

Dans ce troisième article inspiré de sa thèse *Syntaxe, vision parafovéale et processus de lecture. Contribution du modèle structural à la pédagogie* (voir A.L. n°90, juin 05, pp.66-81 et 91, sept. 05, pp.31-41), Denis Foucambert montre comment et combien les variations syntaxiques et lexicales perçues en vision périphérique participent à la lecture.

L'importance de la vision périphérique pour le modèle structural.

Position du problème

Les expérimentations présentées jusqu'ici permettent de mieux comprendre les relations entre l'oubli de lettres et le rôle organisateur des mots dans la phrase. Néanmoins, d'autres facteurs semblent susceptibles d'influer sur cette relation. Nous rappelions dans les numéros précédents les hypothèses de Koriat où il suggère une influence possible de la vision périphérique sur l'effet d'oubli de lettres (Koriat & Greenberg, 1996).

Ce que nous allons tenter de découvrir, c'est si le lecteur peut, à l'aide d'informations périphériques présentées très rapidement, établir des cadres syntaxiques et jusqu'à quelle distance cette construction est possible. Autrement dit, le lecteur est-il sensible à des variations syntaxiques et/ou lexicales dans la zone parafovéale et jusqu'à quelle distance ?

L'objectif de cette expérimentation est donc de tester l'hypothèse nulle qui stipule que les informations parafovéales ne permettent pas de prédire la structure syntaxique.

Méthode

Cette épreuve se passe sur un ordinateur dont la résolution d'écran est fixée en 800x600 pixels. Tous les affichages se font en Times New Roman de taille 13. Un expérimentateur se tient perpendiculairement au sujet. Un temps d'aprentissage est prévu.

Déroulement de l'épreuve

- Un point apparaît sur l'écran, qu'il est demandé au sujet de fixer.
- Après 500 ms, une phrase apparaît et reste affichée pendant 300 ms.
- Le sujet dit à l'expérimentateur le mot qui est apparu à la place du point de fixation. Si ce n'est pas le bon mot, l'expérimentateur passe à l'item suivant.
- Si le sujet a correctement donné le mot, trois phrases s'affichent.
- Le sujet doit sélectionner la phrase qu'il lui semble avoir vue initialement.
- On passe à l'item suivant.

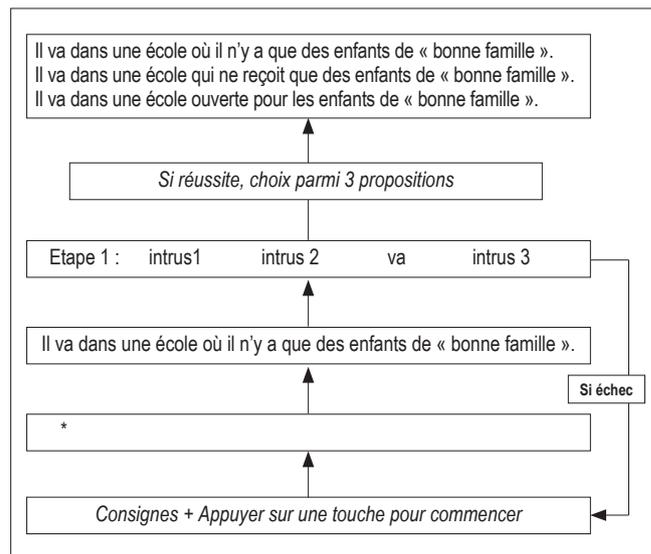


Figure 1 : déroulement de l'épreuve

Construction des triplets de phrases

Le plan expérimental comprend trois sources de variations :

■ **La fréquence du mot fixé.** Il peut être soit rare, soit fréquent. On peut faire l'hypothèse que l'identification d'un mot rare (comme botanique) influe sur la suite de la tâche en consommant davantage de « ressources cognitives » que la reconnaissance d'un mot fréquent (comme grand).

■ **Le rôle du mot** par où commence la variation entre les trois phrases. Deux catégories sont possibles : soit le rôle du mot qui initie la variation des phrases est sémantique (adjectif, verbe, nom), soit il est syntaxique (pronom relatif, conjonctions,...). Les deux autres phrases varient à partir de ce mot en le remplaçant pour l'une, par un mot à rôle sémantique et pour l'autre par un mot à rôle syntaxique.

