



**Mémoire de Recherche Master 2 PEPSCO
2016-2017**

**Signalisation d'un contenu textuel visuospatial :
effet sur la compréhension**

Présenté par
Frédéric Maillet (20804776)

Sous la direction de
Patrice Terrier (Pr)
et
Julie Lemarié (MCF, HDR)

UFR DE PSYCHOLOGIE

—
Département Psychologie Cognitive et Ergonomie

Sommaire

Résumé-Abstract

Introduction

Signalisation et compréhension	1
Imagerie mentale, visuelle et spatiale	4
Imagerie visuospatiale et langage écrit	5
Problématique et hypothèses	7

Méthodologie

Participants	9
Matériel	9
Procédure	10

Résultats

Entraînement	12
Scores de réussite	12
Temps de traitement	13

Discussion et perspectives

Bibliographie

Annexes

A Détail du matériel utilisé

B Prétest de la signalisation prosodique

C Prétest du matériel d'imagerie visuelle

D Vividness of Visual Imagery Questionnaire (VVIQ)

Résumé

Mise en forme visuelle des textes, la signalisation textuelle participe à l'effort de compréhension, entendue comme la construction d'une représentation cohérente de la situation décrite par le texte. Parce que la lecture ne mobilise pas seulement des représentations verbales ou propositionnelles, comprendre un texte peut avoir comme exigence de requérir à la fois la construction d'une représentation imagée d'un référent visuospatial, ainsi que le traitement d'un contenu textuel, d'abord objet visuel et dont le traitement d'une signalisation semble impliquer la composante visuospatiale de la mémoire de travail. Nous avons ici cherché à vérifier dans quelle mesure la signalisation visuelle d'un contenu textuel visuospatial pouvait interférer avec la compréhension.

Pour cela, nous avons proposé la lecture de la description verbale d'une configuration spatiale suivie d'une tâche de rotation mentale. Nous avons manipulé la présence et la pertinence d'une signalisation visuelle avec l'objectif de mettre en évidence une interférence dans une mesure on-line de la compréhension. L'analyse des résultats obtenus est en cohérence avec la littérature sur la signalisation, mais ne permettent ni de confirmer ni d'infirmer notre hypothèse. Ceci invite à repenser des paradigmes expérimentaux qui intègrent les sous-composantes du calepin visuospatial dans des protocoles plus écologiques.

Abstract

Visual textual formatting signalisation enhance reader comprehension which refer to the elaboration of a coherent representation of the situation described by the text. As reading not only involves verbal or propositional representation, reading comprehension may require the construction of an imagery representation of a visuospatial referent as well as a textual visual signalisation processing which seem to involve working memory visuospatial component. Our work here deals with the idea that visual signalisation of a visuospatial textual content may interfere with reading comprehension.

In order to check this hypothesis, we asked participants to read the verbal description of a spatial configuration before answering a mental rotation task. We controlled the presence and relevance of a textual signalisation of the spatial description, aiming to look for processing time interference. Although consistent with current literature, our results analysis fails to settle the argument. They lead us to look for more ecological experimental design likely to further investigate working memory visuospatial sub-components involved in signalisation processing.

Remerciements

Est déjà presque loin, mais toujours aussi vif, le souvenir d'un regroupement SED, option cognitive, en L3. Julie Lemarié introduisait les modèles de la compréhension du langage écrit. Cette révélation de la révolution cognitive mobilise mon parcours depuis et je suis heureux et très honoré que Julie Lemarié ait alors accepté de m'accompagner, car c'est le mot qui convient, dans la mise en oeuvre d'une démarche de recherche. Je lui adresse mes plus profonds remerciements ainsi qu'à Patrice Terrier pour avoir bien voulu s'associer à notre démarche et dont les réflexions pertinentes ont largement contribué à éclaircir mon argumentation.

On réalise probablement mal la complexité d'organisation d'un Master dans sa nouvelle maquette et je ne saurais trop remercier Maja Becker et Aline Chevalier pour leur présence, toujours réactive, et leur bienveillance.

Le temps du Master est un temps contraint. Par des bornes, dont une limite extrême culmine aujourd'hui, et par sa répartition nécessaire entre les différentes tâches assignées qui entraîne parfois une dilution des disponibilités. Aussi dense qu'elle ait pu être, cette année achève formellement une belle aventure de découverte de la psychologie en général et de la psychologie cognitive en particulier. Savoir être sérieux sans se prendre au sérieux est certainement ce qui convient le mieux à l'esprit de notre promotion dont l'ambiance chaleureusement malicieuse a grandement contribué à confirmer l'orientation que j'ai souhaité donner à mon parcours : la recherche.

A tous mes camarades de promo, merci... et à bientôt pour de nouvelles aventures :)

Introduction

Souvent réalisée visuellement, la signalisation renvoie à une mise en forme textuelle différenciée, un ensemble de signaux textuels — titres, sous-titres, mots en gras... — qui souligne l'organisation du texte et en favorise la compréhension (pour une revue voir [Lemarié, 2016](#)). Cette compréhension est décrite comme un « [...] *processus dynamique de construction en mémoire d'une représentation cohérente de la situation évoquée et à laquelle viennent s'ajouter les inférences générées* » ([Blanc et Brouillet, 2005](#), p.30). S'agissant du format de cette représentation, on peut judicieusement remarquer que, si produire un texte consiste souvent à traduire des représentations visuelles et spatiales en texte ([Olive et Passerault, 2012](#)), alors sa lecture doit nécessairement impliquer un processus inverse de reconstruction des représentations initiales, visuelles et spatiales, décrites par l'auteur (voir [Heurley, 1994](#); [Lemarié, 2016](#)).

Comprendre un texte peut donc avoir comme exigence de requérir à la fois la construction d'une représentation imagée d'un référent visuospatial, ainsi que le traitement d'un contenu textuel, d'abord objet visuel et dont l'organisation et la signalisation peut aussi être de l'ordre du visuospatial. Nous avons en effet montré que le traitement d'une signalisation visuelle (mots en gras) semblait bien impliquer la composante visuospatiale de la mémoire de travail ([Maillet, 2015](#)).

La perspective que l'on dessine ici invite alors naturellement à se demander dans quelle mesure la signalisation visuelle d'un contenu textuel visuospatial peut interférer sur la compréhension. Pour cela nous avons proposé la lecture d'un contenu textuel visuospatial, dans trois conditions de signalisation, suivie d'une tâche de rotation mentale. Une modalité orale contrôlée permettait de vérifier que l'effet, s'il était obtenu, était spécifique de la modalité visuelle (lecture). Notre variable d'intérêt est la compréhension, dont nous avons deux indicateurs :

1. une mesure on-line : le temps de traitement (lecture ou écoute) du contenu textuel visuospatial ; ce temps est interprété comme l'effort nécessaire à la construction d'une représentation cohérente de la situation visuospatiale décrite dans le texte ;
2. une mesure off-line : le score de réussite à la tâche de rotation mentale ; il ne peut être significativement meilleur que le hasard que si le participant a construit une représentation cohérente.

Signalisation et compréhension

Dans les textes expositifs, l'apparence visuelle d'une page est rarement homogène : des titres et sous-titres, une indentation des paragraphes, des listes, des mots soulignés ou en gras, ou en italique... traduisent l'intention de l'auteur d'accompagner le lecteur dans la compréhension de son texte et donc « *Du point de vue fonctionnel, les signaux (i. e. les éléments signalant l'organisation des textes) peuvent être vus comme des dispositifs participant à la construction du discours, des indications métadiscursives, des instructions de lecture et de traitement du texte, des traces reflétant les processus de rédaction, des indices révélant les intentions de l'auteur...* » ([Ho-Dac et al., 2012](#), p. 6). SARA (Signal Available Relevant Accessible information, [Lemarié et al., 2008](#)) propose un cadre conceptuel descriptif de la signalisation textuelle qui associe une analyse textuelle du signal en lien

avec une analyse du processus de lecture. Ce modèle se démarque d'une définition classique du signal comme étant essentiellement de nature visuelle (Meyer, 1975, in Lemarié et al., 2008) en proposant qu'en tant qu'implémentation dans le texte d'une métaphore issue du prototexte, le signal puisse être implémenté sur un continuum allant d'une réalisation purement discursive à une réalisation purement visuelle. Ces dispositifs de mise en forme visuelle du texte sont couramment employés sans toujours faire explicitement référence à la signalisation textuelle. Ainsi, Olive et Barbier (2017) mettent en évidence dans une tâche de prise de notes un effet facilitateur de l'utilisation de listes à puces (par rapport à un texte « brut » sans mise en forme) : la mise en forme du texte permet aux étudiants de mieux appréhender la structure du texte et réduit la charge cognitive associée à la lecture. Les signaux sont donc des éléments signalant l'organisation des textes, et invitent donc le lecteur à construire une structure conceptuelle hiérarchique du texte et leur effet facilitateur dans la construction d'une représentation de la structure thématique n'est plus à démontrer (pour une revue voir Lemarié, 2016).

Néanmoins, cette construction résulte de l'interaction texte-lecteur et « [...] *ce n'est pas parce qu'une fonction informative est rendue disponible par un signal au lecteur, qu'elle provoquera forcément les effets attendus a priori ;* » (Lemarié, 2016, p.23). Trois dimensions d'analyse sont proposées afin de caractériser cette interaction : disponibilité, pertinence et accessibilité (voir Figure 1) et nous retiendrons ici la pertinence « [...] *qui renvoie à l'utilité des fonctions informatives du signal en relation avec les buts poursuivis par le lecteur [...]* » (Lemarié, 2016, p.23). Cette pertinence du signal reste toutefois intimement liée à son accessibilité comme le démontre Lorch et al. (2013) dans une tâche de rappel de la structure thématique d'un texte : une signalisation visuelle (titres) donne de meilleures performances qu'une signalisation purement textuelle (phrases thématiques). En revanche, lorsque les exigences du lecteur sont élevées au regard de la tâche, la signalisation des contenus pertinents est utile et permet de diminuer la charge cognitive induite par la compréhension du texte (Britton et al., 1982 ; Loman & Mayer, 1983 in Lemarié, 2016).

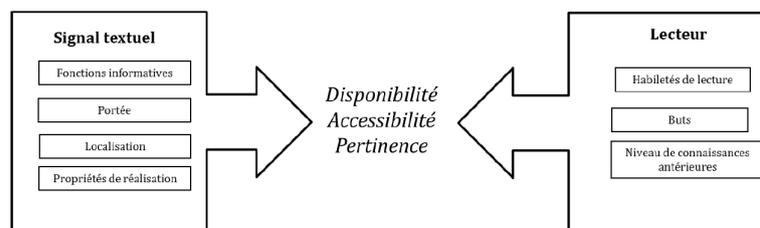


FIGURE 1. – Synthèse des dimensions de caractérisation des signaux dans l'interaction texte-lecteur (tiré de Lemarié, 2016, p.23)

Puisque la mise en forme visuelle d'un signal apporte une information au lecteur, il semblait pertinent de questionner les processus et les structures qui permettent au lecteur d'associer information visuelle et verbale et légitimement s'interroger sur les interactions au sein du calepin visuospatial. C'est ce travail que nous avons initié au cours du M1 (Maillet, 2015) en défendant l'hypothèse générale qu'une partie tout au moins du traitement cognitif du signal visuel est dévolue à la composante visuospatiale de la mémoire de travail. Afin de mettre en évidence ce traitement, nous avons utilisé comme indicateur le souvenir de la localisation des mots d'un texte lu. Bien qu'une reconstruction de la structure narrative soit toujours possible, ce souvenir de localisation repose en priorité sur une mémoire visuospatiale qui associe l'information textuelle à sa localisation dans le texte (pour une revue voir Le Bigot et al., 2010). La démarche expérimentale s'organisait donc autour de la recherche de

l'effet de la signalisation sur la compétence du lecteur à se souvenir préférentiellement de la localisation de mots signalés visuellement. Une première hypothèse proposait que les mots signalés par une mise en forme visuelle seraient mieux localisés que les mots non signalés. Nous avons ensuite placé le lecteur en situation de double tâche au sein du calepin visuospatial en introduisant un contenu textuel sollicitant l'imagerie visuelle. Ceci nous permettait de tester une seconde hypothèse de dégradation de la performance de localisation.

Soixante-deux étudiants ont participé volontairement à cette expérience en passations collectives. Chaque participant lisait successivement deux textes, à contenu spatial et non spatial, avant de procéder à une tâche de localisation de mots dictés : le lecteur devait inscrire chacun d'eux sur une feuille quadrillée à l'endroit précis où il pensait l'avoir lu. La moitié des mots était signalés dans le texte (mise en italique), l'autre moitié était non signalée. Un mot était correctement localisé si la case origine de la localisation rappelée était située à l'intérieur d'un carré d'environ 4 cm de côté autour du début du mot cible (Figure 2).

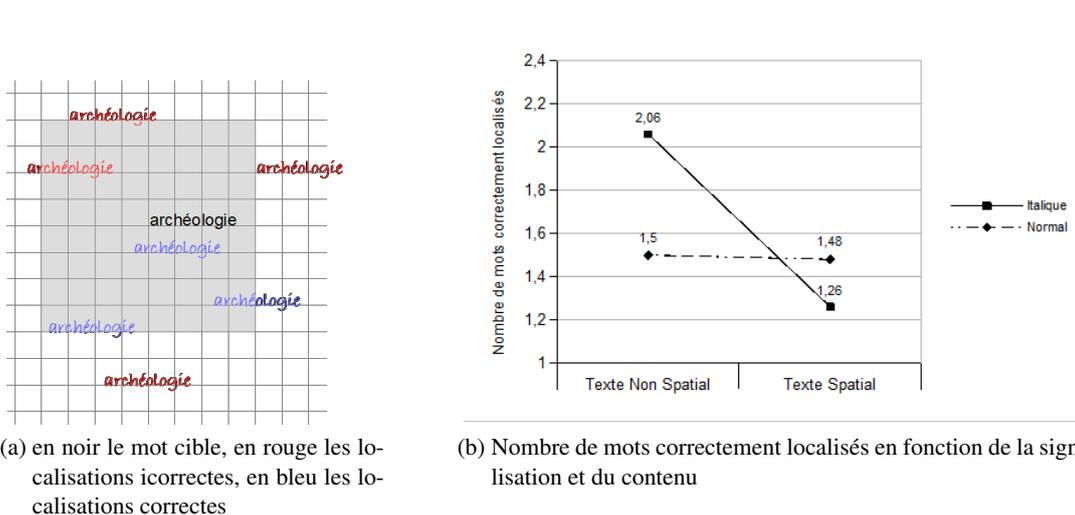


FIGURE 2. – Expérience de M1 (tiré de Maillet, 2015)

Une analyse de variance prenant en compte les deux facteurs constituant cette étude révélait un effet d'interaction « contenu * signalisation » ($F(1,61) = 9,914$; $p = .003$). Une comparaison par paire rapportait une différence significative du facteur signalisation lorsque le contenu est non spatial ($p < .001$) et du facteur contenu lorsque les mots sont signalés ($p < .001$) :

- Les mots signalés par une mise en forme visuelle sont mieux localisés que les mots non signalés ;
- Lorsque le texte a un contenu spatial, on observe une action très spécifique de l'effort d'imagerie visuelle sur le contenu signalé.

Cette dégradation observée de la performance de localisation des mots signalés dans le cas du texte spatial suggérait donc un traitement spécifique de la signalisation visuelle par le calepin visuospatial.

Une critique intéressante de ce travail a porté sur la différence de consigne entre les lectures des textes spatiaux et non spatiaux. En effet, afin de mobiliser certainement une représentation visuospatiale du contenu textuel, nous invitons nos participants à lire simplement « pour comprendre » dans la condition non spatiale, mais à lire « pour dessiner » dans la condition visuospatiale. Nous avons donc reproduit cette expérimentation à l'identique en homogénéisant les consignes de lecture

« pour comprendre » dans les deux conditions. Les résultats ne mettent plus en évidence d'interférence dans la condition spatiale. Bien que cela puisse renvoyer à l'engagement du lecteur dans la tâche en fonction de la consigne donnée, nous avons toutefois proposé une lecture qui reste compatible avec l'hypothèse initiale : l'absence de consigne explicite peut avoir incité les lecteurs à limiter l'effort de représentation visuospatiale. Dès lors, l'absence d'effet constaté lors de cette réplication peut s'expliquer par l'absence de compétition pour les ressources visuospatiales.

