

Le rôle du schème et de la forme globale du mot dans la reconnaissance des mots écrits en arabe

Assia Boumaraf Ghania Droua-Hamdani Fouzia Badaoui

CRSTDLA

CRSTDLACRSTDLA

chikha.boumaraf57@gmail.com gh.droua@post.com foubadaoui@yahoo.fr

Résumé

Le but de cette étude est de mettre en évidence les caractéristiques morphologiques de la langue arabe et plus précisément à explorer le rôle du schème vocalique et de la forme globale des mots dans la reconnaissance des mots arabes écrits. A travers l'utilisation du paradigme d'amorçage masqué (Ferrand and Grainger, 1994; Grainger and Jacob, 1999 ; Segui and Grainger, 1990), nous avons mené deux expériences ; la première porte sur le schème vocalique et la deuxième traite la forme globale du mot écrit. Les résultats montrent l'absence de l'effet du schème vocalique et plaident pour un effet inversé de la forme globale sur la reconnaissance des mots écrits.

Mots clés:

Schème vocalique, forme globale du mot, reconnaissance, mot écrit, arabe.

المخلص:

تهدف هذه الدراسة إلى تسليط الضوء على الخصائص الصرفية للغة العربية وكيفية تأثير كل من الوزن الصرفي والشكل العام للكلمة في التعرف على الكلمات العربية المكتوبة، وذلك من خلال إجراء تجربتين بالاعتماد على تصميم المعالجة المسبقة بطعم: تجربة الأولى تخص الوزن الصرفي، والثانية متعلقة بالشكل العام للكلمة. أسفرت النتائج عن غياب أثر للوزن الصرفي في التعرف على الكلمات المكتوبة، مع وجود أثر عكسي للشكل العام للكلمة.

الكلمات المفتاحية:

الوزن الصرفي، الشكل العام للكلمة، التعرف، الكلمات المكتوبة، العربية

Abstract

The purpose of this research is to highlight the morphological characteristics of the Arabic language and more precisely to explore the role of the word pattern and the overall form of the word in the recognition Arabic written words. By using the masked priming paradigm (Ferrand and Grainger, 1994; Grainger and Jacob, 1999; Segui and Grainger 1990), we conducted two experiments: the first concerns the word pattern and the second deals with the global word shape. Results showed no word pattern effect, and a reversed global word on the recognition of arabic written word.

Keywords:

Word pattern, global word shape, recognition, written words, Arabic

Introduction

La racine et le schème sont les unités de base pour la formation des mots en langue arabe.

Les études menées sur le lexique mental en langues sémitiques ont mis l'accent sur le rôle joué par la racine dans la reconnaissance des mots écrits (Prunet, Béland et Idrissi, 2000, et Frost, Forster et Deutsch, 1997 pour l'hébreu). Cependant, le statut du schème reste encore ambigu dans ce domaine. En effet, Mimouni, Kehayia & Jarema 1998 en testant 24 sujets normaux et 02 aphasiques à travers le paradigme d'amorçage auditif associé à une tâche de décision lexicale des verbes, ont démontré un effet d'amorçage pour la racine des verbes mais pas pour le schème pour une SOA (Le délai qui sépare la présentation du distracteur de celle de la cible est appelé SOA, ce qui correspond en anglais à l'abréviation de Stimulus Onset Asynchrony) (Patrick Bonin, 2003, p.34) de 250ms. Mahfoudhi (2007) confirme la présence d'un effet significatif pour la racine des verbes trilitères (conjugués au passé) et en même temps, l'absence d'un effet de schème, que ce soit pour les schèmes sains et des *الأوزان* (Grainger et coll 2003). Ces résultats corroborent ceux de Prunet, Béland & Idrissi (2000) qui ont analysé les erreurs commises par les sujets normaux lors d'un *tongue split* et les erreurs commises par un aphasique (ZT) et ont abouti au fait que les erreurs de métathèse touchent sélectivement la racine des mots et non pas le schème. Par ailleurs, Abu-Rabia & Awwad (2004) n'ont trouvé aucun effet distinctif de la racine et du schème sur la réponse du sujet lorsque la durée de présentation de l'amorce est de 50ms. De même, Boudelaa & Marslen-Wilson (2005) ont pu démontrer un effet significatif de la racine sur les verbes et sur les noms lorsque les durées de présentation de l'amorce sont de 32, 48, 64 et 80ms et un effet significatif pour le schème uniquement pour les durées de 48 et 64 ms. Dans le même cadre Grainger, Dichy, El Halfaoui & Bamhamed (2003) ont mené deux expériences. La première avait pour objectif la vérification de la présence d'un effet d'amorçage orthographique sur les mots arabes (figure ci-dessous). Cet effet a été observé dans des langues alphabétiques latines comme le français et l'anglais. La seconde expérience visait à tester le rôle de la fréquence des mots et leur racine dans la reconnaissance des mots arabes écrits.