■ **La distance de la variation.** Les phrases se différencient au bout de 9, 15, 21 ou 27 signes entre le milieu du mot à fixer et le début du mot qui introduit la variation. Pour des questions de faisabilité, une tolérance d'un signe est acceptée dans ces distances. Ce qui est recherché ici, c'est l'éventualité d'une distance maximale à partir de laquelle l'effet ne serait plus perceptible. L'unité de mesure que nous avons choisie est le nombre de signes plutôt que l'angle qui sépare le point de fixation du mot introduisant la variation. L'emploi d'une police proportionnelle où la largeur des lettres varie nous impose ce choix, dans la mesure où il semble impossible de fabriquer des triplets de phrases possédant un angle de variation identique. En outre, nous ne contrôlons pas la position de la tête du lecteur par rapport à l'écran. Cependant, à titre illustratif, nous avons calculé l'angle moyen pour chacune des distances de variation, en faisant l'hypothèse que les yeux du lecteur se trouvent à 40 cm de l'écran. Le tableau 1 les présente, ainsi que le nombre moyen de mots séparant le point de fixation du mot introduisant la variation.

Signes	Angle	Mots
9 signes	3°59'	2,33
15 signes	5°18'	3,42
21 signes	7°1'	4,87
27 signes	8°4'	4,75

Tableau 1 : Equivalences entre le nombre de signes, l'angle moyen d'écartement et le nombre moyen de mots.

La mise en relation des différents angles rapportés dans ce tableau avec ce que nous rapportions dans un article précédemment (AL n°91, page 31, figure 2) montre qu'à partir de la distance de 15 signes la qualité des identifications est inférieure à 50%.

■ En conséquence, ce plan aboutit à seize combinaisons différentes ; nous produirons deux triplets par combinaison soit un total de trente-deux triplets. Leur récapitulation exhaustive serait trop longue, mais nous en donnerons ici deux exemples :

Exemple n°1 : il s'agit ici d'un triplet avec un mot d'appel fréquent (*grand*), avec une distance de 9 signes entre le milieu du mot fixé et le début des mots qui différencient les phrases (*donne, doit, dont*). La phrase à retrouver, d'abord affichée à l'écran en flash, a, comme mot débutant la variation, un mot à rôle sémantique (*donne*). La lettre en gras montre la place du point de fixation. La variation dans les deux autres phrases commence pour l'une par un mot à rôle sémantique (*doit*) et pour l'autre par un mot à rôle syntaxique (*dont*).

Le grand garçon **donne** la bourse pleine d'o...
 Le grand garçon **doit** la bourse pleine d'or...
 Le grand garçon **dont** la bourse pleine d'or...
 9 signes

Exemple n°2 : il s'agit ici d'un triplet avec un mot d'appel fréquent (*monsieur*), avec une distance de 21 signes entre le mot d'appel et le mot qui différencie les phrases (*devant, devenu, avec*), et dont la phrase affichée à l'écran en flash contient comme mot débutant la variation un mot à rôle syntaxique (*devant*). La lettre en gras montre la place du point d'appel ; *devenu* a un rôle sémantique, *avec* un rôle syntaxique.

C'est le **monsieur** en costume qui est devant le premier minist...
 C'est le **monsieur** en costume qui est devenu le premier minist...
 C'est le **monsieur** en costume qui est avec le premier minist...
 21 signes

Les triplets sont construits pour avoir la plus grande similarité visuelle. On notera que, dans le cas où, dans la phrase présentée en flash, le mot différenciant a un rôle syntaxique, le mot différenciant à rôle sémantique d'une des deux autres phrases a une apparence visuelle très proche (*devant/devenu*), alors que l'autre mot à rôle syntaxique est fort différent (*devant/avec*). En revanche, quand, dans la phrase présentée en flash, le mot différenciant a un rôle sémantique, les trois mots différenciant du triplet sont visuellement proches.