Imagerie mentale, visuelle et spatiale

L'accent mis sur les processus langagiers conduit souvent à considérer le comportement verbal explicite, celui qui est observé, comme modèle des processus verbaux implicites de la pensée (Paivio, 1971). Pourtant, on ne peut plus simplement considérer le système cognitif en terme de caractéristiques abstraites, verbales ou propositionnelles, comme le proposait Pylyshyn (1981). Aristote considérait déjà qu'il ne peut y avoir de pensée sans imagerie (Aristote, *de Anima*, traduction de Hicks, R. D., 1907). Ainsi, « *Théoriquement, les images organisent la pensée, ce qui conduit à une meilleure compréhension.* » (Kolker et Terwilliger, 1986) et Shepard illustre l'intervention de l'imagerie dans la pensée en décrivant la façon dont il comptait le nombre de fenêtres dans sa maison : « ... je dois imaginer la maison vue de différents côtés ou de l'intérieur des différentes pièces, et compter ensuite les fenêtres qui se présentent dans ces diverses images mentales. Un traitement purement verbal ne semble pas approprié » (1967, p.203, notre traduction). Et cela semble effectivement une habileté universelle : dans une étude auprès de 500 adultes (non-étudiants), ils sont 97% à rapporter des expériences d'imagerie visuelle, 93% d'imagerie auditive, 74% d'imagerie motrice, 70% tactile, 67% gustative, 66% olfactive, douleur 54% et température 43% (McKellar 1965a,b ; pour des résultats similaires chez des étudiants voir Marks, 1972) . Les enfants sont aussi 93% à rapporter des expériences d'imagerie visuelle (Carey, 1915).

L'imagerie mentale est définie comme notre habileté à générer, inspecter et manipuler une représentation mentale interne d'objets, événements et scènes qui ne sont pas physiquement présents (Sack et Schuhmann, 2012). Si nous retiendrons cette usuelle distinction entre percept et image mentale, distinction qui offre les possibilités d'une confortable opérationnalisation, Holt (1972) nous rappelle que (1) le percept n'est pas moins une image construite que ne l'est l'imagerie mentale, et que (2) cette dernière renvoie aussi à un « *contenu phénoménologique de nature sensorielle ou quasi sensorielle* » (Holt, 1972, p.10, notre traduction). Cette imagerie mentale est habituellement déclinée selon la modalité sensorielle qui fonde la représentation générée. Ainsi, les processus impliqués dans la génération, l'inspection et la manipulation d'images visuelles en l'absence de stimulus correspondants renvoient à l'imagerie visuelle mentale, ou imagerie visuelle (Finke, Ronald A., 1989). L'imagerie spatiale (ou visuospatiale) renvoie à l'inspection et l'évaluation de caractéristiques spatiales (distance, position relative, configuration...) et/ou la manipulation spatiale (rotation, déplacement, réorientation...) d'images visuelles générées mentalement (Sack et Schuhmann, 2012).

L'imagerie visuelle est alors une « *expérience cognitive caractérisée par l'activation d'une représentation d'un objet ou d'une scène en l'absence du stimulus correspondant* » (Fourtassi et al., 2016, p.1, notre traduction). Nous l'avons déjà souligné, elle se distingue de la perception par l'absence de stimulus, mais aussi des processus attentionnels (Thompson et al., 2011), et de la mémoire

visuelle par notre capacité à manipuler la représentation visuelle ainsi construite (Kosslyn et Koenig, 1992). L'imagerie visuelle est donc centrale à la cognition visuelle, laquelle est définie comme « l'utilisation de l'imagerie visuelle mentale dans la pensée » par Kosslyn et Koenig (1992, p. 128, notre traduction) qui ont suggéré que l'imagerie visuelle puisse être la conséquence d'un traitement opportuniste : elle a pu se développer initialement pour la reconnaissance d'objets, mais, une fois cette habileté disponible, elle aurait été recrutée pour supporter la réflexion (Kosslyn et Koenig, 1992).

Les travaux sur la cognition visuelle ont mis en évidence un codage isomorphique entre l'objet et sa représentation et, en support au caractère analogique de l'imagerie visuelle, il a été montré que les structures neurologiques visuelles les plus périphériques (aire V1) sont activées lors de l'évocation d'images mentales, aires qui sont associées à des représentations directement rétinotopiques (Gregory, 2004). Ainsi, pour Kosslyn (1994), un buffer visuel permet de traiter les informations visuelles en provenance de la rétine tout en servant de support aux représentations visuelles. Celles-ci peuvent alors subir quatre types de traitements : un processus de génération sous-tend la création de l'image dans le buffer visuel à partir des informations stockées dans la mémoire visuelle à long terme ; un processus d'inspection opère l'exploration des patterns pour récupérer les informations pertinentes (forme et configuration spatiale) ; enfin, d'autres processus transforment l'image (rotation, translation, zoom...) ou en permettent le maintien (rafraîchissement de l'image au sein du buffer). Le calepin visuospatial, composante visuospatiale de la mémoire de travail (*Working Memory* ou WM, voir Repovš et Baddeley, 2006, pour une revue), utilise les mêmes structures que la perception ou l'imagerie visuospatiale, laquelle est considérée comme un service prototypique fourni par le calepin visuospatial (Zimmer, 2008). Le calepin visuospatial se déclinerait en deux sous-composantes (Logie, 1995), l'une dédiée au stockage passif des objets visuels (le cache visuel), l'autre dédiée au rafraîchissement nécessaire au traitement de l'information visuelle (scribe interne). On précisera avec Beech (1984) que le calepin visuospatial semble bien puiser ses ressources autant dans l'une ou l'autre de ses deux sous-composantes, mais pas dans les composantes verbales de la mémoire de travail, et que, lorsque le texte à un contenu visuospatial, l'effort d'imagerie visuelle repose sur la composante visuelle de la mémoire de travail (voir par exemple Keogh et Pearson, 2011 ; Tong, 2013).

Imagerie visuospatiale et langage écrit

Il est admis qu'à la lecture d'un texte décrivant un environnement visuospatial, nous pouvons construire une représentation mentale de cet environnement absent (Olive et al., 2010 ; Kintsch, 1998 ; Denis, 1982) et que cette représentation mentale possède des caractéristiques fonctionnelles similaires à la perception visuelle (Denis et al., 2002). « *Qu'il s'agisse d'une carte ou d'un texte décrivant une configuration spatiale, les sujets s'engagent, dans les deux cas, dans la construction de représentations spatiales [...]* » (Blanc et Tapiero, 2000, p. 242). Mais, alors que la carte ou le dessin fournissent directement toutes les informations permettant la construction d'une représentation qualifiée de « modèle spatial », le texte implique la construction d'une représentation de niveau intermédiaire, un niveau sémantique (Blanc et Tapiero, 2000, p.243).

S'agissant d'évaluer la capacité d'un texte à induire une imagerie visuelle, si Gambrell (1981) a proposé une catégorisation basée sur le nombre de propositions du texte, la plupart des auteurs utilisent une catégorisation préalable des textes par des lecteurs (voir par exemple Eddy et Glass, 1981).

L'approche de [Kolker et Terwilliger \(1986\)](#) est notable : en se basant sur les travaux de Paivio sur la concrétude des mots ([Paivio et al., 1968](#); [Toglia et Battig, 1978](#)), les auteurs proposent la construction d'un score d'imagerie à priori du texte en calculant un ratio de mots à forte, versus faible, imagerie. Évalué ensuite par une épreuve de compréhension auprès d'un échantillon d'enfants de niveau CM2, ce score semble effectivement selon les auteurs différencier les textes à forte ou faible imagerie. Dans le cadre qui nous intéresse ici, s'agissant d'évaluer plus spécifiquement la capacité d'un texte à induire une imagerie plutôt spatiale que visuelle (voir supra), des éléments d'analyse linguistique de l'articulation textuelle des propositions nous offrent les possibilités d'une définition opérationnelle. Les organisateurs, ou connecteurs, spatiaux (au nord, au sud, à gauche, à droite...) sont des éléments fréquents dans la structuration de cette catégorie de textes descriptifs : « *la fréquence élevée d'organisateur spatiaux dans la description tient principalement au fait que toutes les séquences qui décrivent des paysages, des villes ou des maisons, en sont naturellement saturées.* » ([Adam, 1990](#), p. 163). Ainsi, nous proposons de retenir une catégorisation de textes « à contenu visuospatial », c'est-à-dire susceptible d'induire une imagerie visuospatiale, fonction du rapport de connecteurs spatiaux sur le nombre total de connecteurs du texte.

		3	4	
	1	2	5	
		7	6	
		8		

In the starting square put a 1
In the next square to the right put a 2

In the next square up put a 3
In the next square to the right put a 4

In the next square down put a 5
In the next square down put a 6

In the next square to the left put a 7
In the next square down put an 8

FIGURE 3. – Matrice et instructions de disposition spatiale (tiré de [Brooks, 1967](#), p.291)

Puisque donc un contenu textuel est susceptible d'induire une imagerie visuospatiale, un certain nombre d'auteurs se sont interrogés très tôt sur la qualité de la représentation, et donc de la compréhension du texte, élaborée à partir d'indications verbales. Dans une série de quatre études désormais célèbres, [Brooks \(1967\)](#) propose à ses participants de mémoriser la disposition spatiale de 8 chiffres répartis dans une grille 4x4 (Figure 3). Les indications spatiales étaient présentées soit à l'écrit, soit à l'écrit accompagné d'une lecture orale. Les participants devaient ensuite répéter les indications de disposition. Les résultats mettent en évidence un conflit entre la lecture et la visualisation mentale de l'organisation spatiale décrite. Cet effet est obtenu même si le lecteur a une connaissance préalable de la disposition spatiale, mais il n'apparaît pas dans la condition orale. Pour l'auteur, la lecture entre donc spécifiquement en compétition avec la représentation mentale de l'information spatiale et dégrade l'organisation spatiale, à l'encodage comme à la restitution.

Dans les travaux qui ont suivi, [Brooks \(1970\)](#) propose une tâche de rotation mentale d'une configuration spatiale en forme de croix (voir Figure 4) sous trois conditions : lecture, écoute et visualisation de la croix elle-même. Les résultats confirment que seule la lecture interfère avec l'activité spatiale de manipulation des relations spatiales. En particulier, [Brooks \(1970\)](#) souligne que, dans la condition visuelle de présentation les participants doivent aussi lire le nom des éléments pour assurer la tâche de rotation mentale, mais que ce mode de présentation n'interfère pas avec la tâche. Pour l'auteur, c'est donc bien le traitement d'une information verbale textuelle qui est à l'origine de

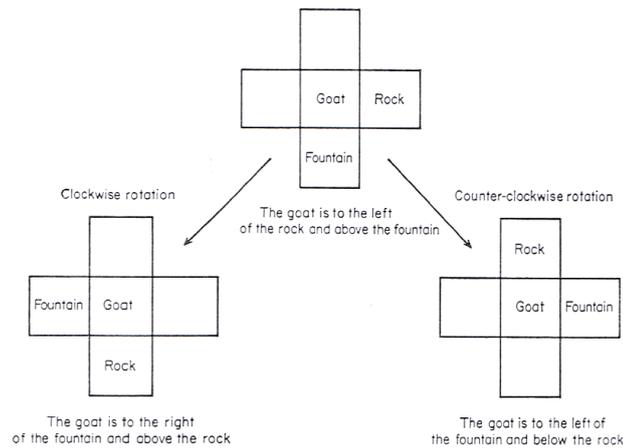


FIGURE 4. – Configuration spatiale en forme de croix (tiré de Brooks, 1970, p.92)

l'interférence observée.

Problématique et hypothèses

Nous avons vu que la signalisation textuelle pouvait avoir un effet facilitateur sur la compréhension du langage écrit, mais que cet effet dépendait d'un certain nombre de facteurs, notamment l'accessibilité et la pertinence du signal (Lemarié, 2016). Nous avons vu que l'imagerie mentale en général, visuelle et spatiale en particulier, participait de la pensée et qu'un contenu textuel pouvait induire une représentation visuospatiale. Comme par ailleurs la capacité des images à refléter adéquatement les objets qu'elles représentent est affectée par la charge de traitement créée par la structure intrinsèque de la description (Denis et Cocude, 1992), on devrait s'attendre à ce que la signalisation textuelle puisse faciliter la compréhension d'un texte à contenu visuospatial, bien que cet effet puisse être modéré par la pertinence de la signalisation.

Pourtant si, comme nous le pensons, la signalisation visuelle est traitée, au moins partiellement, par cette composante de la mémoire de travail (Maillet, 2015), le traitement de la signalisation devrait rentrer en compétition avec la construction d'une représentation d'un texte à contenu visuospatial. Il a par ailleurs déjà été mis en évidence une interférence entre l'imagerie visuospatiale et le traitement d'un message verbal sous forme écrite — recodage des informations spatiales textuelles — mais pas sous forme orale (Brooks, 1967, 1970).

L'hypothèse générale défendue ici est donc que le traitement d'une signalisation visuelle interfère spécifiquement dans la lecture avec la compréhension d'un texte à contenu visuospatial. Afin de mettre en évidence cette interférence, nous présentons à nos participants la description verbale d'une situation visuospatiale sous deux modalités, lecture et écoute, et nous manipulons la présence et la pertinence d'une signalisation pour contraster les effets attendus : nous fixons une accessibilité forte de cette signalisation en choisissant la mise en gras d'éléments pertinents ou non pertinents dans le texte.

Notre variable d'intérêt est la compréhension, dont l'évaluation « *requiert de s'interroger non pas sur ce que l'individu a retenu du texte mais bien ce qu'il en a construit* » (Blanc et Brouillet, 2005, p.28). Pour cela nous utilisons le temps de traitement dévolu par le participant, à la lecture ou à

l'écoute, de la description verbale, comme indicateur de l'effort de construction d'une représentation cohérente de la situation décrite, et donc comme mesure on-line de la compréhension. Le score de réussite à une tâche de rotation mentale de la représentation construite constitue une mesure off-line de cette compréhension et invite les participants à construire une représentation aboutie.

Selon notre hypothèse générale, le traitement d'une signalisation visuelle devrait mettre le lecteur en situation de double tâche au sein du calepin visuospatial et donc dégrader la compréhension. Nous devrions alors observer un allongement du temps de traitement pour maintenir la performance. Une signalisation non pertinente devrait majorer cet effet (voir Figure 5). Nous qualifierons cette hypothèse de « visuelle » et nous ne devrions observer ce pattern de données que dans la modalité lecture.

Les données de la littérature sur la pertinence de la signalisation invitent pourtant à une hypothèse plutôt canonique de la signalisation : par rapport à l'absence de signalisation, une signalisation pertinente favorise la compréhension et devrait donc diminuer le temps de traitement. À l'inverse, une signalisation non pertinente devrait dégrader la compréhension et donc rallonger le temps nécessaire à la construction de la représentation (voir Figure 5). Nous qualifierons cette hypothèse de « sémantique » et nous nous attendons à ne l'observer que dans la modalité orale.

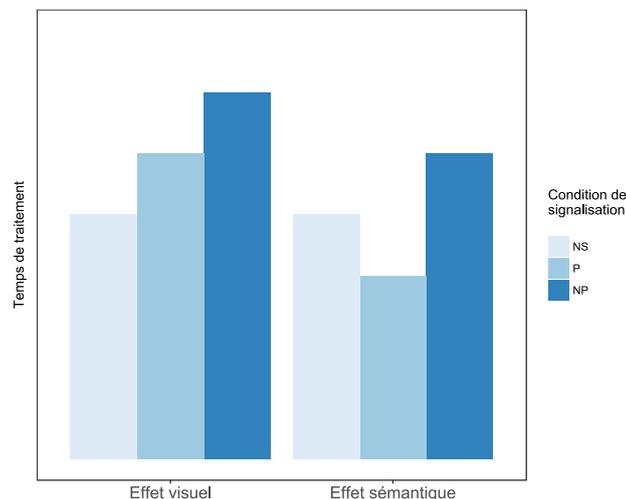


FIGURE 5. – Représentation de nos deux hypothèses opérationnelles selon l'effet attendu

Méthodologie

L'expérimentation a été réalisée sur ordinateur. Chaque participant devait initialement valider une page de consentement à participer anonymement et qui lui rappelait la possibilité qu'il conservait de quitter l'expérience, à tout moment, sans justification. Un débriefing permettait ensuite d'éclairer, s'il le souhaitait, le participant, et de recueillir certains éléments d'appréciation que l'on retrouvera discutés dans les résultats.