Les résultats de l'expérience 1 ont montré une quasi-absence d'effet de l'amorçage orthographique sur les mots arabes tri-consonantiques. Cependant, ceux de l'expérience 2 ont exposé un effet significatif de la racine dans la reconnaissance des mots, même lorsque ces derniers sont de faible fréquence. Les auteurs expliquent que la faiblesse des effets d'amorçage orthographique observés dans les langues sémitiques est due au rôle réduit que joue la lettre dans l'identification des mots chez les lecteurs experts. De même, ils accordent une grande importance à la forme globale des mots dans cette tâche et citent à cet effet «*En arabe et en hébreu, les informations relatives à la forme globale des mots (son enveloppe) sont présentes de manière systématique et avec peu de variabilité,.....bien que le rôle de ce type d'information reste à étudier dans une perspective expérimentale* ». (Grainger et coll, 2003, P. 86).

Les données issues de la pathologie du langage sont en faveur de l'effet la forme globale du mot sur la reconnaissance des mots. En effet, cet effet a été déjà observé chez une patiente arabophone, âgée de 68, atteinte d'une aphasie de Broca et d'une dyslexie profonde suite à un accident vasculaire cérébral. En effet, HY a accepté les pseudo-homographes (c'est-à-dire les non-mots qui ont les mêmes formes globales des vrais mots) comme étant de vrais mots dans une tâche de décision lexicale mais elle a rejeté les pseudo-homographes (c'est-à-dire les non-mots se prononcent de la même manière que des vrais mots). De même, HY a commis des erreurs de lexicalisation lors de lecture des non-mots pseudo-homographes (Boumaraf & Macoir, sous presse).

Notre présente étude qui s'inspire de la recherche de Grainger et al, 2003 a pour objectif de mettre en lumière le rôle du schème et l'effet de la forme globale sur la reconnaissance des mots écrits en arabe. L'article est organisé comme suit : section 1 expose des notions sur la langue arabe, section 2 la méthodologie adoptée dans l'étude à savoir l'échantillon d'étude, les stimuli et le paradigme et le déroulement de l'expérience. Section 3 présente les résultats obtenus. La recherche se termine par une discussion et une conclusion.

1. Notions sur la langue Arabe

1.1. Description du système graphique arabe

L'alphabet arabe ('abgadiyya) comprend vingt neuf lettres fondamentales. Certaines lettres arabes sont présentées avec des points, ces derniers ont une valeur distinctive. En fait, ils constituent la seule marque de distinction entre plusieurs graphèmes, puisque les phonèmes (ح, ج, خ) (h, g, x) ne sont pas différenciés à l'écrit que par la position du point, ce qui est également le cas pour (ث, ت, ص, ض, ذ, ر, ز, س, ش, ط, ظ) (1/4, ®, °, s, †, d, ‡, ©,)

Par conséquent, un seul point peut différencier deux mots ou plus, par exemple: (حديث / جديد) [ħadi:d-ġadi:d] (traduction: fer et nouveau)

- L'alphabet arabe, comme les alphabets des autres langues sémitiques est consonantique; le lecteur doit connaître la structure de la langue pour restituer les voyelles (brèves), car seules les voyelles longues sont représentées par des lettres (la lettre "و" (w) représente la voyelle ū mais aussi la consonne (w) et les glides (aw, iw).
- L'écriture arabe est une écriture monocamérale : les notions de majuscule et de minuscule n'existent pas. Le sens de l'écriture est de droite à gauche.
- En général, les lettres arabes s'attachent entre elles, qu'ils agissent d'une écriture manuscrite ou imprimée. La lettre change partiellement de forme selon la position qu'elle occupe. Exemple, la lettre ṣn (س) (س) se transcrit (س) en position antérieure, (س) en position finale, (س) en position médiane.
- Le système d'écriture de la langue est très transparent; il y a une très grande corrélation entre le phonème et la lettre qui lui correspond.

