (voir panneau 3 de la figure 3). L'ampleur de cet effet est tout aussi remarquable, allant du simple au triple (approximativement 2400 ms contre 7300 ms). Lorsque les prénoms possèdent les mêmes marques, le sujet établirait préférentiellement un lien avec le leurre sans les problèmes d'incompatibilité morphologique. L'établissement de la coréférence s'effectuerait ultérieurement par un rajustement moins énergivore : le recours à l'image ne serait plus aussi nécessaire.

**Revue pour la recherche en éducation**, actes de colloque, Acfas 2010, p. 1-18  
@2011, Ottawa, Canada

La comparaison des panneaux 1 (zone TEXTE) et 2 (sous-zone CORÉFÉRENT) révèle une ressemblance notable. Dans les deux zones, on a obtenu le même type d'interaction. Comme les deux autres sous-zones du texte n'indiquent pas d'effet d'interaction analogue, il semble légitime d'identifier la sous-zone CORÉFÉRENT comme le lieu textuel où se manifeste l'effet d'interaction. Pour le moment, faute d'avoir pu récupérer certaines informations relativement aux trajectoires visuelles, on peut faire l'hypothèse qu'en version «anaphorique avec genre identique», la coréférence erronée provisoirement établie entraînerait des retours en arrière vers la sous-zone du texte qui enregistre le temps prolongé nécessaire à la résolution du problème.

Par ailleurs, on constate que les deux sous-zones de l'image, contrairement à la zone globale IMAGE, ne révèlent pas d'effet significatif. Pour l'instant, on n'a donc pas réussi à dégager de lieux précis où les sujets tireraient des informations relativement aux données linguistiques. Cependant, présenter d'abord un pronom, puis interposer dans la chaîne thématique un prénom-leurre de genre différent, impliquerait davantage de difficulté pour résoudre la coréférence, d'où le recours nécessaire à l'image. Celle-ci contribuerait à clarifier le modèle mental que le lecteur met en place. L'examen exhaustif des trajectoires visuelles pourra éventuellement apporter des informations complémentaires quant à l'incidence de la complexité linguistique sur la cueillette des informations picturales en direct.

### *Prospective*

Cette première étude nous a permis de faire quelques constats. Avant d'étendre notre programme de recherche, nous devons solutionner les difficultés que nous avons rencontrées avec la technologie GazeTracker relativement à l'extraction des informations pertinentes des parcours oculaires. Nos résultats devront, par conséquent, être complétés. Par ailleurs, la fréquence d'échantillonnage du système FaceLab à 60 Hz pourrait être trop faible, selon quelques collègues. D'autres systèmes ont des fréquences d'échantillonnage variant entre 500 et 2000 Hz. Cette situation exige d'être mieux évaluée. De façon indépendante du système utilisé, nous avons aussi fait le constat que les zones d'intérêt dans le texte auraient dû être agrandies ou multipliées afin de mieux détailler les parcours oculaires. Par exemple, les petits mots comme les pronoms sont habituellement peu fixés, étant perçus avec la ou les fixations sur le mot précédent (Rayner, 1998). Effectivement, peu d'observations ont été mesurées à la zone PRONOM. Il faudrait également systématiser l'emplacement de chaque zone dans tous les textes afin de mieux contrôler leurs positions relativement aux déplacements des yeux sur la ligne de texte : par exemple, certaines zones PRONOM étaient en bout de ligne, d'autres au milieu. Concernant la conception du matériel, la relation de genre devra être explorée pour chaque situation (texte-image). Dans le contexte de notre étude, les situations étant soit «Genre identique», soit «Genre différent», il est possible que la relation de genre ait été confondue avec celles des situations, ce qui rend difficile la distinction entre l'effet de genre et celui des situations présentées. Une recherche exploratoire implique une certaine phase d'appropriation, de tâtonnement et aussi ... d'erreurs. Nos résultats sont probablement questionnables, mais la démarche développée ici ne s'en trouve pas pour autant invalidée. Notre réflexion s'est enrichie et des correctifs seront apportés pour les prochaines expériences.

**Revue pour la recherche en éducation**, actes de colloque, Acfas 2010, p. 1-18  
@2011, Ottawa, Canada

## **Conclusion**

La comparaison des zones textuelles et picturales au moyen d'une technique d'observation des mouvements oculaires peut être révélatrice de la fonction que les lecteurs accordent à l'image par rapport à certains phénomènes linguistiques. Les lecteurs ont-ils un intérêt en soi pour observer l'image, indépendamment des propriétés linguistiques du texte? Ou, au contraire, leur intérêt pour l'image n'est-il conditionné que par les caractéristiques du texte, afin de solutionner les difficultés de compréhension éprouvées lors de la lecture (une fonction compensatoire) ? Les résultats de l'étude semblent soutenir l'idée que le recours à l'image s'explique par des difficultés liées à la cohérence textuelle.