Participants

Le matériel et la procédure, décrits ci-après, ont tout d'abord été prétestés en ligne sur différents réseaux sociaux sur Internet (voir C). L'expérience a été réalisée en présence d'un expérimentateur et a mobilisé 39 étudiants de l'université Toulouse Jean-Jaurès et Paul Sabatier. Trois observations ont été retirées (temps de lecture inférieur à la seconde) et un temps de lecture extrême pour un des participants (interruption dans la tâche) a été corrigé et remplacé par la moyenne du participant. Nous avons donc 36 observations valides (âge moyen 23 ans, 9 hommes).

Matériel

La tâche proposée consistait en la lecture, ou en l'écoute, de trois phrases décrivant une scène visuelle que les participants étaient invités à se représenter avant de répondre à une question de localisation (voir exemple Figure 6). Les trois phrases précisaient les positions relatives de quatre objets concrets dans l'espace d'un carré, les deux premières situant les objets sur l'axe horizontal gauche-droite, la troisième levant l'ambiguïté de position verticale dessus-dessous (détail de la construction des jeux de phrases en Annexe A).

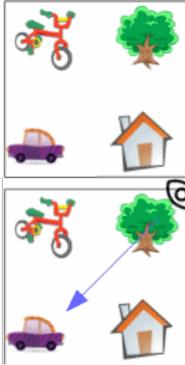
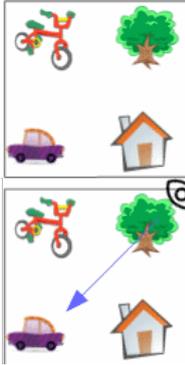
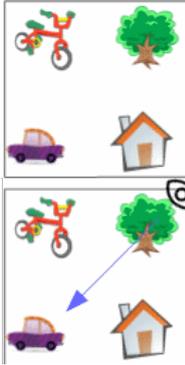
Non signalé	Le vélo rouge est à gauche de l'arbre vert	
	La maison blanche est à droite de la voiture bleu	
	L'arbre vert est au-dessus de la maison blanche	
Signal pertinent	Le vélo rouge est à gauche de l'arbre vert	
	La maison blanche est à droite de la voiture bleu	
	L'arbre vert est au-dessus de la maison blanche	
Signal non pertinent	Le vélo rouge est à gauche de l'arbre vert	
	La maison blanche est à droite de la voiture bleu	
	L'arbre vert est au-dessus de la maison blanche	
Question	Si vous êtes à la place de l'arbre et que vous regardez la voiture, la maison est-elle à votre gauche ou à votre droite ?	

FIGURE 6. – Exemple de situation visuospatiale dans les trois conditions de signalisation et question de localisation

Trois conditions de signalisation textuelle ont été mises en œuvre : des phrases sans signalisation (NS), une signalisation pertinente (P) et une signalisation non pertinente (NP). Trois essais (scènes visuelles différentes) ont été proposés par condition, soit un total de neuf scènes (NS1, NS2,

NS3, P1, P2, P3, NP1, NP2 et NP3) dont les présentations ont été aléatoirement contrebalancées (détail des scènes visuelles en Annexe A).

L'introduction d'une condition orale dans notre design expérimental nous a amenés à traduire à l'oral la signalisation d'un mot signalé dans un texte. On trouvera en Annexe B (p. V) les détails de l'implémentation choisie. On retiendra que nous avons utilisé une synthèse vocale pour assurer l'homogénéité des productions orales et que l'emphase prosodique disponible sur le moteur de synthèse vocale nous a servi d'équivalent pour le signal textuel. Un prétest nous a permis de valider cette équivalence (voir Annexe B, p. V).

Les trois phrases décrivant une scène étaient présentées simultanément, à la lecture ou à l'écoute, et le participant décidait quand il était prêt à passer à la question. Celle-ci lui suggérait de prendre un point de vue particulier dans la scène, avant de lui demander d'indiquer la position d'un des objets par rapport à lui-même (Figure 6, question). Cette question induit une manipulation de la représentation construite afin de répondre au mieux à la localisation demandée et vise ainsi à garantir l'effort nécessaire à une bonne construction de cette représentation mentale. Notre variable d'intérêt est donc le temps de lecture ou d'écoute de la description spatiale, considéré comme reflétant l'effort de construction de la représentation.

Ce matériel a été prétesté dans deux expériences proposées sur deux groupes Facebook professionnels (orthophonie et orthoptie). Une première expérience visait à tester l'association de la signalisation visuelle (mots en gras) avec la signalisation prosodique retenue. Les résultats de cette expérience permettent de considérer le choix de la signalisation prosodique retenue comme suffisamment isomorphe à la signalisation textuelle associée (voir Annexe B). La seconde expérience visait à valider, en condition lecture seulement, le matériel et la tâche d'imagerie visuospatiale. (voir Annexe C).

Procédure

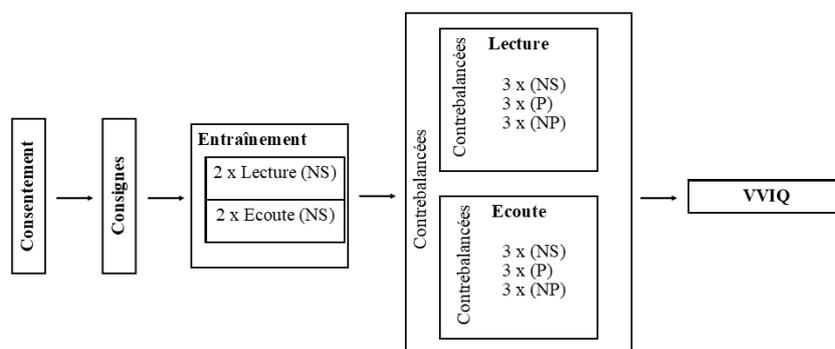


FIGURE 7. – Procédure expérimentale

La lecture des consignes, assortie d'un exemple illustré (Figure 6), introduisait la procédure (Figure 7). Quatre essais d'entraînement, non signalés, deux en modalité « lecture » et deux en modalité « écoute », dans cet ordre, permettaient au participant de se familiariser avec la tâche demandée avant de démarrer l'expérience proprement dite. L'ordre de présentation des modalités (lecture et écoute) était aléatoirement contrebalancé entre les participants, ainsi que les conditions de signali-

sation et les essais au sein de chaque modalité. Denis et Cocude (1989) remarquent que la qualité de la manipulation des images mentales dépend de l'exposition au matériel verbal : la qualité du modèle spatial construit s'accroît avec le nombre d'exposition. La répétition semble donc un paramètre important permettant d'aboutir à la construction d'une représentation visuospatiale susceptible d'être correctement manipulée. Nous choisissons donc de laisser à nos participants le soin de lire, ou d'écouter, autant de fois qu'ils le souhaitent, le matériel verbal présenté.

Il était demandé aux participants d'enchaîner, à leur rythme, mais sans interruption, les neuf essais d'une modalité. Une pause était aménagée entre les modalités et avant de passer au questionnaire d'habileté d'imagerie visuelle (VVIQ). L'ensemble de la procédure a été implémenté dans un programme JAVA qui assurait la succession des tâches et l'enregistrement des temps et des réponses.

Dans ce plan factoriel $2 * 3$ à mesures répétées, les facteurs intra-sujets sont la modalité de présentation (lecture ou écoute) et la signalisation : texte non signalé (NS), signalisation pertinente (P) et signalisation non pertinente (NP).

Les variables dépendantes sont le temps de traitement (lecture ou écoute) de la description de la scène visuospatiale et le score de réussite à la tâche de rotation mentale. Les variables contrôles sont l'âge, le genre et l'habileté d'imagerie visuelle (VVIQ).

Résultats

L'analyse a été réalisée avec le logiciel R (version 3.4.0, [R Core Team, 2017](#))¹ et a été conduite sur l'ensemble des observations, puis répétée sur le sous-ensemble de réponses correctes à la tâche de rotation mentale. Comme il n'y a pas de différence significative entre ces deux analyses, les résultats rapportés le sont pour l'ensemble des observations.

Entraînement

La consigne (temps moyen de lecture 40 sec., $SD = 16$ sec.) était illustrée d'un exemple et suivie de deux essais d'entraînement par modalité, sans signalisation (Training 1 et 2), toujours dans cet ordre, et dont les résultats sont rapportés dans le tableau 1. Les participants semblent avoir passés significativement plus de temps sur le premier entraînement oral ($t(36) = 3.02$; $p = .004$) que sur le précédent entraînement lecture. Les tests de proportionnalités (tous $p > .100$) confirment l'homogénéité des scores entre les conditions de signalisation et les modalités de traitement. Avec un score moyen à l'issue de la phase d'entraînement de près de 70%, significativement meilleur que le hasard (tests binomiaux, tous $p < .001$), on peut conclure qu'à l'issue de la phase d'entraînement, les participants ont compris et réussissent correctement la tâche demandée.

Scores de réussite

Le score de réussite à la tâche de rotation mentale constitue notre mesure off-line de la compréhension. Les participants réussissent significativement mieux que le hasard, autour de 75 % (tests binomiaux, tous $p < .001$), l'ensemble des situations proposées (Tableau 1). Ces scores de réussite ne diffèrent pas entre conditions et modalités (tests de proportionnalités, tous $p > .100$). Cette performance se maintient tout au long de l'expérience, quelles que soient les conditions, ce qui permet de conclure qu'à l'issue de leur traitement de la description textuelle ou orale, les participants ont construit une représentation suffisamment aboutie pour leur permettre d'assurer une performance continue.

Tableau 1

Temps de lecture moyens (écart-types) et taux de réponses correctes par conditions de signalisation et de modalités (toutes observations)

Condition	PréTest (Lecture)		Lecture		Ecoute	
	Temps	Score	Temps	Score	Temps	Score
Training 1	61.00 (36.65)	78 %	30.61 (12.28)	67 %	50.48 (21.18)	69 %
Training 2	56.94 (33.16)	65 %	35.95 (19.67)	64 %	41.24 (23.09)	75 %
NS	36.89 (17.43)	75 %	28.43 (14.50)	70 %	37.42 (14.68)	79 %
P	36.44 (19.28)	80 %	30.55 (14.63)	71 %	34.35 (11.56)	76 %
NP	43.51 (24.00)	75 %	31.54 (14.78)	81 %	41.19 (15.99)	74 %

1. Les références des packages R utilisés sont rapportées en bibliographie

Temps de traitement

Il est important de préciser que, comme dans le cadre du prétest, les distributions empiriques (Figure 9, p. 15) ne satisfont pas les critères de normalités (Shapiro-Wilk et Anderson-Darling², tous $p < .001$). Les temps moyens de traitement pour l'ensemble des participants selon les modalités sont rapportés dans le tableau 1 et la figure 8. Les résultats du prétest sont rappelés à titre d'information. Nous avons appliqué un test de Friedman, stratifié par sujets, pour évaluer les différences de temps de lecture ou d'écoute selon les conditions de signalisation.

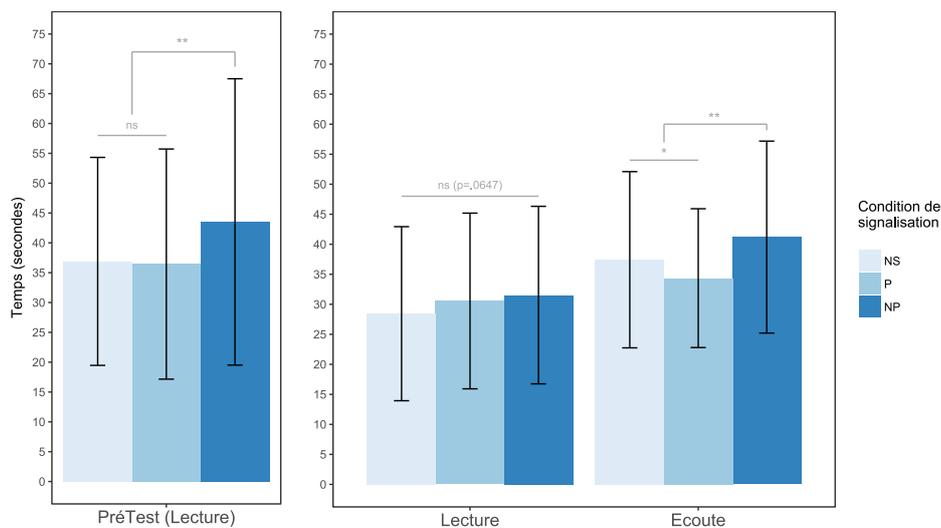


FIGURE 8. – Moyennes par conditions de signalisation et modalités (barres d'écart-tpes, significativités au test de Friedman)

Dans la condition Lecture, ce test n'est que marginalement significatif ($\chi^2(2) = 5.39, p = .0676$) mettant en évidence l'absence de réelle différence entre les conditions. Dans la condition Ecoute, le test de Friedman met en évidence une différence significative ($\chi^2(2) = 3.536, p = .001$) entre les conditions. Les tests post-hoc précisent que le temps de traitement de la condition non pertinente est significativement plus élevé que celui de la signalisation pertinente ($p = .0011$) et que le temps de traitement de la signalisation pertinente est significativement plus court que celui de l'absence de signalisation ($p = .0131$).

Une approche par comparaison de modèles linéaires généralisés à effets mixtes (ajustement par une distribution Gamma) permet d'analyser le rôle des variables contrôles. Notre modèle général a été le suivant :

$$Temps = \alpha + \beta_1.Signal + \beta_2.Genre + \beta_3.VVIQ + \beta_4.Score + \frac{1}{Sujet} + \frac{1}{Essai} + \epsilon$$

Dans ce modèle, *Temps* est notre variable dépendante (le temps de traitement), α sa valeur de base moyenne quand aucun signal n'est présent dans le texte ; β_1 représente la pente moyenne de l'effet du signal dans chacune de ses conditions ; β_2 représente la pente moyenne de l'effet du

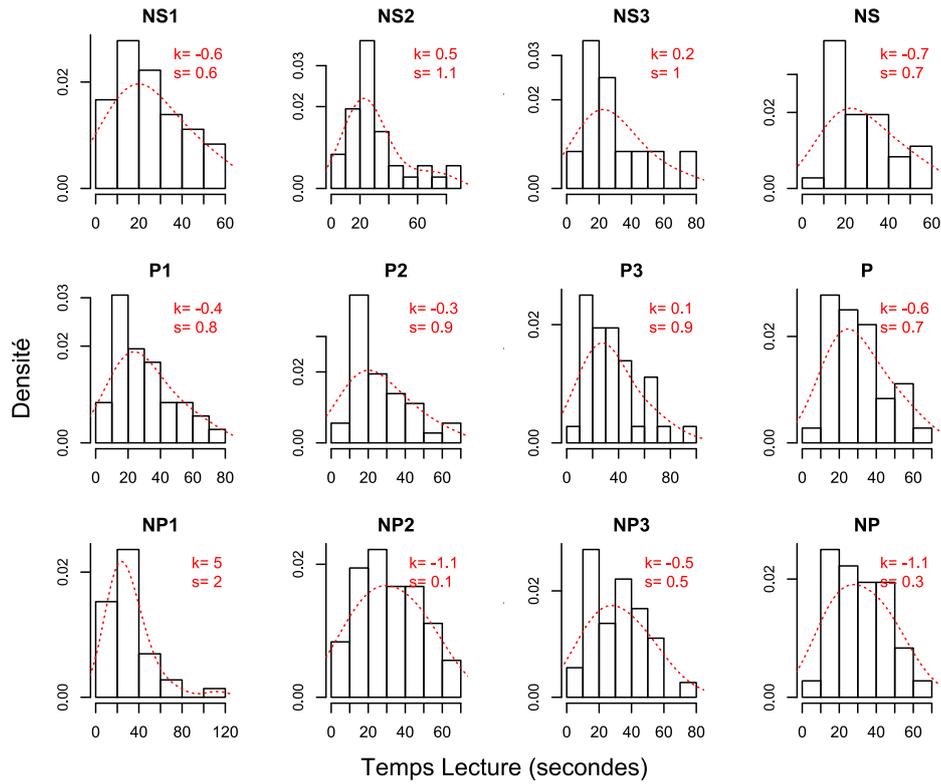
2. Bien que les résultats soient identiques entre ces deux tests, le test d'Anderson-Darling est ici théoriquement mieux adapté, car il accorde plus de poids à la queue de distribution

genre; β_3 représente la pente moyenne de l'effet d'habileté visuelle (voir Annexe D, p. XV); β_4 représente la pente moyenne de l'effet du score (réponse juste). Afin d'approcher l'hypothèse d'indépendance des observations, $\frac{1}{\text{Sujet}}$ et $\frac{1}{\text{Essai}}$ correspondent à la prise en compte dans le modèle des effets aléatoires introduits par les sujets eux-mêmes et la répétition des essais (trois essais par condition de signalisation). α est le résidu. Parmi les distributions disponibles, nous avons utilisé une distribution Gamma comme s'ajustant au mieux à nos distributions empiriques. Nous avons évalué la contribution de chacun des paramètres par la méthode du maximum de vraisemblance dans une série de modèles imbriqués afin de sélectionner le modèle le plus adapté.