L'orthographe arabe

(1) *L'orthographe de surface* est renforcée par les signes diacritiques où les consonnes et les voyelles longues sont représentées par des lettres alors que les voyelles brèves sont représentées par des symboles ;

L'orthographe de surface est utilisée dans le Coran, les livres scolaires, les matériaux pédagogiques pour les premières années de l'apprentissage de la lecture et de l'écriture, et dans les cas où l'absence des signes diacritiques entraîne une ambiguïté.

Exemple

كَتَبَ/kataba/ il a écrit كُتِبَ/kutiba/ il a été écrit.

كُتُبَ/kutub/des livres.

- *L'orthographe profonde* dépourvue de signes diacritiques : les consonnes et les voyelles longues sont présentes, les voyelles brèves sont absentes. S. Bentin & R. Ibrahim (1996).

1.2. Les caractéristiques de la morphologie des langues sémitiques

La morphologie des langues sémitique est une morphologie non-concaténative, c'est-à-dire que les mots ne sont pas formés par la simple juxtaposition des morphèmes dans une séquence linéaire comme pour le mot : trans.form.ation.

Le mot en arabe est la synthèse de la combinaison de deux unités abstraites, à savoir, la racine et le schème.

➤ La racine : (la matière)

La racine arabe est consonantique et elle est généralement trilitère (C1, C2, C3).

(k, t, b) (ك، ت، ب) (‘, l, m) correspond au domaine de la science, et la racine (ع، ل، م) Par exemple : la racine correspond au domaine de l'écriture.

➤ Le schème (le moule) :

فاعِل C'est le modèle sur lequel se moule la racine pour former le mot. Par exemple: le schème /fa:ʕel/ (agent) correspond au masculin singulier du nom. correspond au masculin singulier passé du verbe. (le /f/ correspond à la première فَاعِل Le schème / faʕala/ consonne, le /ʕ/ correspond à la deuxième consonne, et le /l/ correspond à la troisième consonne).

Il est important de souligner que les consonnes du mot ne sont pas toutes relatives à la racine ; le schème lui aussi peut avoir des consonnes qui peuvent occuper la position antérieure, finale ou médiane. Par exemple, la consonne (taʔ) bouclée correspond toujours au schème puisqu'elle n'est pas permanente,

contrairement aux consonnes de la racine qui sont permanentes : l'absence d'une seule consonne de la racine détruit le mot.

La racine et le schème sont des unités abstraites qui n'existent pas d'une façon isolée dans la parole. La linéarité de la parole ne permet pas de faire apparaître ces deux éléments. Seule l'analyse verticale, par la mise en correspondance ou (qiya:s)¹, permet d'extraire la racine et le schème.

2. Méthodologie

2.1. Population

L'étude a été menée sur 33 sujets arabophones volontaires, âgés de 20 à 49 ans ayant un niveau d'instruction supérieur, et une vision parfaite ou corrigée, n'ayant reçu aucune information concernant le but des deux expériences.

2.2. Paradigme d'amorçage subliminal ou rapide :

Dans le cadre de cette étude nous avons mis sur pied deux expériences par la technique de l'amorçage rapide ou masqué mis en œuvre par (Ferrand & Grainger, 1994, Grainger & Jacob, 1999, Segui & Grainger, 1990).