## Références

- Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S., & Trueswell, J. C. (2000). The rapid use of gender information: evidence of the time course of pronoun resolution from eyetracking. *Cognition*, 76, B13-B26.
- Bock, M., & Milz, B. (1977) Pictorial context and the recall of pronoun sentences. *Psychological Research*, 39, 203-220.
- Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still Improve students' learning from text. *Educational Psychology Review*, 14 (1), 5-26.
- Clark, H. H., & Sengul, C. J. (1979). In search of referents for nouns and pronouns. *Memory & Cognition*, 7, 35-41.
- Ehrlich, K., & Rayner, K. (1983). Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Verbal Learning & Verbal Behavior*, 22, 75-87.
- FaceLab 4 user manual (2008). Australie : Canberra.
- Filik, R., & Sanford, A. J. (2008). When is cataphoric reference recognised? *Cognition*, 107, 1112-1121.
- Garnham, A. (2001). *Mental models and the interpretation of anaphora*. Hove, England : PsychologyPress.
- GazeTracker Software Guide, v8 (2008). Eye Response Technologies Inc.
- Glenberg, A. M., & Kruley, P. (1992). Pictures and anaphora : evidence for independent processes. *Memory and Cognition*, 20(5), 461-471.
- Goldsmith, E. (1984). *Research into Illustration: An Approach and a Review*. Cambridge University Press.
- Gordon, P. C., & Hendrick, R. (1997). Intuitive knowledge of linguistic co-reference. *Cognition*, 62, 325-370.
- Gyselinck, V. (1996). Illustrations et modèles mentaux dans la compréhension de textes. *L'Année Psychologique*, 96, 495-516.
- Gyselinck, V., & Tardieu, H. (1999). The role of illustrations in text comprehension : what,

- when, for whom, and why? In H. van Oostendorp and S. R. Goldman (Eds.), *The Construction of Mental Representations During Reading* (pp. 195-218). N.J. : Lawrence Erlbaum Associates.
- Hannus, M., & Hyönä, J. (1999) Utilization of illustrations during learning of science textbook passages among low- and high-ability children. *Contemporary Educational Psychology*, 24, 95–123.
- Hegarty, M. (1992). The mechanics of comprehension and comprehension of mechanics. In K. Rayner (Ed.), *Eye movements and visual cognition* (pp. 428-443). New York: Springer-Verlag.
- Hupka, W. (1989). *Wort und Bild - Die Illustrationen in Wörterbüchern und Enzyklopädien. With an English Summary / Avec un résumé français*. Tübingen: Niemeyer (Lexicographica Series Maior, Band 22).
- Johnson-Laird, P. N. (1980). Mental models in cognitive science. *Cognitive Science*, 4, 71-115.
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading : From eye fixations to comprehension. *Psychological Review*, 87(4), 329-354.
- Kaiser, E., Runner, J. T., Sussman, R. S., & Tanenhaus, M. K. (2009). Structural and semantic constraints on the resolution of pronouns and reflexives. *Cognition*, 112 (1), 55-80.
- Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, 49, 133-159.
- Kazanina, N., Lau, E. F., Lieberman, M., Yoshida, M., & Phillips, C. (2007). The effect of syntactic constraints on the processing of backwards anaphora. *Journal of Memory & Language*, 56, 384-409.
- Kennison, S. M., Fernandez, E. C., & Bowers, J. M. (2009). Processing differences for anaphoric and cataphoric pronouns: Implications for theories of discourse processing. *Discourse Processes*, 46, 25-45.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95(2), 163-182.
- Knoeferle, P., Crocker M. W., Scheepers C., & Pickering M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: evidence from eye-movements in depicted events. *Cognition*, 95, 95-127.
- Revue pour la recherche en éducation**, actes de colloque, Acfas 2010, p. 1-18  
@2011, Ottawa, Canada

- Knoeferle, P., & Crocker, M. W. (2006). The coordinated interplay of scene, utterance, world knowledge: evidence from eye tracking. *Cognitive Science*, 30, 481-529.
- Levin, J. R., Anglin, G. J., & Carney, R. N. (1987). On empirically validating functions of pictures in prose. In D. M. Willows & H. A. Houghton (Eds.), *The Psychology of Illustration: I. Basic Research*, (pp. 51-85). New York: Springer.
- Libben, M. R., & Titone, D. A. (2009). Bilingual lexical access in context: Evidence from eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(3), 381-390.
- Rayner, K., & Pollatsek, A. (1987). Eye movements in reading: a tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading*. (pp. 327-362). Hillsdale, NJ: Erlbaum.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372-422.
- Rayner, K., Rotello, C. M., Stewart, A. J., Keir, J., & Duffy, S. A. (2001). Integrating text and pictorial information: Eye movements when looking at print advertisements. *Journal of Experimental Psychology: Applied*, 7, 219-226.
- Rayner, K. (2009). Eye movements in reading and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, 62(8), 1457-1506.
- Reinwein, J. (1998 - ...). L'illustration et le texte : revue analytique des recherches expérimentales (Illustrations and text : an analytical review of experimental studies). Site web accessible à l'URL : <http://www.images-words.net/> (module 1).
- Roy-Charland, A., Saint-Aubin, J., & Evans, M. A. (2007). Eye movements in shared book reading with children from kindergarten to grade four. *Reading & Writing*, 20, 909-931.
- Underwood, G. (2005). Eye fixations on pictures of natural scenes: Getting the gist and identifying the components. In G. Underwood (Ed.), *Cognitive processes in eye guidance* (pp. 163-187). New York: Oxford University Press Inc.
- Underwood, G., Jebbett, L., & K. Roberts. (2004). Inspecting pictures for information to verify a sentence: eye movements in general encoding and in focused search, *The Quarterly Journal of Experimental Psychology*, 57A (1), pp. 165-182.
- Van Gompel, R. P. G., & Liversedge, S. P. (2003). The influence of morphological information on cataphoric pronoun assignment. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 29(1), 128-139.
- Revue pour la recherche en éducation**, actes de colloque, Acfas 2010, p. 1-18  
@2011, Ottawa, Canada

Vonk, W., (1984). Eye movements during comprehension of pronouns. In A. G. Gale & F. Johnson (Eds.), *Theoretical and Applied Aspects of Eye Movement Research*, (pp. 203-212). Amsterdam: Elsevier Science Publishers B.V.