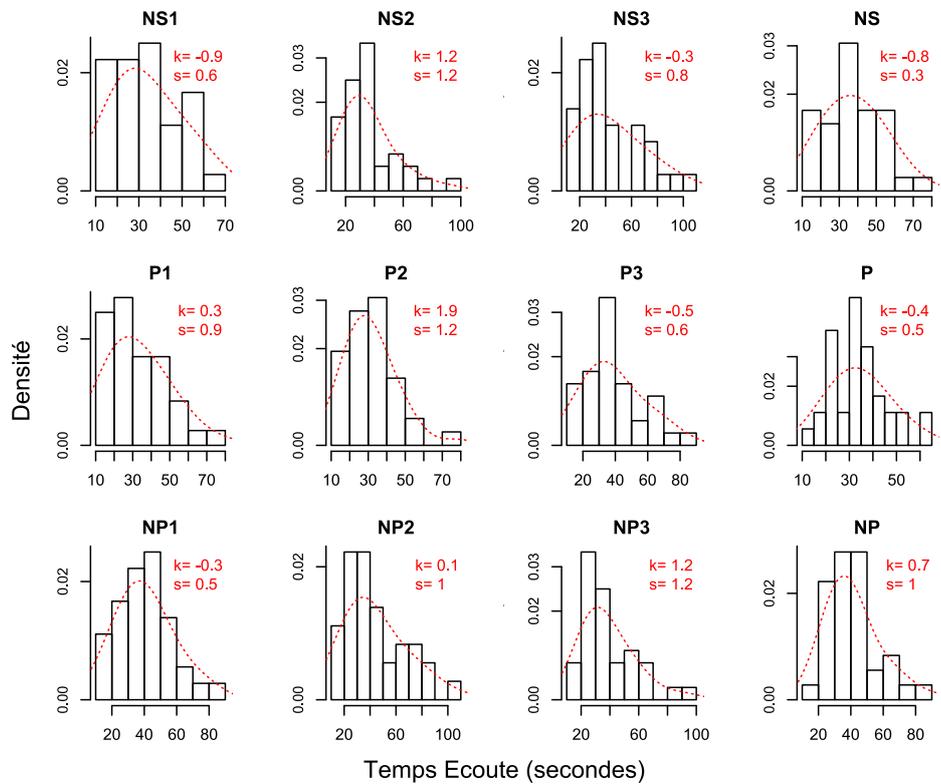
Dans la modalité « lecture », il ressort de cette analyse que seul le signal apporte une contribution marginale à l'ajustement du modèle aux données ($\chi^2(2) = 4.950, p = .0841$). Ni le genre ($\chi^2(1) = 0.328, p = .567$), ni l'habileté d'imagerie ($\chi^2(1) = 0.352, p = .553$) ne sont prédictif. L'essentiel des facteurs aléatoires est expliqué par le facteur sujet, 25 % de la variance contre moins de 1 % pour les essais, laissant entendre qu'il n'y a pas de différence significative dans la construction des situations visuospatiales proposées.

Dans la modalité « écoute », le signal apporte une contribution à l'ajustement du modèle aux données ($\chi^2(2) = 15.104, p < .001$). Ni le genre ($\chi^2(1) = 0.181, p = .670$), ni l'habileté d'imagerie ($\chi^2(1) = 1.081, p = .299$) ne sont prédictifs. L'essentiel des facteurs aléatoires est expliqué par le facteur sujet (25 % de la variance) contre moins de 1 % pour les essais, laissant entendre qu'il n'y a pas de différence significative entre les essais proposés : le choix des objets ne semble pas avoir induit de variabilité significative dans les situations spatiales, ce qui est cohérent avec des résultats rapportant une faible interaction entre l'habileté d'imagerie et la concrétude des mots ((Paivio, 1971, p. 508).

Suivant les recommandations de certains auteurs (voir par exemple Whelan, 2008; Balota et Yap, 2011), une analyse des distributions empiriques a été réalisée et aboutit à un ajustement par une distribution ex-Gaussienne. On en trouvera le principe dans l'analyse du prétest (voir Annexe C, p. X) mais son interprétation reste difficile et discutable et ne modifie pas les conclusions que l'on peut tirer des éléments ci-dessus.



(a) Modalité lecture



(b) Modalité Orale

FIGURE 9. – Distributions empiriques pour chaque essai et agrégées par conditions (NS, P, NP)

Discussion et perspectives

Ce travail a souhaité vérifier dans quelle mesure la signalisation visuelle d'un contenu textuel visuospatial pouvait interférer sur la compréhension, avec l'hypothèse que cette interférence provenait d'une compétition pour les ressources de la composante visuospatiale de la mémoire de travail. En effet, la signalisation lorsqu'elle est réalisée visuellement semble, au moins partiellement, traitée par le calepin visuospatial (Maillet, 2015), ce qui est aussi le cas de la construction d'une représentation visuospatiale à partir d'un contenu textuel visuospatial. Pour cela, nous avons proposé à nos participants la lecture ou l'écoute de la description verbale d'une configuration spatiale avant de les soumettre à une tâche de rotation mentale de cette configuration. Nous avons manipulé la présence et la pertinence d'une signalisation de la description verbale, dans les deux modalités, lecture et écoute, afin de contraster les effets attendus. La mesure du temps de traitement dévolu à la construction de la représentation visuospatiale constituait une mesure on-line de la compréhension dont le score de réussite à la tâche de manipulation spatiale constituait une mesure off-line.

Les scores de réussite pour l'ensemble des conditions de signalisation et des modalités de traitement sont homogènes et significativement supérieurs au hasard, ce qui laisse entendre que les participants ont compris la tâche demandée et ont réellement cherché à y répondre correctement, ce qui peut assurer la validité de la mesure on-line de la compréhension par le temps de traitement. Dans la modalité lecture, sur un plan strictement descriptif, les temps de traitement sont ordonnés dans le sens de notre hypothèse visuelle : $NP > P > NS$ (Figure 13). Toutefois, l'analyse statistique ne rapporte aucune différence significative entre les conditions de signalisation ce qui ne permet pas de conclure au regard de nos hypothèses. Dans la modalité écoute, nous observons une diminution du temps de traitement dans la condition de signalisation pertinente (P) par rapport à l'absence de signalisation (NS). A l'inverse, une signalisation non pertinente (NP) majore ce temps de traitement (Figure 13). Ce pattern de résultat $NP > NS > P$ est plutôt en faveur de l'hypothèse sémantique que nous attendions dans cette modalité (voir représentation des hypothèses Figure 8).

En conclusion, dans la modalité orale nous pouvons considérer que l'hypothèse canonique d'un effet sémantique de la signalisation sur la compréhension d'un texte à contenu spatial est confirmée. A contrario, dans la modalité lecture, nous n'observons aucun pattern réellement marqué et significatif permettant de trancher pour l'une ou l'autre des deux hypothèses envisagées.

Nous pouvons invoquer deux explications principales à ces résultats mitigés. La première renvoie à la tâche relativement coûteuse que nous avons proposée à nos participants : l'effort important et répété d'imagerie visuelle a probablement favorisé la mise en place de stratégies alternatives que notre design expérimental ne permettait pas de contrôler. De nombreux auteurs soulignent d'ailleurs qu'un double codage verbal et visuospatial des informations visuelles ne peut être complètement évité (voir Sadoski et Paivio, 2013), et certains de nos participants ont aussi évoquées des stratégies kinesthésiques d'encodage des positions. Le coût de la tâche, notamment dans le maintien de la représentation, pourrait donc avoir limité l'engagement du participant et masquerait donc l'effet recherché. On peut toutefois judicieusement observer que rien ne devrait empêcher d'observer ce coût dans la modalité orale.

Une deuxième explication renvoie à la distinction proposée par Logie (1995) : selon cet auteur, le calepin visuospatial se déclinerait en deux sous-composantes, l'une dédiée au stockage passif

des objets visuels (le cache visuel), l'autre dédiée au rafraîchissement nécessaire au traitement de l'information visuelle (scribe interne). Chose importante, l'information contenue dans le cache visuel serait codée sous une forme visuelle qui inclue les relations spatiales inter-item. Or, la tâche proposée à nos participants mobilisait essentiellement la composante active du calepin visuospatial, laquelle assure le rafraîchissement de la représentation. Mais le traitement de la signalisation visuelle peut n'avoir impliqué en dernier ressort que le cache visuel : si la signalisation visuelle porte la structure et l'organisation du texte, cette structure peut simplement être stockée dans le cache visuel puisque ce dernier comprend les relations spatiales entre les items.

A la lumière de ces réflexions, une relecture de nos résultats est donc possible : l'absence d'effet constaté dans la modalité lecture pourrait s'expliquer par l'absence de compétition entre les ressources disponibles au sein même du calepin visuospatial : le traitement de la signalisation mobilisant le cache visuel ; et l'imagerie visuospatiale, dans le cadre de notre tâche, mobilisant essentiellement le scribe interne. Cette explication peut sembler plus séduisante, car elle reste compatible avec les résultats observés dans la modalité orale.

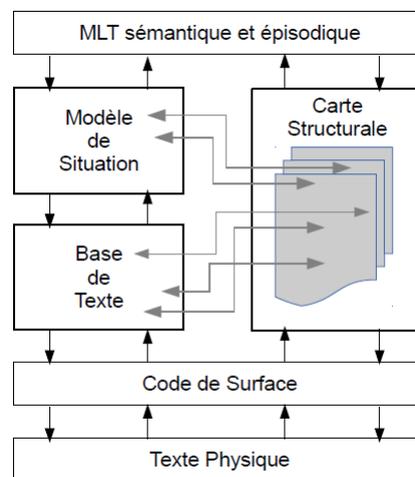


FIGURE 10. – Une représentation possible du modèle de la carte structurale (d'après [Payne et Reader, 2006](#))

L'étude de la signalisation textuelle est importante, car elle participe de la compréhension du langage écrit, et ceci est un enjeu majeur de l'intégration sociale et professionnelle contemporaine. Parce que la lecture ne mobilise pas seulement des représentations verbales ou propositionnelles abstraites, il nous semble nécessaire d'approfondir notre connaissance des structures qui sous-tendent le traitement de la signalisation, notamment celles qui constituent la composante visuospatiale de la mémoire de travail, afin de pouvoir mieux cibler la mise en oeuvre des différents procédés de signalisation en fonction du type de représentations que l'interaction contenu textuel–lecteur–tâche est susceptible de mobiliser. Réinvestir dans ce cadre la proposition de [Payne et Reader \(2006\)](#) selon laquelle le lecteur construit une double représentation du texte, linguistique et structurale (voir Figure 10), pourrait nous rapprocher d'un modèle de la lecture intégrant la signalisation.

Le travail conduit ici met en évidence la difficulté, soulignée par certains auteurs, de mobiliser les composantes visuospatiales en évitant des stratégies alternatives, notamment verbales (voir [Olive et al., 2010](#); [Sadoski et Paivio, 2013](#)). Les paradigmes expérimentaux devront être affinés, notamment au regard de leur validité écologique. On peut par exemple souligner avec [Willems et Jacobs \(2016\)](#),

une différence entre « simulation mentale » et « imagerie visuelle ». Nous avons vu que cette dernière renvoyait à la création délibérée d'une image relativement détaillée alors que la première renverrait à une représentation sensorimotrice plus indéterminée, mais plus spontanément invoquée en lecture. De plus, la simulation mentale semble plus importante au cours d'une lecture complète, contrairement à la lecture de phrases non reliées. Ceci interpelle des situations expérimentales sollicitant une représentation appauvrie et mobilisant des phrases ou des textes courts, décontextualisés. Un texte plus complet serait à même d'induire des résultats plus riches, augmentant la validité écologique.

Bibliographie

- Adam, J.-M. (1990). *Éléments de linguistique textuelle : théorie et pratique de l'analyse textuelle*. Philosophie et langage. Mardaga, Liège, 2e édition edition.
- Baayen, R. H. et Milin, P. (2015). Analyzing reaction times. *International Journal of Psychological Research*, 3(2) :12–28.
- Balota, D. A. et Yap, M. J. (2011). Moving Beyond the Mean in Studies of Mental Chronometry : The Power of Response Time Distributional Analyses. *Current Directions in Psychological Science*, 20(3) :160–166.
- Beech, J. R. (1984). The effects of visual and spatial interference on spatial working memory. *The Journal of general psychology*, 110(2) :141–149.
- Blanc, N. et Brouillet, D. (2005). *Comprendre un texte : L'évaluation des processus cognitifs*. In Press, Paris.
- Blanc, N. et Tapiero, I. (2000). Mode de présentation et organisation de l'information dans l'acquisition de connaissances spatiales. *L'année psychologique*, 100(2) :241–264.
- Brooks, L. R. (1967). The suppression of visualization by reading. *Quarterly Journal of Experimental Psychology*, 19(4) :289–299.
- Brooks, L. R. (1970). An extension of the conflict between visualization and reading. *Quarterly Journal of Experimental Psychology*, 22(2) :91–96.
- Burnham, K. P., Anderson, D. R., et Burnham, K. P. (2002). *Model selection and multimodel inference : a practical information-theoretic approach*. Springer, New York, 2nd ed edition. OCLC : ocm48557578.
- Campos, A. (2011). Internal consistency and construct validity of two versions of the Revised Vividness of Visual Imagery Questionnaire. *Perceptual and Motor Skills*, 113(2) :454–460.
- Campos, A. et Pérez-Fabello, M. J. (2009). Psychometric quality of a revised version Vividness of Visual Imagery Questionnaire. *Perceptual and Motor Skills*, 108(3) :798–802.
- Carey, N. (1915). Factors in the mental processes of school children. *British Journal of Psychology*, 1904-1920, 7(3) :453–490.
- Chaguiboff, J. et Denis, M. (1981). Activité d'imagerie et reconnaissance de noms provenant d'un texte narratif. *L'Année psychologique*, 81(1) :69–86.
- Coyle, T. R. (2003). A review of the worst performance rule : Evidence, theory, and alternative hypotheses. *Intelligence*, 31(6) :567–587.
- Denis, M. (1982). Imaging while reading text : A study of individual differences. *Memory & Cognition*, 10(6) :540–545.

- Denis, M., Beschin, N., Logie, R. H., et Sala, S. D. (2002). Visual perception and verbal descriptions as sources for generating mental representations : Evidence from representational neglect. *Cognitive Neuropsychology*, 19(2) :97–112.
- Denis, M. et Cocude, M. (1989). Scanning visual images generated from verbal descriptions. *European Journal of Cognitive Psychology*, 1(4) :293–307.
- Denis, M. et Cocude, M. (1992). Structural properties of visual images constructed from poorly or well-structured verbal descriptions. *Memory & Cognition*, 20(5) :497–506.
- Eddy, J. K. et Glass, A. L. (1981). Reading and listening to high and low imagery sentences. *Journal of Verbal Learning and Verbal Behavior*, 20(3) :333–345. 00094.
- Finke, R. A. (1980). Levels of equivalence in imagery and perception. *Psychological Review*, 87(2) :113–132.
- Finke, Ronald A. (1989). *Principles of mental imagery*. The MIT Press.
- Foos, P. W. (1980). Constructing cognitive maps from sentences. *Journal of Experimental Psychology : Human Learning and Memory*, 6(1) :25–38.
- Fourtassi, M., Rode, G., et Pisella, L. (2016). Using eye movements to explore mental representations of space. *Annals of Physical and Rehabilitation Medicine*.
- Fox, J. et Weisberg, S. (2011). *car : An R Companion to Applied Regression*. CRAN.
- Galili, T. (2015). *dendextend : an R package for visualizing, adjusting, and comparing trees of hierarchical clustering*. *Bioinformatics*.
- Gambrell, L. B. (1981). Induced Mental Imagery and the Text Prediction Performance of First and Third Graders.
- Gregory, R. L. (2004). Mental Imagery. In *The Oxford Companion to the Mind*. Oxford University Press, second edition.
- Hatzipanayioti, A., Galati, A., et Avraamides, M. N. (2016). Updating spatial relations to remote locations described in narratives. *Memory & Cognition*, 44(8) :1259–1276.
- Hennig, C. (2015). *fpc : Flexible Procedures for Clustering*.
- Herment, S. (2010). Emphase prosodique et emphase syntaxique : le cas de « do » dans un corpus de parole naturelle. *Corela. Cognition, représentation, langage*, (HS-9).
- Heurley, L. (1994). *Traitement de textes procéduraux : étude de psycholinguistique cognitive des processus de production et de compréhension chez des adultes non experts*. PhD thesis, Dijon.
- Hicks, R. D. (1907). *Aristotle De Anima*. Cambridge University Press, 2015 edition.
- Ho-Dac, L.-M., Lemarié, J., Péry-Woodley, M.-P., et Vergez-Couret, M. (2012). Multidisciplinary Perspectives on Signalling Text Organisation : Introduction to the Special Issue. *Discours. Revue de linguistique, psycholinguistique et informatique*, (10).