Le procédé de l'expérience consiste à afficher en premier lieu une étoile pendant 1400ms pour focaliser le regard du sujet sur l'écran, suivi d'une amorce qui reste sur l'écran pour une durée de 25 ms considérée assez brève pour que le sujet n'en perçoive pas la présence. Nous avons constaté que pour l'arabe la durée de 50ms est assez longue pour permettre la mémorisation de l'amorce. Le but de cette démarche est d'étudier le traitement perceptif qui a lieu en dehors de toute prise de conscience par le sujet (Grainger et coll, 2003). Un masque apparaît immédiatement après l'amorce pour une durée de 2500ms, afin d'effacer toute trace mnésique de l'amorce et pour préparer l'apparition du mot cible qui reste présent sur l'écran pour une durée de 2500ms. Le sujet dispose de 4000 ms pour décider si le mot cible est un vrai mot ou pas. A l'issue du résultat, le processus se répète pour les autres items (voir figure 1).

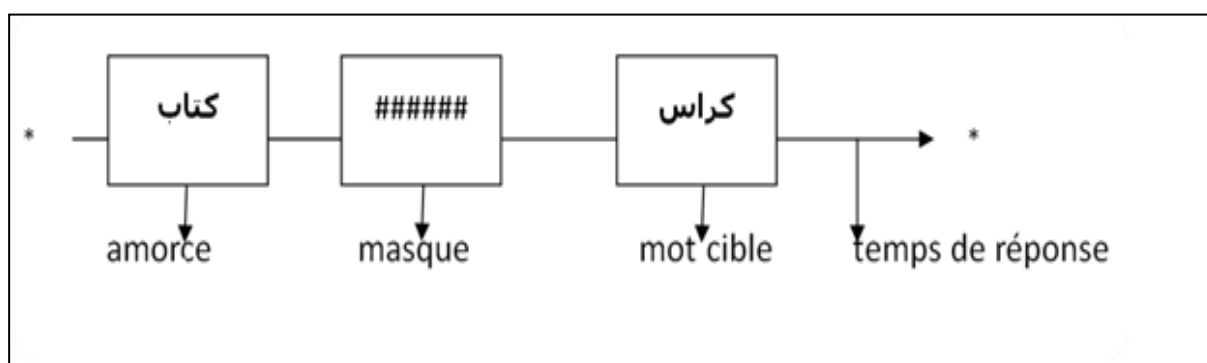


Figure 1 : déroulement de l'expérience

2.3. Stimuli

Le qiya:s est cette mise en équivalence entre deux unités ou plus, afin de déterminer une structure commune (voir Hadj-Salah,¹ 1979).

Notre étude comprend deux expériences dont les objectifs sont l'étude du rôle du schème et de l'effet de la forme globale du mot sur la reconnaissance du mot arabe écrit. Les stimuli relatifs sont :

Stimuli 1

L'expérience comprend 64 mots et 64 non-mots (de 4 à 5 lettres). Les mots utilisés sont des mots fréquents de la langue arabe. En raison de l'indisponibilité d'une table de fréquence pour la langue arabe, nous avons eu recours à des spécialistes pour juger de la fréquence des stimuli utilisés. En plus de cela et afin de familiariser les sujets avec les tâches demandées, six essais ont précédé les deux expériences.

Pour les non-mots, nous avons construit des non-mots prononçables en respectant les règles phonotactiques de l'arabe. Nous avons vérifié tous les non-mots dans le dictionnaire électronique (El ma'ani)² afin de nous assurer qu'ils ne sont pas de vrais mots.

Nous avons testé les cibles (mots et non mots) dans les quatre conditions suivantes :

- 1) Condition1 : le schème, l'amorce et le mot ont le même schème (reliés) ou (non reliés).
- 2) Condition2 : selon le statut de l'amorce. (mot ou non mot).

Conditions expérimentales de la 1ere expérience

Les conditions	Schème	Amorce	Cible
Condition 1	Relié	Non mot دفات عائن	Mot جمال Non mot شالم
Condition 2	Non relié	Non mot دفات عائن	Mot جمال Non mot شالم
Condition 3	Relié	Mot حمراء مقهى	Mot نجلاء Non mot محزى
Condition 4	Non relié	Mot بريق مرضى	Mot خمسة Non mot محقف

Procédures expérimentales

Le sujet est assis sur une chaise à environ 50cm de l'écran d'un ordinateur portable DELL Vosto 3350. Le sujet est appelé à répondre le plus vite possible et à décider si le mot qui apparaît après le masque est un vrai mot ou non en appuyant sur la touche Y pour oui et la touche N pour non.