Enfin, rappelons qu'on oblige le sujet à fixer le mot d'appel par la présence d'un point sur l'écran de l'ordinateur. C'est sa reconnaissance qui valide le triplet. Grâce à cette contrainte, on s'assure que le mot débutant la variation des phrases est bien dans la zone parafovéale.

Population

La population est celle décrite plus précisément dans les Actes de Lecture n°90 (page 74-75) : soixante-quatre adultes, tous enseignants, donc au moins titulaires du baccalauréat, d'âge moyen de 40,7 ans, avec un maximum de 58 ans et un benjamin de 26 ans. L'âge médian est de 41 ans.

Résultats

Pour un des triplets, l'enregistrement n'a pas fonctionné, ce qui ramène le nombre de triplets que nous allons observer à 31.

	Modalité	Pourcentage de réussites	Écart-type
Rôle du mot	Sémantique	31,10	10,12
	Syntaxique	38,14	7,80
Distance	9 signes	36,36	9,09
	15 signes	29,60	8,74
	21 signes	31,20	11,32
	27 signes	40,45	6,66
Fréquence du mot fixé	Fréquent	35,23	7,99
	Rare	33,73	11,33
Total		34,51	9,61

Tableau 2 - Exercice triplet : Pourcentage de réussite en fonction du rôle du mot, de la distance de la variation et de la fréquence du mot fixé.

Première analyse

Comment sont constituées les phrases cibles quand on les retrouve ensuite dans la liste des triplets ? Pour répondre à cette question, nous menons une analyse de variance, avec le pourcentage de réussite pour chacun des triplets comme variable dépendante. Les variables explicatives sont les valeurs qui correspondent aux trois sources de variation : la fréquence du mot fixé, la distance entre le mot fixé et le mot qui initie la variation entre les phrases et enfin le rôle de ce mot, sémantique ou syntaxique. Nous ajoutons dans le modèle l'ensemble des interactions d'ordre 2.

Le tableau 3 montre les résultats univariés issus de l'Anova.

Effet	Somme des carrés	Degré de liberté	Moy. des carrés	F	p<
Distance	530,95	3	176,98	2,259	0,12
Fréquence	9,09	1	9,09	0,116	0,74
Rôle du mot différenciant	390,37	1	390,37	4,983	0,039*
Distance * Fréquence	46,43	3	15,48	0,198	0,9
Distance * Rôle	268,84	3	89,61	1,144	0,36
Fréquence * Rôle	69,61	1	69,61	0,889	0,36
Erreur	1410,05	18	78,34		

Tableau 3 - Tests Univariés de Significativité du Pourcentage de réussite. La ligne en gras montre la contribution significative.

Il montre que le seul facteur qui entretient un rapport significatif avec la réussite est le rôle du mot¹ initiant la variation.

La figure 2 illustre la différence de réussite à l'épreuve en fonction du rôle du mot introduisant la variation entre les phrases : on voit clairement que quand ce mot a un rôle syntaxique (comme dans l'exemple 2 donné plus haut) la réussite est meilleure que quand son rôle est sémantique (l'exemple 1 donné plus haut).

¹ Nous avons fait passer cette épreuve à 49 élèves de sixième et 48 élèves de troisième. Pour ne pas perturber la présentation que nous faisons, qui se focalise sur une population homogène d'adultes, nous ne les avons pas intégrés dans le calcul. Cependant, l'analyse donne des résultats proches de ceux des adultes :

- Pour les élèves de 6^{ème}, le rôle du mot initiant la variation a un effet quasiment significatif ($F(1, 25) = 4,0424, p < 0,055$) avec 29% de réussite quand ce mot a un rôle sémantique et 38% de réussite quand il a un rôle syntaxique.

- Pour les élèves de 3^{ème}, le rôle du mot initiant la variation a un effet quasiment significatif ($F(1, 25) = 3,5123, p < 0,07$) avec 30% de réussite quand ce mot a un rôle sémantique et 37% de réussite quand il a un rôle syntaxique.