- Holt, R. R. (1972). On the nature and generality of mental imagery. In *The function and nature of imagery*, pages 3–33.
- Keogh, R. et Pearson, J. (2011). Mental Imagery and Visual Working Memory. *PLoS ONE*, 6(12) :e29221. 00028.
- Kintsch, W. (1998). *Comprehension : A Paradigm for Cognition*. Cambridge University Press. Google-Books-ID : LuycnLrY3k8C.
- Kolker, B. et Terwilliger, P. N. (1986). Visual imagery of text and children's processing. *Reading Psychology*, 7(4) :267–277.
- Korkmaz, S., Goksuluk, D., et Zararsiz, G. (2014). MVN : An R Package for Assessing Multivariate Normality. *The R Journal*, 6(2) :151–162.
- Kosslyn, S. M. (1994). *Image and brain*. The MIT Press.
- Kosslyn, S. M. et Koenig, O. (1992). *Wet mind : The new cognitive neuroscience*. Simon and Schuster.
- Le Bigot, N., Passerault, J.-M., et Olive, T. (2010). Le souvenir de la localisation des mots d'un texte. *L'Année psychologique*, 110(02) :321. 00003.
- Lemarié, J. (2008). La signalisation de la structure des textes comme élément facilitateur des traitements de l'information. Toulouse.
- Lemarié, J. (2016). *Signalisation, compréhension de textes et apprentissage multimédia*. Rapport HDR, Université Jean-Jaurès, Toulouse II.
- Lemarié, J., Eyrolle, H., et Cellier, J.-M. (2006). Visual signals in text comprehension : How to restore them when oralizing a text via a speech synthesis ? *Computers in Human Behavior*, 22(6) :1096–1115.
- Lemarié, J., Lorch, R. F., Eyrolle, H., et Virbel, J. (2008). SARA : A Text-Based and Reader-Based Theory of Signaling. *Educational Psychologist*, 43(1) :27–48.
- Lemarié, J., Lorch Jr, R. F., et Péry-Woodley, M.-P. (2012). Understanding How Headings Influence Text Processing. *Discours. Revue de linguistique, psycholinguistique et informatique*, (10).
- Logie, R. H. (1995). *Visuo-Spatial working memory*. Lawrence Erlbaum Associates, Incorporated.
- Lorch, R. F., Lemarié, J., et Chen, H.-T. (2013). Signaling topic structure via headings or preview sentences. *Psicología Educativa*, 19(2) :59–66.
- Luce, R. D. (1986). *Response times : Their role in inferring elementary mental organization*. Oxford University Press on Demand.
- Ma, T., Holden, J. G., et Serota, R. A. (2016). Distribution of human response times. *Complexity*, 21(6) :61–69.
- Macaya, M. et Perea, M. (2014). Does bold emphasis facilitate the process of visual-word recognition ? *The Spanish journal of psychology*, 17 :E2.

- Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., et Hornik, K. (2016). *cluster, Cluster Analysis Basics and Extensions*.
- Maillet, F. (2015). *Influence de la signalisation d'un texte et de l'imagerie visuelle sur la mémoire de localisation des mots*. Mémoire de master 1 mention psychologie cognitive, sous la direction de Julie Lemarié (MCF), Université Jean-Jaurès, Toulouse II.
- Marks, D. F. (1972). Individual differences in the vividness of visual imagery and their effect on function. In *The function and nature of imagery*, pages 83–108.
- Marks, D. F. (1973). Visual Imagery Differences in the Recall of Pictures. *British Journal of Psychology*, 64(1) :17–24.
- Marks, D. F. (1995). New directions for mental imagery research. *Journal of Mental Imagery*, 19(3-4) :153–167.
- McKellar, P. (1965a). The investigation of mental images. *Penguin Science Survey. Harmondsworth, England : Penguin Books*.
- McKellar, P. (1965b). Thinking, remembering, and imagining. *Modern perspectives in child psychiatry. Edinburgh : Oliver & Boyd*, pages 170–191.
- Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., et Leisch, F. (2015). e1071 : Misc Functions of the Department of Statistics, Probability Theory Group (Formerly E1071), TU Wien. *CRAN R Projects*. R package version 1.6-7.
- Olive, T. et Barbier, M.-L. (2017). Processing Time and Cognitive Effort of Longhand Note Taking When Reading and Summarizing a Structured or Linear Text. *Written Communication*, 34(2) :224–246.
- Olive, T., Lebrave, J.-L., Passerault, J.-M., et Le Bigot, N. (2010). La dimension visuo-spatiale de la production de textes : approches de psychologie cognitive et de critique génétique. *Langages*, 177(1) :29. 00001.
- Olive, T. et Passerault, J.-M. (2012). The visuospatial dimension of writing. *Written Communication*, 29(3) :326–344.
- Paivio, A. (1971). *Imagery and verbal processes*. Holt, New York. 06798.
- Paivio, A., Yuille, J. C., et Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of experimental psychology*, 76(1p2) :1.
- Palmer, E. M., Horowitz, T. S., Torralba, A., et Wolfe, J. M. (2011). What are the Shapes of Response Time Distributions in Visual Search? *Journal of experimental psychology. Human perception and performance*, 37(1) :58–71.
- Parris, B. A., Dienes, Z., et Hodgson, T. L. (2013). Application of the ex-Gaussian function to the effect of the word blindness suggestion on Stroop task performance suggests no word blindness. *Frontiers in Psychology*, 4.

- Payne, S. J. et Reader, W. R. (2006). Constructing structure maps of multiple on-line texts. *International Journal of Human-Computer Studies*, 64(5) :461–474. 00008.
- Pison, G., Struyf, A., et Rousseeuw, P. J. (1999). Displaying a clustering with CLUSPLOT. *Computational Statistics & Data Analysis*, 30(4) :381–392.
- Pylyshyn, Z. W. (1981). The imagery debate : Analogue media versus tacit knowledge. *Psychological Review*, 88(1) :16–45.
- R Core Team (2016). foreign : Read Data Stored by Minitab, S, SAS, SPSS, Stata, Systat, Weka, dBase, ... *CRAN R Projects*. R package version 0.8-67.
- R Core Team (2017). *R : A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. rbib.
- Repovš, G. et Baddeley, A. (2006). The multi-component model of working memory : Explorations in experimental cognitive psychology. *Neuroscience*, 139(1) :5–21.
- Revelle, W. (2016). psych : Procedures for Psychological, Psychometric, and Personality Research. *CRAN R Projects*. R package version 1.6.12.
- Richardson, J. T. E. (1999). *Imagery*. Psychology Press.
- Rohrer, D. et Wixted, J. T. (1994). An analysis of latency and interresponse time in free recall. *Memory & Cognition*, 22(5) :511–524.
- Sack, A. T. et Schuhmann, T. (2012). Hemispheric Differences within the Fronto-Parietal Network Dynamics Underlying Spatial Imagery. *Frontiers in Psychology*, 3(214).
- Sadoski, M. et Paivio, A. (2013). *Imagery and Text : A Dual Coding Theory of Reading and Writing*. Routledge. Google-Books-ID : Sk4G_0XAfrUC.
- Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H.-M., et Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology : General*, 136(3) :414–429.
- Shepard, R. N. (1967). Recognition memory for words, sentences, and pictures. *Journal of verbal Learning and verbal Behavior*, 6(1) :156–163.
- Taylor, H. A. et Tversky, B. (1992). Descriptions and depictions of environments. *Memory & Cognition*, 20(5) :483–496.
- Thompson, W. L., Hsiao, Y., et Kosslyn, S. M. (2011). Dissociation between visual attention and visual mental imagery. *Journal of Cognitive Psychology*, 23(2) :256–263.
- Toglia, M. P. et Battig, W. F. (1978). *Handbook of semantic word norms*. Lawrence Erlbaum.
- Tong, F. (2013). Imagery and visual working memory : one and the same? *Trends in cognitive sciences*, 17(10) :489–490. 00005.

- Tse, C.-S., Balota, D. A., Yap, M. J., Duchek, J. M., et McCabe, D. P. (2010). Effects of healthy aging and early stage dementia of the Alzheimer's type on components of response time distributions in three attention tasks. *Neuropsychology*, 24(3) :300.
- Whelan, R. (2008). Effective analysis of reaction time data. *The Psychological Record*, 58(3) :475.
- Willems, R. M. et Jacobs, A. M. (2016). Caring about Dostoyevsky : the untapped potential of studying literature. *Trends in cognitive sciences*, 20(4) :243–245.
- Zimmer, H. D. (2008). Visual and spatial working memory : from boxes to networks. *Neuroscience and biobehavioral reviews*, 32(8) :1373–95. 00085.

Annexes

A.Détail du matériel utilisé

Le matériel utilisé dans notre expérience est constitué de situations visuospatiales décrites par trois courtes phrases que le participant doit lire à son rythme avant de répondre ensuite à une question de localisation.

La construction de cartes cognitives à partir de phrases entendues dépend de l'ordre dans lequel les relations sont présentées [Foos \(1980\)](#). Reprenant les conclusions de [Taylor et Tversky \(1992\)](#) qui soulignent que, contrairement à une condition figurative, l'ordre de présentation des informations, dans une description textuelle participe à l'acte de communication des informations spatiales, nous appliquerons une condition de présentation non linéaire ([Blanc et Tapiero, 2000](#)) afin de ne pas favoriser une stratégie verbale pour palier la difficulté de la construction d'imagerie visuelle : le lecteur ne peut alors faire l'économie d'une phrase.

La construction des jeux de phrases a suivi un double schéma contrebalancé entre les situations afin que la disposition spatiale des objets dans l'espace ne corresponde pas strictement à la disposition spatiale des mots dans le texte. En partant de la disposition spatiale des objets (A, B, C, D) un premier pattern de construction a privilégié une description verbale plutôt consistante dans le sens horaire. Un deuxième pattern de construction a privilégié un pattern de description plutôt inconsistent dans le sens antihoraire (voir [Figure A.1](#)). Trois conditions de signalisation ont été mises en oeuvre : une condition sans signalisation, un signalisation des connecteurs spatiaux (à gauche, à droite...), considérée comme pertinente au regard de la tâche à accomplir, et une signalisation des qualificatifs des objets (chat noir) considérée comme non pertinente. La réalisation du signal s'est faite par la mise en gras de l'élément afin d'assurer une accessibilité forte par la prégnance visuelle de la typographie grasse sur le reste du texte et un traitement perceptif plus rapide ([Macaya et Perea, 2014](#)).

On trouvera dans les tableaux qui suivent, l'ensemble des jeux de situations utilisés dans notre expérience, en condition lecture ([Tableau A.1 & A.3](#)) et en condition écoute ([A.2 & A.4](#)).

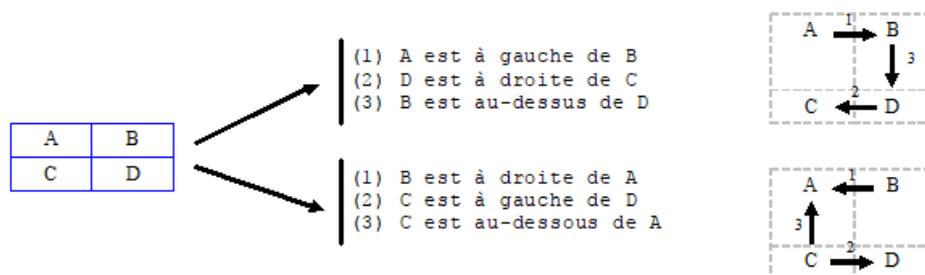


FIGURE A.1. – Mode de construction des descriptions verbales

Tableau A.1

Situations décrites dans la modalité lecture : trois situations par conditions de signalisation

Condition	Situation décrite & question associée
NS	Le chat gris est à gauche du chien marron Le tapis usé est à droite de la niche blanche Le chien marron est au-dessus du tapis usé
	Si vous êtes à la place du tapis et que vous regardez le chat, la niche est à votre ? (gauche)
	Le pain tranché est à droite du torchon blanc La bière brune est à gauche du saucisson sec La bière brune est au-dessous du torchon blanc
	Si vous êtes à la place du torchon et que vous regardez le saucisson, la bière est à votre ? (droite)
P	Le livre rouge est à gauche du cahier vert La trousse ouverte est à droite du stylo noir Le cahier vert est au-dessus de la trousse ouverte
	Si vous êtes à la place du cahier et que vous regardez le stylo, la trousse est à votre ? (gauche)
	La fille brune est à droite du sapin blanc Le garçon blond est à gauche du cadeau rouge Le garçon blond est au-dessous du sapin blanc
	Si vous êtes à la place du sapin et que vous regardez le cadeau, le garçon est à votre ? (droite)
NP	Le moulin fortifié est à gauche de la maison blanche L'école primaire est à droite de la piscine couverte La maison blanche est au-dessus de l'école primaire
	Si vous êtes à la place de la piscine et que vous regardez la maison, le moulin est à votre ? (gauche)
	L'étoile noire est à droite du soleil rouge La lune blanche est à gauche du nuage gris L'étoile noire est au-dessus du nuage gris
	Si vous êtes à la place de l'étoile et que vous regardez la lune, le soleil est à votre ? (droite)
NP	La balle jaune est à gauche de la raquette cassée Le filet blanc est à droite du joueur fatigué Le filet blanc est au-dessous de la raquette cassée
	Si vous êtes à la place du filet et que vous regardez la balle, le joueur est à votre ? (gauche)
	L'assiette creuse est à droite de la fourchette grise La cuillère inox est à gauche du verre vide La cuillère inox est au-dessous de la fourchette grise
	Si vous êtes à la place de l'assiette et que vous regardez la cuillère, la fourchette est à votre ? (droite)
NP	Le téléphone rouge est à gauche de la loupe carrée L'agenda fermé est à droite de la tasse vide La loupe carrée est au-dessus de l'agenda fermé
	Si vous êtes à la place de la loupe et que vous regardez la tasse, l'agenda est à votre ? (gauche)

Tableau A.2

*Situations décrites dans la modalité orale : trois situations par conditions de signalisation
(l'emphase prosodique est reproduite ici par la mise en gras correspondante)*

Condition	Situation décrite & question associée
NS	Le bouclier vernis est à gauche de l'armure lourde Le casque rouillé est à droite de l'épée cassée L'armure lourde est au-dessus du casque rouillé
	Si vous êtes à la place de l'armure et que vous regardez l'épée, le casque est à votre ? (gauche)
	La porte fermée est à droite du tableau peint Le livre rouge est à gauche de la chaise bleue Le livre rouge est au-dessous du tableau peint
	Si vous êtes à la place de la porte et que vous regardez le livre, le tableau est à votre ? (droite)
P	La bague verte est à droite de la montre cassée Le bracelet noir est à gauche du collier rond Le bracelet noir est au-dessous de la montre cassée
	Si vous êtes à la place de la montre et que vous regardez le collier, la bague est à votre ? (gauche)
	L'âne gris est à gauche de la clôture verte Le ruisseau chantant est à droite de la fleur fanée La clôture verte est au-dessus du ruisseau chantant
	Si vous êtes à la place de la clôture et que vous regardez la fleur, l'âne est à votre ? (droite)
NP	La pomme verte est à gauche de l'orange mûre La corbeille vide est à droite de la banane cuite L'orange mûre est au-dessus de la corbeille vide
	Si vous êtes à la place de la corbeille et que vous regardez la pomme, la banane est à votre ? (gauche)
	Le feutre usé est à droite du crayon rouge Le pinceau fin est à gauche du stylo noir Le pinceau fin est au-dessous du crayon rouge
	Si vous êtes à la place de feutre et que vous regardez le pinceau, le crayon est à votre ? (droite)
NP	La chemise blanche est à droite du chapeau rond Le manteau noir est à gauche du foulard rouge Le manteau noir est au-dessous du chapeau rond
	Si vous êtes à la place du manteau et que vous regardez la chemise, le chapeau est à votre ? (gauche)
	Le tableau vide est à gauche de la pendule ronde La cheminée éteinte est à droite de la chaise cassée La pendule ronde est au-dessus de la cheminée éteinte
	Si vous êtes à la place de la chaise et que vous regardez la pendule, la cheminée est à votre ? (droite)
NP	La feuille verte est à gauche de l'abeille fatiguée La fleur jaune est à droite de la ruche cachée L'abeille fatiguée est au-dessus de la fleur jaune
	Si vous êtes à la place de la feuille et que vous regardez la fleur, l'abeille est à votre ? (gauche)