Les paires des quatre conditions sont réparties suivant un ordre différent d'un sujet à l'autre.

Pour les 64 réponses, nous avons obtenu 32 réponses en oui et 32 réponses en non.

Pour la réalisation de l'expérience, nous avons utilisé le logiciel Psychopy v.1.79 de Jon Peirce 2012, université de Nottingham.

Analyse des résultats

Dans un premier temps les moyennes des temps de réaction des réponses correctes pour chaque item et pour chaque participant ont été calculées avant de procéder à l'analyse statistique proprement dite. Nous avons éliminé les participants qui ont commis plus de 20% d'erreurs.

Une analyse de variance a été réalisée pour dégager les différences statistiquement significatives.

Une étude statistique des temps de réactions a été effectuée pour déceler les relations qui existent entre l'amorce et le mot cible (tableau suivant).

Les moyennes calculées

Facteurs	Moyennes	SD
Schème		
Schème relié	4.0417	0.98
Schème non-relié	4.0452	0.96
Amorce		
Amorce mot	4.0728	0.95
Amorce non-mot	4.0141	1
Cible		
Cible mot	3.51	0.71
Cible non-mot	4.5769	1.3

Résultats Statistiques : calcul de l'ANOVA

Facteurs	F	P
Schème	F(1,28)=.005	p=.94
Amorce	F(1,28)=.012	p=.34
Cible	F(1,28)=47.54	p=.0001

Il n'existe pas d'effets significatifs pour les facteurs indépendants schème, amorce

Les résultats montrent qu'il n'y a pas un effet significatif du schème, ni des statuts de l'amorce mais un effet de lexicalité est observé (performance meilleurs lorsque la cible est un mot VS non-mot).

Expérience 2 :

Stimuli 2

Cette expérience comprend 64 mots et 64 non mots, ici aussi nous avons utilisé des mots fréquents de la langue arabe et des non-mots, respectant les règles phono-tactiques de l'arabe.

Procédures de l'expérience 2

Les mêmes procédures techniques utilisées dans la première expérience ont été envisagées pour la deuxième expérience.

Conditions expérimentales de la 2eme expérience

	Forme globale	Amorce	Cible
Condition 1	reliée	Non mot قرآن زرهية	Mot قرآن Non mot زرمية
Condition 2	Non reliée	Non mot جندق شيدم	Mot قفاز Non mot فلدن
Condition 3	reliée	Mot فانوس صاروخ	Mot قاموس Non mot ضازوج
Condition 4	Non reliée	Mot عطور كمال	Mot طالب Non mot نامض

Dans les 4 conditions de cette 2eme expérience, nous avons manipulé d'une part le statut de la cible (mot, non mot) et de l'autre la forme globale des mots. Le but de cette expérience est de vérifier si le fait de présenter une amorce qui partage la même forme globale du mot cible va faciliter son traitement.

Nous faisons l'hypothèse que l'effet de la forme globale va être observé lorsque l'amorce (quel que soit son statut) partage la même forme globale du mot cible (vrai mot). Nous faisons l'hypothèse que cet effet va être observé uniquement lorsque la cible est un vrai mot par ce qu'il s'agit des formes globales (représentations graphiques) stockées dans le lexique orthographique.

Résultats

Les moyennes calculées

Facteurs	Moyennes	SD
Forme Globale		
Forme reliée	4.0375	.86
Forme non-reliée	3.7861	.71
Amorce		
Amorce mot	3.9332	.75
Amorce non-mot	3.8904	.78
Cible		
Cible mot	3.4714	.74
Cible non-mot	4.3521	.91

Résultats Statistiques : calcul de l'ANOVA

Facteurs	F	P
Forme Globale	F(1,27)=8.798	p=.006
Amorce	F(1,27)=.850	p=.365
Cible	F(1,27)=44.91	p=.0001

En plus de l'effet de la lexicalité qui a été déjà noté dans la première expérience, les résultats montrent une différence significative entre les temps des réactions des reliées et des paires non reliées. Ces dernières ont eu des temps des réaction plus courts que les paires qui partagent la même forme globale.