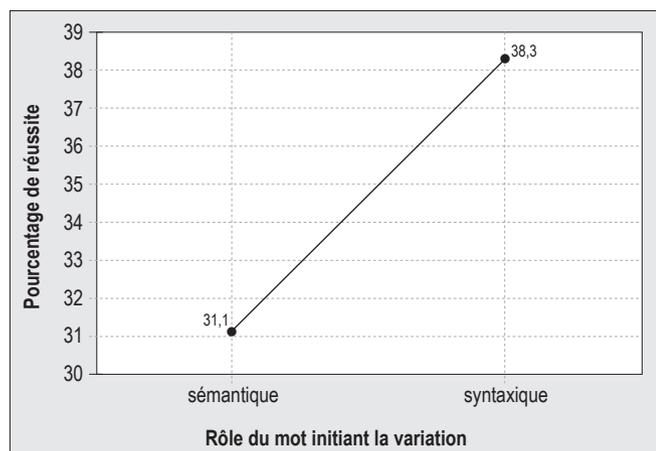


Figure 2 - Pourcentages (des moindres carrés) de réussite en fonction du rôle du mot introduisant la différence à l'intérieur des triplets.

On note également que la distance n'est pas significative ($p < 0,12$), ce qui est corroboré par les tests Post Hoc qui ne montrent pas d'oppositions significatives entre les différentes modalités². On remarquera quand même que les distances importantes ne semblent pas poser plus de problèmes que les distances courtes, puisque c'est à la distance de 27 signes que les sujets obtiennent les meilleurs résultats, la distance de 9 signes venant après.

Les échecs

Quand le sujet n'a pas correctement répondu, c'est-à-dire quand il n'a pas retrouvé la bonne phrase, il a néanmoins dû faire un choix. Si le mot différenciant a un rôle sémantique, les trois mots seront visuellement proches (« *sourd, soumis, sous* », « *salie, saline, selon* »,...). Quand le mot différenciant a un rôle syntaxique, les autres possibilités de réponses possèdent comme mot différenciant soit un mot à rôle sémantique qui est visuellement proche, soit un mot à rôle syntaxique qui est visuellement éloigné (« *par, pur, en* » ; « *mais, mains, et* »,...).

Dans la phrase à retrouver, le mot portant la variation	Dans les deux phrases intruses
a un rôle syntaxique a un rôle sémantique	Le mot à rôle syntaxique ne lui ressemble pas Le mot à rôle sémantique lui ressemble
	Le mot à rôle syntaxique lui ressemble Le mot à rôle sémantique lui ressemble

D'une certaine manière, on tend un piège au lecteur : une des phrases intruses est visuellement plus proche de la phrase projetée puisque les deux mots différenciants se ressemblent (dans ce cas, la structure n'est pas respectée) alors que l'autre lui ressemble moins puisque les mots différenciants sont visuellement très différents (dans ce cas, la structure est identique). L'alternative est claire : le lecteur va-t-il faire le choix de l'apparence visuelle contre la structure ou celui de la structure contre l'apparence visuelle ?

Par définition, ce schéma n'a pas à s'appliquer aux phrases dont le mot initiant la variation a un rôle sémantique puisqu'on fait l'hypothèse que c'est la perception et le traitement non conscients des mots fonctionnels qui sont possibles dans la vision parafovéale.

Quand un sujet a échoué à un triplet dont le mot différenciant est fonctionnel, nous garderons comme information la nature de sa réponse : a-t-il choisi la phrase visuellement ou structurellement proche ? Pour chacun de ces triplets, nous connaissons donc le pourcentage de réponses privilégiant l'apparence visuelle ou la structure de la phrase.

Un simple calcul du T de Student nous permet de savoir si les différences de moyennes sont significatives entre ces deux types de réponses. Les variances des deux groupes (Apparence proche vs. Rôle identique) étant homogènes (Test de Lévène : $F(1,28) = 0$; $p = 1$), nous pouvons accepter le résultat de l'analyse du Test t pour des Echantillons Indépendants) présenté dans le tableau 4.

	Moy. Apparence.	Moy. Rôle.	valeur t	df	p	F Levene	DI Levene	P Levene
Apparence Proche vs. Rôle Identique	42,01	57,99	-3,10	28	0,004	0,00	28	1

Tableau 4 - Résultats de la comparaison des catégories de réponse en cas de mauvaise réponse.