Tableau A.3

Scènes visuospatiales en entraînement dans la modalité lecture : deux situations non signalées

Condition	Situation décrite & question associée
Training (NS)	La fenêtre ouverte est à gauche de la table ronde
	L'armoire basse est à droite de la chaise bleue
	La fenêtre ouverte est au-dessus de la chaise bleue
	Si vous êtes à la place de la table et que vous regardez la chaise, l'armoire est à votre ? (gauche)
	Le vélo rouge est à gauche de la maison blanche
	L'arbre vert est à droite du camion bleu
	La maison blanche est au-dessous de l'arbre vert
	Si vous êtes à la place de l'arbre et que vous regardez le vélo, le camion est à votre ? (droite)

Tableau A.4

Scènes visuospatiales en entraînement dans la modalité écoute : deux situations non signalées

Condition	Situation décrite & question associée
Training (NS)	Le carré bleu est à gauche du rond jaune
	Le triangle vert est à droite de l'étoile rouge
	Le rond jaune est au-dessus du triangle vert
	Si vous êtes à la place du rond et que vous regardez l'étoile, le triangle est à votre ? (gauche)
	L'oiseau bleu est à droite du nuage gris
	L'arbre vert est à gauche du vélo rouge
	L'arbre vert est au-dessous du nuage gris
	Si vous êtes à la place du vélo et que vous regardez le nuage, l'oiseau est à votre ? (droite)

B.Prétest de la signalisation prosodique

Afin de transcrire à l'oral la signalisation de certains mots écrits dans une phrase, comme dans :

Le chat est **à gauche** de la souris

nous avons étudié une forme possible de signalisation prosodique que ce prétest visait à valider en proposant l'écoute de phrases comportant aucun, un, deux ou trois mots signalés par cette emphase. Il a déjà été montré que lorsque la signalisation textuelle était restituée à l'oral, elle permettait la construction d'une représentation adaptée [Lemarié et al. \(2006\)](#), ceci que la restitution soit discursive ou prosodique. C'est donc bien la présence du signal, quelle que soit sa réalisation qui semble pertinente. Une signalisation prosodique peut être un équivalent efficace du signal textuel. L'emphase consiste en la mise en relief d'un ou plusieurs éléments d'un énoncé oral et peut se traduire dans la prosodie par des modifications de la hauteur mélodique de la voix (plus aiguë), de la durée des phonèmes (on allonge le mot), de l'intensité (on augmente le volume) ou encore de l'attaque du mot emphatique ([Herment, 2010](#)). Les participants devaient alors associer cette phrase écoutée à une réalisation textuelle comportant aucun, un, deux ou trois mots signalés en gras. Réalisé sous Qualtrics, ce questionnaire a été proposé sur deux groupes Facebook professionnels (orthophonie et orthoptie) et a mobilisé 54 participants. Avec 19 questionnaires incomplets, 35 observations valides ont pu être extraites (âge moyen 35 ans, 29 femmes). L'ensemble de l'analyse a été réalisé avec le logiciel R (version 3.4.0, [R Core Team, 2017](#)).

Synthèse vocale et signalisation

Afin de contrôler la signalisation prosodique, celle-ci a été réalisée par voix de synthèse avec un API (*Application Programming Interface*) de synthèse vocale mis à disposition par Microsoft : le SAPI5 (*Speech API version 5*) permet l'intégration d'une synthèse vocale dans les applications Windows, selon un principe similaire à l'extension SSML (*Speech Synthesis Markup Language*) de la syntaxe XML qui met à disposition la synthèse vocale dans une application Web.

La spécification SAPI 5.4 intègre quelques balises permettant de modifier la prosodie. Nous avons utilisé la balise <emph> qui induit une emphase sur la partie du texte qu'elle encadre, comme :

Le chat est <emph>à gauche</emph>de la souris

La synthèse vocale prononcera alors avec insistance le connecteur spatial. La voix par défaut implémentée sous Windows 10 (Hortense Desktop) nous a semblé donner de bons résultats. La comparaison des ondes sonores générées par la synthèse vocale pour la phrase « Le livre rouge est au-dessus du cahier quadrillé » met en évidence la modification de trois paramètres lors de l'emphase du connecteur spatial « au-dessus » : la durée des phonèmes s'allonge, l'intensité augmente ainsi que, dans une certaine mesure, la hauteur (plus grave ici, voir spectre des fréquences [Figure B.2](#)).

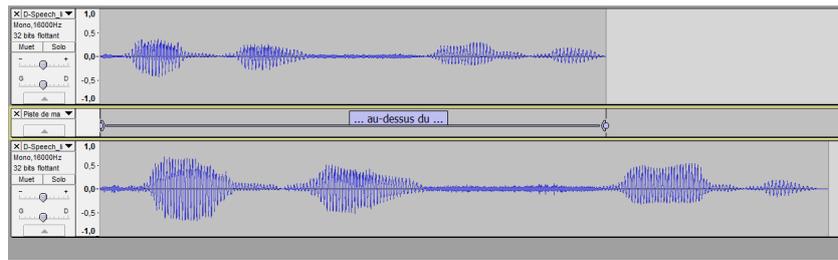
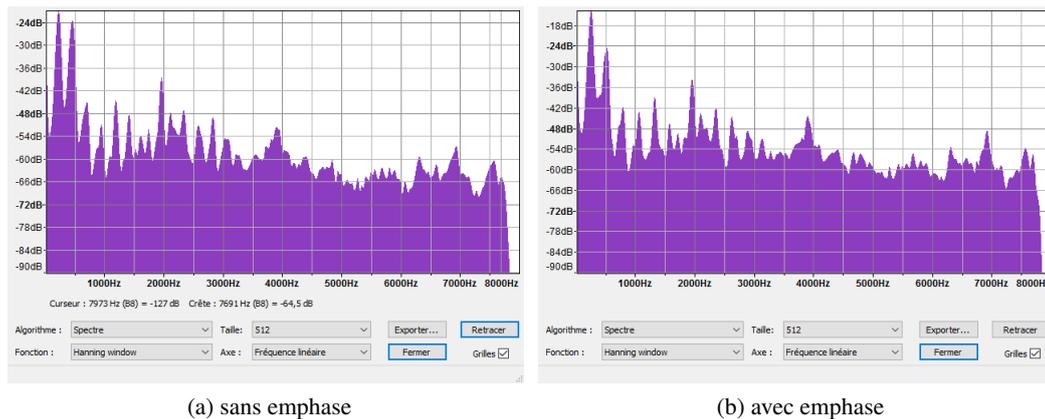


FIGURE B.1. – Comparaison des ondes sonores d'un extrait sans emphase (en haut) avec le même extrait avec emphase (en bas) (tracé par Audacity v2.1.0)



(a) sans emphase

(b) avec emphase

FIGURE B.2. – Analyse des fréquences des extraits sonores avec et sans emphase : augmentation de l'intensité dans les basses fréquences (~250 Hz)

Méthodologie

Nous avons testé quatre conditions de signalisation :

S0 : Aucun signal : « Le livre rouge est au-dessus du cahier quadrillé »

S1 : Un signal : « Le livre rouge est **au-dessus** du cahier quadrillé »

S2 : Deux signaux : « Le livre **rouge** est au-dessus du cahier **quadrillé** »

S3 : Trois signaux : « Le livre **rouge** est **au-dessus** du cahier **quadrillé** »

Chaque condition a été réalisée en synthèse vocale, l'emphase prosodique étant portée par l'équivalent signalé textuellement (voir supra, §B). Afin de valider l'association de la signalisation textuelle et de cette emphase prosodique, deux tâches successives ont été proposées :

1. Une tâche d'association croisée : les participants devaient associer les signalisations prosodiques S0, S1 et S2 d'une même phrase à leurs équivalents textuels. Cette tâche a été proposée trois fois. Présentée en premier, elle faisait office d'entraînement : en autorisant l'écoute comparative des fichiers audio, elle permettait au participant de bien comprendre l'association demandée (voir Figure B.3a).
2. Une tâche d'association simple : les participants devaient associer une seule signalisation prosodique, dans l'une des conditions S0, S1, S2 et S3, à leurs équivalents textuels (voir Figure B.3b). L'écoute comparative de la prosodie n'était dès lors plus possible. Trois essais par conditions ont été proposés, soit 12 phrases audio.

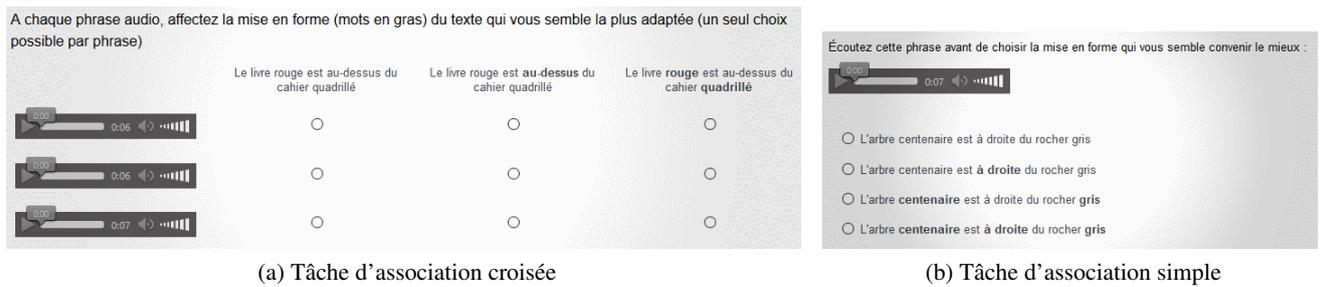


FIGURE B.3. – Tâches d’associations d’emphase prosodique et de signalisation textuelle

Les participants pouvaient écouter chaque fichier audio autant de fois qu’ils le souhaitaient, avant de valider leur réponse. Au sein de chaque tâche, les conditions de signalisation étaient aléatoirement contrebalancées et nous attendions une concordance entre l’emphase prosodique et la signalisation textuelle.

Résultats

Les participants ont mis en moyenne 7 minutes ($SD = 4$) à réaliser l’ensemble du questionnaire. Dans la tâche d’association croisée, il n’a pas été possible d’individualiser le nombre d’écoutes de chaque fichier audio. Dans la tâche d’association simple, il y a eu en moyenne 2 écoutes par fichiers audio, avec un maximum allant jusqu’à 5 pour les deux premières conditions de signalisation (S0 et S1), alors que ce maximum descend entre 3 et 4 pour les deux dernières conditions de signalisation (S2 et S3). Une écoute comparative n’étant pas possible dans cette tâche, il peut sembler plus facile de valider la présence d’une emphase prosodique plutôt que son absence. Le doute inclinant à vérifier, ceci peut expliquer le comportement de nos participants.

Tableau B.1

Pourcentages de bonnes réponses par condition

		(a) Association croisée			(b) Association simple			
		Signaux textuels			Signaux textuels			
		S0	S1	S2	S0	S1	S2	S3
Signaux audio	S0	98.1	1.9	0.0	88.6	2.9	5.7	2.9
	S1	1.0	97.1	1.9	0.0	96.2	1.0	2.9
	S2	1.0	1.0	98.1	1.9	0.0	93.3	4.8
	S3				0.0	2.9	11.4	85.7

Les résultats de la première comme de la seconde tâche sont en accord avec nos prévisions (voir tableau B.1), confirmant une correspondance quasi parfaite entre l’emphase prosodique et la signalisation textuelle (*tests de Fisher*, tous $p < .001$). L’analyse des résidus de Pearson confirme la correspondance des signalisations prosodiques et textuelles.

C.Prétest du matériel d'imagerie visuelle

Ce prétest à visée exploratoire tendait à évaluer les possibilités et les contraintes du matériel expérimental. L'analyse a été réalisée avec le logiciel R (version 3.4.0, [R Core Team, 2017](#)).

Participants

Le prétest a été réalisé sous Qualtrics et proposé sur deux groupes Facebook professionnels (orthophonie et orthoptie). Il se terminait par le questionnaire d'évaluation des habiletés d'imagerie, le VVIQ. Il a mobilisé 304 participants, dont seulement 96 questionnaires (16 % d'hommes) ont été retenus valides après nettoyage des données (suppression des questionnaires incomplets et de quatre valeurs extrêmes). Le temps total de réalisation est en moyenne de 20 minutes ($SD = 7.1$ minutes).

Matériel et procédure

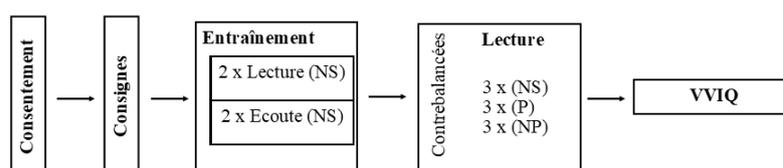


FIGURE C.1. – Procédure expérimentale du prétest

Le prétest consistait en la lecture de trois phrases décrivant une scène visuelle que les participants étaient invités à se représenter avant de répondre à une question de localisation (voir procédure Figure C.1 et exemple Figure C.2). Les trois phrases précisaient les positions relatives de quatre objets concrets dans l'espace d'un carré, les deux premières situant les objets sur l'axe horizontal gauche-droite, la troisième levant l'ambiguïté de position verticale dessus-dessous (voir construction du matériel en Annexe A). Trois conditions de signalisation textuelle ont été mises en œuvre : des phrases sans signalisation (NS), une signalisation pertinente (P) et une signalisation non pertinente (NP) et trois scènes visuelles différentes ont été proposées par condition pour un total de neuf scènes (NS1, NS2, NS3, P1, P2, P3, NP1, NP2 et NP3) dont les présentations ont été aléatoirement contrebalancées. Les trois phrases étaient présentées simultanément et le lecteur décidait quand il était prêt à passer à la tâche de rotation mentale. Celle-ci lui suggérait de prendre un point de vue particulier dans la scène, avant de lui demander d'indiquer la position d'un des objets par rapport à lui-même (Figure C.2, question). Cette question induit une manipulation de la représentation construite afin de répondre au mieux à la localisation demandée et vise ainsi à garantir l'effort nécessaire à une bonne construction de cette représentation mentale. Notre variable d'intérêt est donc le temps de lecture de la description spatiale, considéré comme reflétant l'effort de construction de la représentation.

A la fin du test, le participant devait répondre au questionnaire d'habileté d'imagerie visuelle VVIQ dont l'analyse spécifique est reportée en Annexe D.

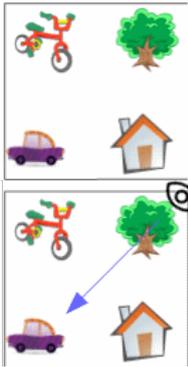
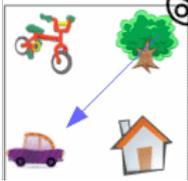
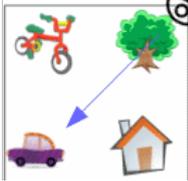
Non signalé	Le vélo rouge est à gauche de l'arbre vert	
	La maison blanche est à droite de la voiture bleue	
	L'arbre vert est au-dessus de la maison blanche	
Signal pertinent	Le vélo rouge est à gauche de l'arbre vert	
	La maison blanche est à droite de la voiture bleue	
	L'arbre vert est au-dessus de la maison blanche	
Signal non pertinent	Le vélo rouge est à gauche de l'arbre vert	
	La maison blanche est à droite de la voiture bleu	
	L'arbre vert est au-dessus de la maison blanche	
Question	Si vous êtes à la place de l'arbre et que vous regardez la voiture, la maison est-elle à votre gauche ou à votre droite ?	

FIGURE C.2. – Exemple de situation visuospatiale dans les trois conditions de signalisation et question de localisation

Résultats

Restreindre cette analyse aux seules réponses correctes à la tâche de rotation mentale ne modifie pas les conclusions obtenues, aussi les résultats ci-dessous sont donnés pour l'ensemble des observations.