Ces résultats plaident pour un effet renversé de la forme globale des mots. Un effet semblable a été observé lorsque le stimuli était orthographiquement similaire à un mot fréquent; ce phénomène est nommé « effet de fréquence du voisinage orthographique » (voir Grainger & Segui, 1990).

Cette expérience a mis en valeur l'importance que joue la forme globale dans la reconnaissance des mots en arabe.

Conclusion

Au cours de cette recherche, nous avons essayé de mettre en lumière des facteurs qui peuvent influencer le traitement des mots écrits en arabe, mis à part la racine ; en l'occurrence le schème et la forme globale du mot.

Nos résultats montrent la non pertinence du schème dans la reconnaissance des mots écrits malgré l'effet significatif de l'interaction entre le schème et le statut de l'amorce d'une part. Et d'autre part, l'importance de la forme globale dans le traitement des mots écrits. Cependant, des études plus approfondies sont nécessaires

Revue al-lisaniyyat Vol 18 N 1

pour confirmer les résultats obtenus afin de déceler les mécanismes relatifs à la reconnaissance des mots écrits en arabe.

Bibliographie :

- Abu-Rabia, S & Awwad, Y. 2004. Morphological structures in visual word recognition: The case of Arabic. *Journal of Research in Reading*, 27, 321-336.
- Bentin, S., & Ibrahim, R. 1996. New evidence of phonological processing during visual word recognition: the case of Arabic. *Journal of Experimental Psychology: Learning Memory and Cognition*, 22(2), 309–323.
- Bonin, Patrick. 2003. Production verbale de mots. De Boeck et Larcier. Bruxelles.
- Boudelaa, S & Marslen-Wilson. 2005. “Discontinuous morphology in time: Incremental masked priming in Arabic”. *Language and Cognitive Processes* 20(1-2):207-260
- Boumaraf, A., 2010. Les troubles du langage observés au niveau de l’écrit chez les aphasiques arabophones », in *Al-lisaniyyat*, N°14-15, 97-136.
- Ferrand, L., Grainger, J., & Segui, J. (1994). Masked form priming in picture and word naming. *Memory and Cognition*, 22, 431-441.
- Frost, R., Forster, K.I., and Deutsch, A. 1997. What can we learn from the morphology of Hebrew: A masked priming investigation of morphological representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 829-856.
- Grainger, J., Dichy, J, El Halfaoui, M. & Bamhamed, M. 2003. Approche expérimentale de la reconnaissance du mot écrit en arabe, *Faits de langues*, 22, 77-86.
- Grainger, J. & Jacobs, A.M. (1999). Temporal integration of information in orthographic priming. *Visual Cognition*, 6, 461-492.
- Grainger, J. and Juan S. (1990). “Lexical decision and masked identification latencies neighborhood Frequency effects in visual word recognition: A comparison of lexical decision and masked identification latencie
- s
- Hadj Salah, A. 1979. Linguistique arabe et linguistique générale, Essai d'épistémologie du ‘ilm al ‘arabiyya, thèse pour le doctorat (deux tomes), Paris-Sorbonne.
- Mahfoudhi, A., 2007. Roots and patterns in Arabic lexical processing. *Perspectives on Arabic Linguistics XIX: Papers from the Nineteenth Annual Symposium on Arabic Linguistics*, Benjamins
- Mimouni, Z., E. Kehayia, and G. Jarema. 1998. ‘The Mental Representation of Singular and Plural Nouns in Algerian Arabic as Revealed through Auditory Priming in Agrammatic Aphasic Patients’. *Brain and Language* 61: 63–87.
- Pillon, A, 1993. La mémoire des mots : ses unités, son organisation. Bruxelles, Pierre Mardaga,
- Prunet, J-F., R. Béland, and A. Idrissi. 2000. ‘The Mental Representation of Semitic Words’. *Linguistic Inquiry* 31(4): 609–48.

Segui, J, & Grainger, J. (1990). Priming word recognition with orthographic neighbors: Effects of relative prime-target frequency. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 65-76.