² Les deux modalités s'éloignant le plus significativement sont la distance de 15 signes contre celle de 27 signes avec plus d'oubli pour 27. La significativité du test de HSD de Tuckey ne donne qu'un $p < 0,08$.

La figure 3 montre très clairement que les sujets choisissent plus souvent la phrase qui a une structure identique bien que l'apparence visuelle en soit plus éloignée³. Ce résultat montre toute l'importance des informations parafovéales dans la construction du squelette syntaxique de la phrase.

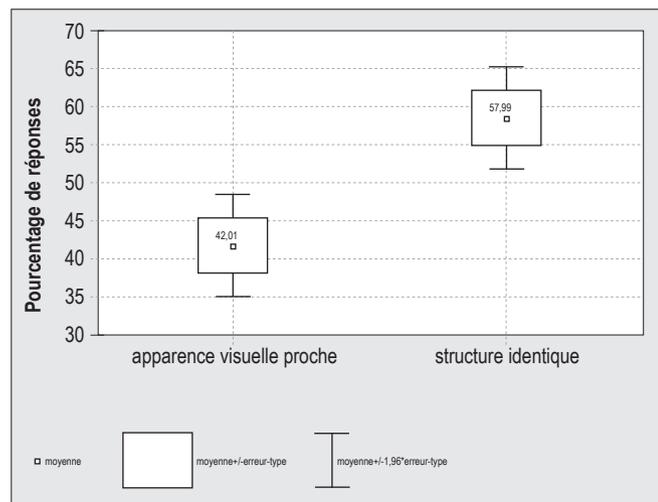


Figure 3 - Pourcentage de réponses aux triplets échoués selon l'apparence visuelle ou la structure identique.

Discussion

L'ensemble des résultats présentés ici permet de rejeter notre hypothèse nulle qui prévoyait que les informations parafovéales ne permettent pas de prédire la structure syntaxique. Au contraire, il semble bien que le lecteur est susceptible de prendre appui sur les informations issues de la vision périphérique pour construire le squelette syntaxique de la phrase.

La particularité de cette expérience, c'est de proposer un protocole où le choix se fait entre trois phrases également plausibles, et où il ne peut s'expliquer que par l'utilisation des informations parafovéales. La reconnaissance par le lecteur de la structure ne peut se faire ni à partir des informations perçues fovéalement, ni à l'aide du contexte linguistique dans lequel s'insère la phrase. La seule information dont il a conscience concerne le mot qu'il devait fixer et cette information ne peut lui servir en rien pour prévoir la suite de la

phrase, puisqu'elle est présente dans les trois phrases proposées. Pour prendre une décision, il faut donc que le lecteur s'appuie sur les informations visuelles périphériques.

■ **Le premier résultat** que nous obtenons montre que les triplets les mieux réussis sont ceux où le sujet a pu repérer, dans la phrase présentée initialement, une structure syntaxique bien particulière. Quand, de manière tangible, les mots fonctionnels organisent la structure de la phrase, les sujets sont mieux capables de la retrouver, comme s'ils avaient été sensibles à sa présence et que finalement, à ce stade, c'est du squelette syntaxique dont ils se souviennent.

■ **Le second résultat**, qui concerne l'analyse des erreurs, est sans doute le plus intéressant. On aurait pu s'attendre à ce que les sujets soient attirés par la proximité d'une apparence visuelle. Il n'en est rien : ce qui prime dans le choix de la réponse, c'est la reconnaissance d'une structure à laquelle on a dû être sensible dans la phrase initialement projetée. Si on y a été sensible, c'est de manière totalement non consciente puisque l'œil se trouvait fixé dans la phrase bien en amont du lieu critique. Ce résultat est tout à fait congruent avec ceux que nous avons rappelés dans un article précédent (*A.L.* n°91) qui montraient que les informations non conscientes peuvent servir de supports à des manipulations cognitives (Dehaene et al., 1998 ; Naccache & Dehaene, 2001).

Finalement, le fait marquant de cette expérience, c'est que les sujets organisent leurs réponses à travers la perception qu'ils ont du cadre syntaxique de la phrase. Car nous avons veillé à ce que les mots à rôle sémantique soient les plus proches possibles (si ce n'est identiques) dans les trois phrases du triplet : ce ne sont donc pas les informations sémantiques différentes qui guident le choix de la phrase, c'est uniquement la prégnance d'une organisation syntaxique bien particulière.