La consigne (temps moyen de lecture 63 sec., $SD = 33$ sec.) était illustrée d'un exemple et suivie de deux essais d'entraînement, sans signalisation (Training 1 et 2), toujours dans cet ordre, et dont les résultats sont rapportés dans le tableau C.1. Les participants mettront quelque quatre secondes de moins à répondre lors de la seconde question ($p = .047$) mais les temps de lecture sont similaires. Le taux de réponses correctes est sensiblement différent entre les deux questions (tests binomiaux, $p = .028$) mais est significativement supérieur au hasard pour les deux questions. On pourra donc considérer qu'à l'issue du bloc d'entraînement, les participants ont compris les instructions et sont à même de réaliser correctement l'épreuve.

Tableau C.1

Temps de traitement moyens (écart-types) et taux de réponses correctes par conditions de signalisation (toutes observations)

Condition	Prétest (Lecture)	
	Temps	Score
Training 1	61.00 (36.65)	78 %
Training 2	56.94 (33.16)	65 %
NS	36.89 (17.43)	75 %
P	36.44 (19.28)	80 %
NP	43.51 (24.00)	75 %

Les scores de bonnes réponses (Tableau C.1) sont tous supérieurs au hasard (tests binomiaux, tous $p < .05$) et il n'y a pas de différences entre conditions de signalisation (tests de proportionnalités, $p = .957$). A l'issue de leur traitement de la description textuelle, les participants ont donc construit une représentation suffisamment aboutie pour leur permettre d'assurer une performance continue.

Les distributions des temps de traitement sont largement biaisées vers la droite (skewness moyen = 1.7, voir Figure C.3a) et ne satisfont aucun des tests de normalité (Shapiro-Wilk et Anderson-Darling¹, tous $p < .001$). Nous avons donc utilisé comme test non paramétrique le test de Friedman, stratifié par sujets, pour évaluer les différences de temps de traitement selon les conditions de signalisation. Celui-ci met en évidence une différence significative entre les conditions ($\chi^2(2) = 3.464$, $p = .002$). Les tests post-hoc précisent que le temps de traitement de la condition non pertinente est significativement plus élevé que celui de la signalisation pertinente ($p = .002$) ainsi que significativement plus élevé que celui de l'absence de signalisation ($p = .003$), mais qu'il n'y a pas de différence significative entre les conditions non signalée et pertinente ($p = .975$).

Nous avons évalué la contribution de chacun des paramètres par la méthode du maximum de vraisemblance dans une série de modèles linéaires généralisés à effets mixtes imbriqués afin de sélectionner le modèle le plus adapté en prenant le modèle général suivant :

$$\text{Temps} = \alpha + \beta 1.\text{Signal} + \beta 2.\text{Genre} + \beta 3.VVIQ + \beta 4.\text{Score} + \frac{1}{\text{Sujet}} + \frac{1}{\text{Essai}} + \varepsilon$$

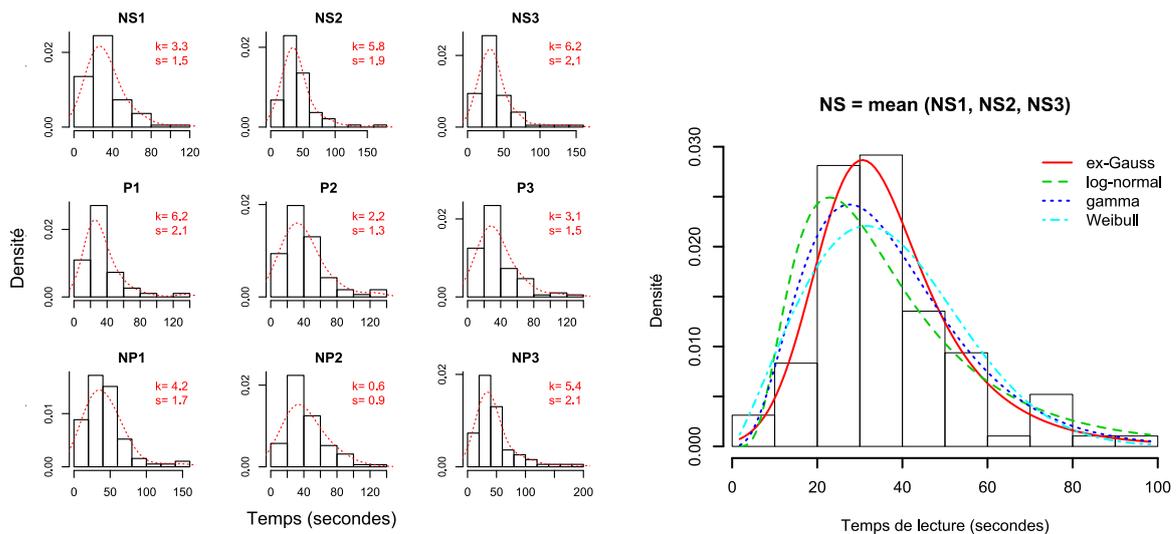
Dans ce modèle, *Temps* est notre variable dépendante (le temps de traitement), α sa valeur de base moyenne quand aucun signal n'est présent dans le texte ; $\beta 1$ représente la pente moyenne de l'effet du signal dans chacune de ses conditions ; $\beta 2$ représente la pente moyenne de l'effet du genre ; $\beta 3$ représente la pente moyenne de l'effet d'habileté visuelle (voir Annexe D, p. XV) ; $\beta 4$ représente la pente moyenne de l'effet du score (réponse juste). Afin d'approcher l'hypothèse d'indépendance des observations, $\frac{1}{\text{Sujet}}$ et $\frac{1}{\text{Essai}}$ correspondent à la prise en compte dans le modèle des effets aléatoires introduits par les sujets eux-mêmes et la répétition des essais (trois essais par condition de signalisation). α est le résidu. Parmi les distributions disponibles, nous avons utilisé une distribution Gamma comme s'ajustant au mieux à nos distributions empiriques. Nous avons évalué la contribution de chacun des paramètres par la méthode du maximum de vraisemblance dans une série de modèles imbriqués afin de sélectionner le modèle le plus adapté.

Le signal est un prédicteur significatif des données ($\chi^2(2) = 23.634$, $p < .001$) et les résultats du modèle sont compatibles et ne modifient pas les résultats du test de Friedman ci-dessus. Ni le genre, ni le score ne sont de bons prédicteurs des données observées. L'habileté d'imagerie visuelle semble participer significativement aux résultats ($\chi^2(2) = 9.466$, $p = .002$) : les participants « bons imaginants » mettent quelques 9 secondes de plus à traiter la description visuospatiale, semblant donc s'être engagé plus profondément dans la construction d'une représentation. Ceci est cohérent avec les résultats de Denis (1982) qui met en évidence que les lecteurs bons imaginants mettent plus de temps à lire un texte descriptif que des lecteurs faibles imaginants.

Analyse de la distribution des temps de traitement

Dans le cas de distributions très biaisées, Whelan (2008) souligne qu'une analyse basée sur les indicateurs de tendance centrale est inadapté, la capacité à détecter des différences pouvant être faible, et préconise alors une analyse de l'ensemble de la distribution, plus informative qu'une normalisation par transformation (voir aussi Balota et Yap, 2011). Une transformation log-normale, souvent préconisée dans le cas de temps de réponse, a toutefois été initialement tentée sans permettre

1. Bien que les résultats soient identiques entre les deux tests, le test d'Anderson-Darling est ici théoriquement mieux adapté, car il accorde plus de poids à la queue de distribution



(a) Tous essais avec densité estimée, kurtosis (k) et skewness (s) (b) Condition sans signalisation avec densités théoriques

FIGURE C.3. – Histogrammes des temps de lecture

la normalisation de l'ensemble des distributions empiriques de cette expérience ce qui interdisait la comparaison des conditions. Une analyse de la distribution dans son ensemble a donc été préférée.

Une approche graphique (Distributions théoriques : Figure C.3b; graphe de Cullen & Frey : Figure C.4) est en faveur d'une composante exponentielle et nous avons évalué l'adéquation de quatre modèles — log-normal, Gamma, Weibull et ex-Gaussienne — par la méthode du maximum de vraisemblance (*maximum likelihood*). Théoriquement, cette méthode impose l'indépendance des mesures par participants (Palmer et al., 2011). Il n'existe néanmoins pas de moyen simple de tester la violation de cette hypothèse que l'on ne peut complètement garantir ici : ainsi que les participants ont pu le rapporter, notre design expérimental favorise fatigue et apprentissage. Nous illustrerons la suite de cette analyse en prenant comme exemple le temps de lecture agrégé² dans la condition NS (texte non signalé), sachant que les conclusions ont été vérifiées et valables pour l'ensemble des temps de réponse qui ont été mesurés.

Le tableau C.2 rapporte l'ajustement des données (*goodness-of-fit*) aux distributions théoriques. Les critères d'ajustement AIC (Aikake's Information Criterion) et BIC (Bayesian Information Criterion), évalués à la lumière des recommandations de Burnham et al. (2002, p. 70), sont en faveur d'un ajustement optimal à une distribution exponentielle gaussienne (*ex-Gaussienne*). Celle-ci se décrit comme la somme (la convolution) de deux variables indépendantes à trois paramètres μ , σ et τ : les deux premiers (μ et σ) correspondent à la moyenne et à l'écart-type de la composante Gaussienne. Le troisième (τ) reflète la moyenne et l'écart-type de la composante exponentielle. Cette distribution est connue pour avoir un bon ajustement aux *temps* de réponse et ses composantes sont parfois attribuées à différents processus mentaux (voir par exemple Palmer et al., 2011; Parris et al., 2013). Ainsi, Luce (1986, chap. 6) décrit l'ex-Gaussienne en terme de processus de décision : la composante gaussienne de la distribution refléterait les processus de décision et la partie exponentielle la réponse motrice. Les paramètres d'ajustement ont été calculés pour l'ensemble des temps de lecture et l'on peut apprécier à

2. Les temps de lecture par condition sont obtenus par agrégation des temps de réponse de chaque participant à chacune des trois scènes qui constituent la condition, c.-à-d. NS = moyenne (NS1, NS2, NS3)

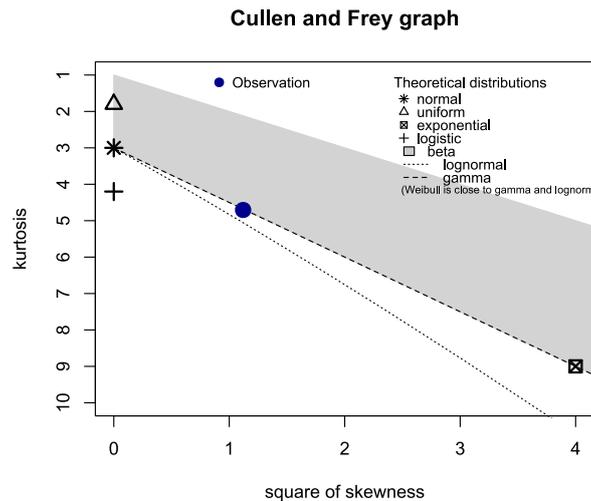


FIGURE C.4. – Pré-analyse graphique de la distribution des temps de lecture

titre d'exemple l'ajustement de l'ex-Gaussienne à la distribution empirique des temps de lecture dans la condition non signalée (Figure C.5).

Tableau C.2

Ajustement des distributions théoriques aux temps de lecture (condition NS)

Critères d'ajustement	Log-normal	Gamma	Weibull	ex-Gaussian
AIC (Aikake's Information Criterion)	842.0297	819.7326	817.6619	809.0992
BIC (Bayesian Information Criterion)	847.1584	824.8613	822.7906	816.7923

Les paramètres d'ajustements ex-Gaussiens se rapportant aux temps de lecture agrégés par condition sont rapportés Tableau C.3 et Figure C.6. Les moyennes arithmétiques sont aussi rappelées ce qui permet de vérifier la propriété d'additivité des moyennes de l'ex-gaussienne : la moyenne arithmétique doit correspondre à la somme de μ et de τ .

Tableau C.3

Moyenne arithmétique classique et paramètres d'ajustements ex-Gaussiens des temps de lecture et taux de réponses correctes par conditions

Condition	Moyenne	Paramètres ex-Gaussiens			Réponses correctes
	arith.	μ	σ	τ	
Training 1	61.00 (36.65)	27.58 (3.70)	14.19 (2.77)	33.43 (4.82)	78 %
Training 2	56.94 (33.16)	22.74 (3.23)	10.31 (2.68)	34.22 (4.64)	65 %
NS	36.89 (17.43)	22.45 (2.06)	9.29 (1.42)	14.54 (2.36)	75 %
P	36.44 (19.28)	18.17 (1.82)	6.91 (1.37)	18.27 (2.51)	80 %
NP	43.51 (24.00)	23.28 (2.18)	9.64 (1.51)	20.23 (2.83)	75 %

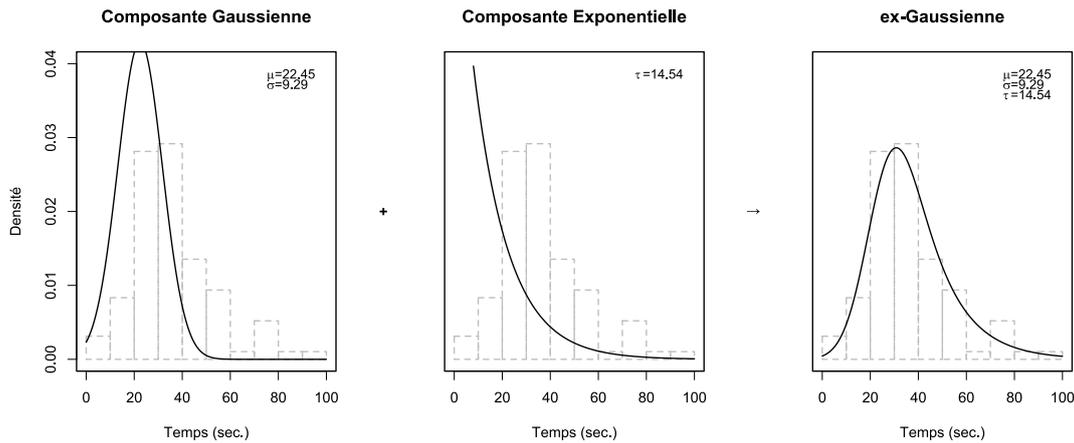


FIGURE C.5. – Composantes de la distribution ex-Gaussienne ajustée aux temps de lecture (condition NS, distribution empirique en grisé)

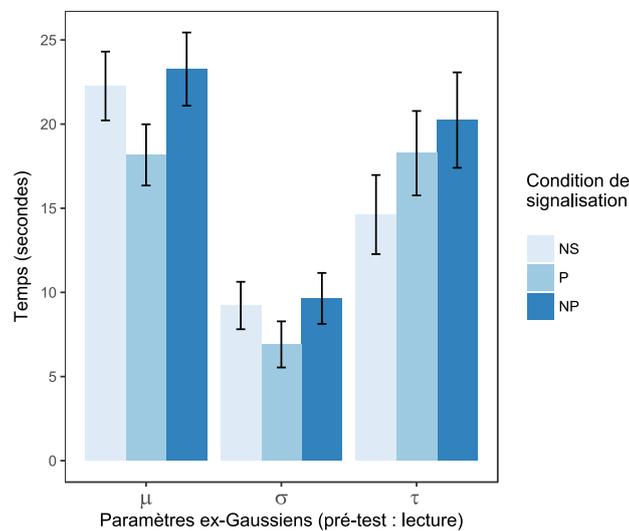


FIGURE C.6. – Valeurs des paramètres ex-Gaussiens pour les temps de lecture par condition (en secondes avec écart-types)

Une régression linéaire a été évaluée pour les deux composantes afin de prédire le temps de traitement en fonction de la signalisation :

Composante Gaussienne : $F(2, 190) = 181.75, p < .001, \eta^2G = .558$, les tests post-hoc (t-test, correction Bonferroni) indiquent des différences toutes significatives : $NP > NS > P$ ($p < .001$)

Composante Exponentielle : $F(2, 190) = 123.02, p < .001, \eta^2G = .493$, les tests post-hoc (t-test, correction Bonferroni) indiquent des différences toutes significatives : $NP > P > NS$ ($p < .001$)

Discussion

Ce prétest visait à valider le matériel et la procédure envisagée. Il met en évidence une mortalité expérimentale importante (près de 66 %), probablement parce que la tâche est couteuse, notamment dans le maintien de la représentation construite. Cette mortalité expérimentale peut induire un biais de sélection dans cet échantillon, seuls les participants motivés ou performants ayant poursuivi l'expérience jusqu'au bout. L'analyse des scores de réussite à la tâche de rotation mentale met en évidence un engagement suffisant dans le traitement de la description verbale. L'analyse par test paramétrique ou régression linéaire semble en faveur d'un effet sémantique de la condition de signalisation. Si ce résultat n'est pas incompatible avec nos hypothèses, ce n'est pas ce lui que nous attendons dans la modalité lecture. Le biais de sélection nous invite à la prudence dans l'analyse prématurée du prétest.

L'interprétation de l'ajustement des données par une distribution ex-Gaussienne n'est pas sans difficulté et si certains auteurs mettent en garde contre toute interprétation en terme de processus psychologiques (voir par exemple [Palmer et al., 2011](#); [Parris et al., 2013](#)), on peut toutefois souligner la stabilité des paramètres ex-gaussiens pour un individu, avec des corrélations de .717, .509 et .872 pour μ , σ et τ respectivement, dans une tâche de décision lexicale (voir [Balota et Yap, 2011](#)). [Schmiedek et al. \(2007\)](#) ont étudié les relations entre les capacités de mémoire de travail et les paramètres ex-gaussiens. Des deux composantes, la composante exponentielle (paramètre τ) est celle qui corrèle le plus avec la mémoire de travail. Ceci a été répliqué chez l'adulte ([Tse et al., 2010](#)) avec des tâches différentes. [Balota et Yap \(2011\)](#) soulignent par ailleurs que ce lien robuste relevé entre τ et la mémoire de travail est cohérent avec la loi de la pire performance ([Coyle, 2003](#)). Dans la tâche proposée, c'est bien la mémoire de travail qui est mobilisée dans la construction de la représentation. Une interprétation possible alors des résultats obtenus pourraient être similaire à celle proposée par [Rohrer et Wixted \(1994\)](#) :

- la composante gaussienne renverrait à une phase courte d'initiation distribuée normalement : le lecteur prend connaissance des éléments, le traitement pourrait être essentiellement verbal ; on remarque que le pattern de données obtenu est en faveur de notre hypothèse sémantique.
- la composante exponentielle pourrait alors renvoyer à une phase plus longue d'élaboration de la représentation et dont le traitement pourrait être essentiellement visuospatial ; on remarque que le pattern de données obtenu est en faveur de notre hypothèse visuelle.

Cette analyse par composantes de nos distributions empiriques reste très largement discutable (pas de consensus dans la littérature malgré un intérêt croissant ces dernières années) et nous ne la retiendrons pas ici, car de très nombreuses critiques peuvent être formulées à l'égard des interprétations que nous avons esquissées. Néanmoins, il nous semble qu'elle ne doit pas être complètement écartée, car elle peut offrir les possibilités de nouveaux paradigmes expérimentaux.

D.Vividness of Visual Imagery Questionnaire (VVIQ)

Il existe une certaine variabilité interindividuelle dans les capacités d'un lecteur à construire et/ou manipuler une image visuelle et, s'agissant de juger de la qualité de la construction d'une représentation d'imagerie visuelle, il nous a semblé pertinent d'évaluer l'habileté des lecteurs à mobiliser une image mentale claire et nette. David Marks propose une définition de cette netteté (*vividness*) : « Par "netteté" nous entendons une combinaison de clarté et de vivacité. Plus une image est nette, au mieux elle approxime un percept réel. » (Marks, 1972, p.83, notre traduction).

En tant qu'expérience phénoménologique, l'imagerie visuelle est une expérience subjective dont il n'existe pas de comportement non verbal observable et elle ne peut donc être étudiée que grâce aux rapports verbaux. Un certain nombre de questionnaires ont été établis, dont le plus utilisé est probablement le Vividness of Visual Imagery Questionnaire (VVIQ, Marks, 1973; Richardson, 1999). Le VVIQ a été élaboré pour distinguer les sujets susceptibles de générer des images visuelles à partir de descriptions verbales. Il est postulé que des scores élevés à ce questionnaire prédisaient une tendance plus élevée à construire des images visuelles à partir du texte. De fait, l'étude de Chaguiboff et Denis (1981) invitait à une lecture d'un texte narratif de 2000 mots à propos d'un voyage automobile suivie d'une épreuve de reconnaissance de mots dans deux conditions : verbale ou figurative. Les sujets à hauts scores au VVIQ obtiennent les scores de reconnaissance verbale les plus élevés, mais aussi les réponses les plus rapides. La reconnaissance figurative n'affecte que les sujets à faible score VVIQ avec des temps de latence élevés. Les auteurs proposent que les sujets à faible pouvoir d'imagerie visuelle (faible score VVIQ par rapport à la moyenne du groupe) ont une représentation du texte relativement pauvre en imagerie. Il semble que l'on puisse considérer que les lecteurs à haut pouvoir d'imagerie s'engagent plus profondément dans l'utilisation et le maintien des images construites. C'est ce que suggère Denis (1982) qui met en évidence que les lecteurs bons imaginants mettent plus de temps à lire un texte descriptif que des lecteurs faibles imaginants, mais rappellent mieux le contenu textuel. Dans le cadre qui nous intéresse ici, on retiendra que le VVIQ est un bon prédicteur des performances de mémoire spatiale des objets (voir Gur et Hilgard, 1975; Marks, 1973 in Finke, 1980).

Ce questionnaire a été proposé en fin de procédure lors du prétest et du test. Les analyses en clusters des réponses ont été réalisées avec le logiciel R R Core Team (version 3.4.0, 2017).

Description

Le VVIQ est une évaluation par questionnaire de la netteté (*vividness*) d'une image mentale visuelle générée à la lecture d'une description verbale. Il a été développé initialement en 1973 par Marks (1973) qui en a proposé en 1995 une deuxième version avec deux fois plus d'items que dans la version initiale et qui inverse l'échelle de cotation afin qu'un score élevé au VVIQ reflète un haut potentiel d'imagerie (Marks, 1995). C'est aujourd'hui un questionnaire largement utilisé (Richardson, 1999) et dont l'évaluation de fiabilité et de validité de construit est considérée comme élevée et similaire pour les deux versions (Campos et Pérez-Fabello, 2009; Campos, 2011).

Il est constitué de seize questions (voir Tableau D.2), regroupées en quatre groupes de quatre items. Le participant est invité à évaluer l'image mentale qu'il forme à la lecture de chaque item. Chaque item du questionnaire est coté sur une échelle de Likert à 5 choix croissants (1 : « aucune image évoquée », 5 : « image parfaitement claire, aussi nette et précise qu'une véritable perception » ; voir Tableau D.1). Le score global (somme des scores de chacun des 16 items) s'étend donc théoriquement d'un minimum de 16 à un maximum de 80.

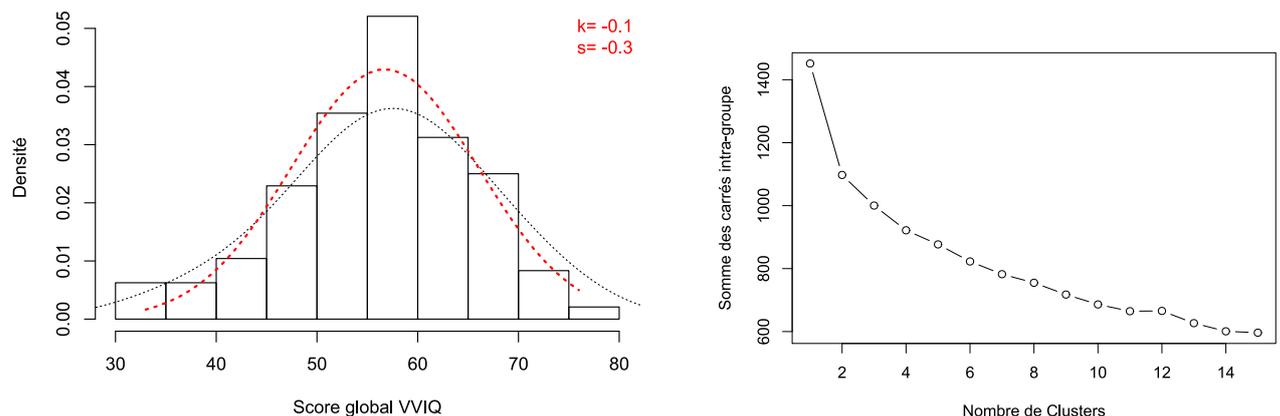
Tableau D.1

VVIQ : échelle d'auto-évaluation pour chacun des items

Label	
1	Aucune image évoquée
2	Image vague et imprécise
3	Image moyennement claire et nette
4	Image relativement claire, presque aussi nette et précise qu'une perception
5	Image parfaitement claire, aussi nette et précise qu'une véritable perception

Analyse en cluster pour le prétest

Cette analyse en cluster visait à déterminer l'existence de groupes spécifiques. Les participants obtiennent un score moyen de 56.78 ($SD = 9.65$), avec un score minimum de 33 et maximum de 76 et la distribution du score total satisfait l'hypothèse de normalité (Figure D.1 ; test de normalité de Shapiro-Wilk, $p=.276$).



(a) Distribution empirique des scores avec densité empirique (pointillé noir) et densité théorique normale (pointillé rouge)

(b) Scree plot

FIGURE D.1. – Prétest :distribution empirique des scores avec densité empirique (pointillé noir) et densité théorique normale (pointillé rouge)

Le dendrogramme D.3a ainsi que le scree plot D.1b sont en faveur de deux clusters ce que tendent à confirmer les résultats donnés par le package R 'NbClust'. L'analyse en deux clusters détermine alors deux groupes d'effectifs 58 et 38 participants dont les scores globaux au questionnaire sont

Tableau D.2

VVIQ : Groupes et Items

Groupe	Items
	Pensez à un paysage de campagne, avec des arbres, des montagnes et un lac. Examinez ensuite attentivement l'image visuelle qui vous vient à l'esprit quand :
	... vous examinez les contours du paysage ... vous examinez la couleur et la forme des arbres ... vous examinez la couleur et la forme du lac ... vous imaginez qu'une rafale de vent souffle sur les arbres et provoque des vagues sur le lac
	Imaginez maintenant le soleil qui se lève. Examinez ensuite attentivement l'image visuelle qui vous vient à l'esprit quand :
	... le soleil se lève à l'horizon dans un ciel brumeux ... le ciel s'éclaircit et le soleil est entouré de bleu ... il fait un ciel d'orage, les nuages sont illuminés par les éclairs ... un arc-en-ciel apparaît
	Pensez à un membre de votre famille ou à une personne amie que vous rencontrez souvent. Examinez ensuite attentivement l'image visuelle qui vous vient à l'esprit quand :
	... vous examinez le contour exact du visage, de la tête, des épaules et du corps ... vous examinez les positions caractéristiques de la tête et le maintien du corps ... vous examinez l'allure et l'amplitude précise de la démarche ... vous examinez les différentes couleurs d'une tenue vestimentaire habituelle
	Pensez maintenant à la devanture d'un petit magasin que vous connaissez bien. Examinez ensuite attentivement l'image visuelle qui vous vient à l'esprit quand :
	... vous examinez l'apparence générale du magasin vu de l'autre côté de la rue ... vous examinez la présentation de la vitrine, avec la couleur, la forme, les détails des articles exposés ... vous imaginez que vous êtes près de l'entrée. Examinez la couleur, la forme et les détails de la porte ... vous imaginez que vous entrez dans le magasin et que vous allez jusqu'au comptoir. Un vendeur vous sert. L'argent change de mains

respectivement de 62.44 et 47.38, soit environ d'un écart-type autour de la moyenne de l'échantillon.

Analyse en cluster pour le test

La distribution du score total satisfait l'hypothèse de normalité (Shapiro-Wilk, $p=.580$). L'analyse en cluster des réponses au questionnaire VVIQ lors du test est identique à celle du prétest avec une analyse en faveur de deux clusters très bien systématisés sur la Figure D.3b. La répartition des clusters donne deux groupes de 18 participants dont les scores globaux au questionnaire sont respectivement de 65.44 et 49.83 (Tableau D.3 et Figure D.2b), à près d'un écart-type autour de la moyenne de l'échantillon (Tableau D.3 et Figure D.2a).

Résultats

On trouvera dans le tableau D.3 la répartition des scores en deux clusters suite aux analyses ci-dessus. Nous qualifierons les deux groupes de participants correspondants de « bons » et « mauvais »

imaginants, relativement distribués symétriquement à ± 1 écart-type de la moyenne de l'échantillon.

Tableau D.3

Scores moyens et écarts à la moyenne empirique (rappel : score $\in [16, 80]$)

	Moyenne théorique	Moyenne empirique (SD)	Bons imaginants	Mauvais imaginants
prétest	48	56.78 (9.65)	62.45 (+0.6 SD)	47.38 (-1.0 SD)
test	48	57.64 (9.44)	65.44 (+0.8 SD)	49.83 (-0.8 SD)

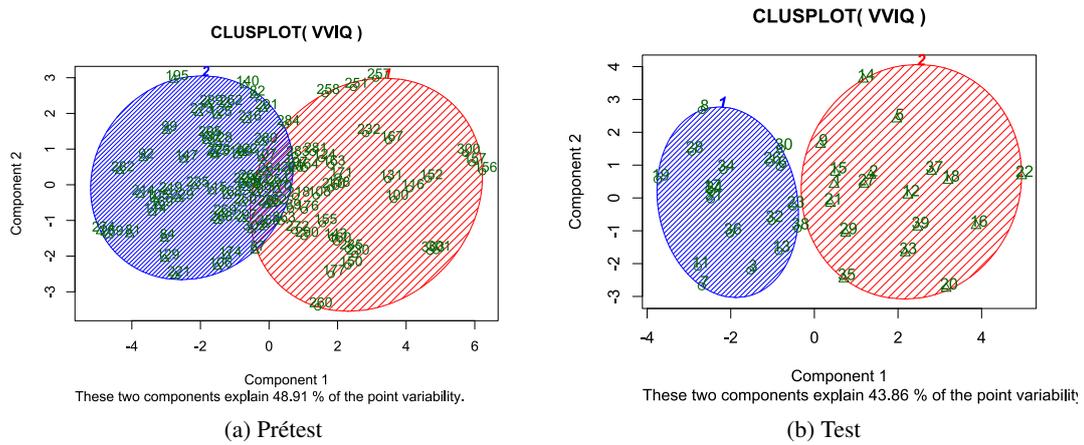
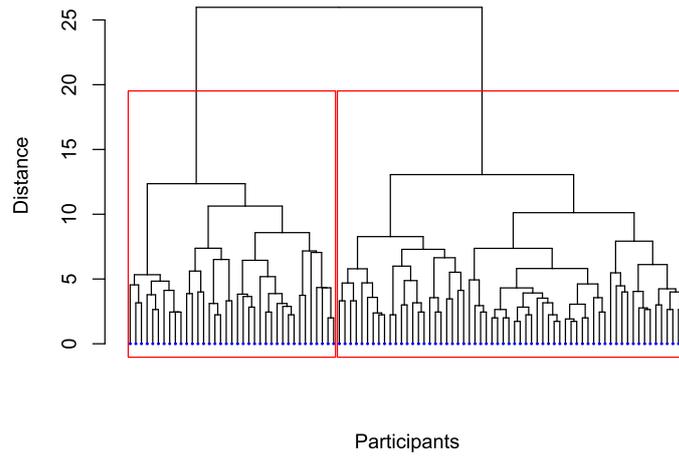
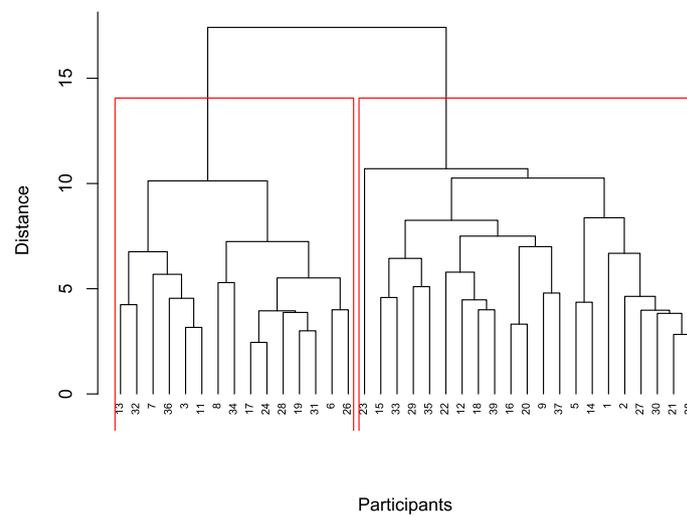


FIGURE D.2. – Clusters



(a) Prétest



(b) Test

FIGURE D.3. – Dendrogrammes