La lecture de ces résultats amène une autre remarque concernant cette fois la distance entre le point de fixation et l'endroit de la variation dans les phrases. Nous n'avons pas pu trouver de limites supérieures dans la mesure où, à

³ Pour les 49 élèves de sixième et les 48 élèves de troisième (voir note précédente) les résultats montrent :

- Pour les élèves de sixième, les réponses données pour une apparence visuelle proche sont moins importantes que pour une fonction similaire (41,1% vs. 58,9%). Cette différence est significative ($t = -2,93$; $df = 28$; $p < 0,007$)

- De même, pour les élèves de troisième, les réponses données pour une apparence visuelle proche sont moins importantes que pour une fonction similaire (41,3% vs. 58,7%). Cette différence est significative ($t = -2,86$; $df = 28$; $p < 0,008$)

27 signes d'écart, on continue à observer un pourcentage de réussite important. Nous pouvons également regretter de n'avoir pas pu observer d'interactions, en raison d'un nombre trop faible d'observations à l'intersection des différents facteurs étudiés⁴. En conséquence, et en s'appuyant sur ces premiers résultats significatifs, d'autres protocoles expérimentaux s'en inspirant seront nécessaires, en prévoyant d'éloigner encore le lieu de la variation et en construisant des ensembles en plus grande quantité.

Nous voudrions rappeler que les résultats présentés ici ne signifient pas que la seule source d'information pour établir un cadre syntaxique soit d'origine parafovéale. Nous avons déjà rapporté les expériences de Saint-Aubin qui fait apparaître les textes mot à mot sur un écran d'ordinateur. Cette présentation empêche toute prise d'information dans la zone périphérique de l'œil. Malgré cela, on observe un effet important de la fonction syntaxique du mot (Saint-Aubin & Klein, 2001). Ce qui se précise, c'est la relation dialectique entre les informations présentes en amont du lieu et du moment de la construction du cadre syntaxique (connaissance de la syntaxe de la langue, relation syntaxe-discours, connaissance du style de l'auteur,...) et la perception subliminale de la majorité des mots qui cristallisent l'organisation de la phrase. Ces deux types d'informations peuvent suffire, séparément, pour prendre une décision syntaxique (quoique avec moins d'efficacité), mais ce n'est que de leur union qu'on peut attendre les meilleures performances.

Pour finir, si ce résultat a de l'importance dans la compréhension des processus de lecture, il en a également dans une optique pédagogique. Nous faisons l'hypothèse qu'un entraînement spécifique devrait permettre une amélioration de la compréhension. Il semble indispensable de réserver, à l'intérieur de cet entraînement, un moment particulier pour travailler la qualité de cette perception subliminale. Cet aspect n'est pas forcément le plus simple, dans la mesure où il semble éloigné des préoccupations des professeurs de français, surtout quand il est question de syntaxe ! Nous verrons dans un article ultérieur, l'apport d'un tel entraînement.

Denis FOUCAMBERT ■■■■

Bibliographie

Dehaene, S., Naccache, L., Le Clec'h, G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G. et al. (1998). Imaging unconscious semantic priming. *Nature*, 395, 597-600.

Koriat, A. and Greenberg, S. N. (1996). The Enhancement Effect in Letter Detection : Further Evidence for the Structural Model of Reading. *Journal of Experimental Psychology : Learning, Memory and Cognition*, 22, 1184-1195.

Naccache, L. & Dehaene, S. (2001). Unconscious semantic priming extends to novel unseen stimuli. *Cognition*, 80, 215-229.

Saint-Aubin, J. & Klein, R. M. (2001). Influence of parafoveal processing on the missing-letter effect. *J. Exp. Psychol. Hum. Percept. Perform.*, 27, 318-334.

⁴ Sur les 32 triplets de départ, on n'en a plus que 16 dont le mot initiant la différenciation a un rôle syntaxique. Au croisement de ce rôle et des 4 distances, on ne trouve plus que 4 triplets.

dossier
LE 5 DERNIER 5
une recherche-action

