Abstractness in Lexical Processing: How far could we go?

Abdessatar Mahfoudhi
University of Ottawa

Abstract: In the traditional theory of Arabic morphology, words are made by the interleaving of roots into patterns, both of which are abstract unpronounceable units. The root, often made of three consonants, carries the core meaning of the word and the pattern, made of the root consonant slots, vowels and consonant affixes, holds the syntactic meaning. The more recent etymon theory proposes that the root should be replaced by the bi-consonantal etymon, which is the segmental manifestation of the more abstract notion of the phonetic matrix. The two segments of the etymon or the two feature bundles of the matrix can be reversible, i.e. (x, y) or (y, x). This study examined the role of roots, etymons, and matrices in lexical processing in three lexical decision experiments. It also examined how abstract the processing of these units could be. Finally, the implications of the results for theories of morphology and lexical processing are discussed. The results suggest that roots and etymons are used in word recognition. Yet, there seems to be a limit on how abstract the processing of etymons and phonetic matrices could be.

Keywords: Arabic morphology, roots, etymons, phonetic matrices, lexical processing, abstractness

1. Introduction

Most theories of Arabic morphology propose that Arabic word-formation is based on the nonlinear mixing of very abstract morphemes. They, however, disagree on which units are at play. The traditional Root and Pattern Theory and Prosodic Morphology suggest that words are made by the interleaving of roots into patterns, both of which are abstract unpronounceable units. The Etymon Theory proposes that the root should be replaced by the bi-consonantal etymon and the related notion of the phonetic matrix. These three units are abstract. The root and the etymon are both discontinous unpronounceable units. Besides, some etymons can have reversible segments, i.e. allomorphs like (b, t) and (t, b). The phonetic matrix is not made up of segments but two feature bundles like {[labial], [coronal]}, which can also be reversible.

The present study examines the cognitive validity of roots, etymons and phonetic matrices in lexical processing. The second objective is to see how abstract the processing of these abstract morphemes could be. That is, the study tries to find out whether priming could be achieved with units with discontinous segments (roots and etymons) or with just shared bundles of features. Furthermore, it examines whether processing could access the underlying order-

CLO/OPL
Janvier/January 2006, Vol. 34: 75-85

free etymon and abstract away from the linear order in the allomorphs of reversible etymons. Finally, the implications of the results for theories of morphology and theories of lexical processing will be discussed.

2. Arabic Morphology: Roots or Etymons and Matrices?

The traditional theories of Arabic morphology (e.g., Cantineau, 1950) postulate that word formation is based on the mapping of roots onto patterns. For instance, the word [**ħ**a**m**a**l**] 'carried' is made of the root {**ħ**, **m**, **l**} and the pattern {CaCaC}. The root carries the core meaning of the word 'carrying' and the pattern has the syntactic meaning 'perfective, active'. Prosodic morphology (e.g., McCarthy, 1981) also uses the root but proposes that the pattern should be an underspecified skeleton and that the vowels and the affixal consonants have a morphemic status. These two theories have recently been challenged by the less common etymon theory (Bohas, 2000), which argues that the root is not able to capture the semantic similarity between words sharing only two consonants and should be replaced by the etymon. The etymon consonants can be reversible (e.g., *ʔiʕta**l**la* 'he/it fell sick' and ***l**awwaʕa* 'he/it tormented'). The etymon itself is the segmental instantiation of a bundle of phonetic features, i.e. a phonetic matrix. The phonetic matrix is made of "une combinaison de traits phonétiques et de noyaux sémiques" (Bohas, 2000:64). It relates via a complex semantic network of words sharing two or more unordered phonetic features, as in (1). Included under the same matrix then are words that share either the same etymon as in (1a-b), i.e. (**b**, **t**) or only place features in two different phonemes (1c-d).

(1) Arabic words sharing the phonetic matrix {[labial], [coronal]}

a.	ħa**b**a**t**a	'to hit'
b.	ha**b**a**t**a	'to hit'
c.	ha**b**a**ʃ**a	'to hit someone and hurt him/her'
d.	ra**f**a**z**a	'to hit; to beat up'

3. Previous Studies

Boudelaa and Marslen-Wilson (2005) found a rather persistent priming effect of the root in lexical decision. However, there is no previous research on the phonetic matrix, and evidence for the etymon is both scarce and controversial. Evidence for the etymon comes from a lexical decision experiment by Boudelaa and Marslen-Wilson (2001). These results were met with much skepticism (Bentin and Frost, 2001).

4. Experiment 1: Roots

4.1 Objectives

The present study tested whether a masked word prime would facilitate the recognition of a word having the same root. Given some evidence for the importance of the root in lexical processing in Arabic (Boudelaa & Marslen-Wilson, 2005), I expected to find a priming effect of the root.

4.2 Participants

The participants were 36 Arabic-speaking[1] students from Tunisia, where all the experiments were conducted. They were aged between 22 and 27 and all had at least 12 years of formal education in Standard Arabic.

4.3 Stimuli and Design

The targets were 48 triliteral Arabic verbs in the third person singular perfective in two sets of data, controlled for familiarity in the absence of reliable frequency counts for Arabic. The first set included 24 targets that were paired with primes from one of these three conditions: +Root +Semantics, +Orthography/+Phonology, and Unrelated. In the first condition, primes and targets had the same root and a transparent semantic relationship. In the second condition, primes shared with targets roughly the same number of letters as in the related conditions. The third condition included urelated primes. The second set of targets were paired with primes from these conditions: +Root-Semantics,

[1] Modern Standard Arabic, the variety examined in this work, is the written variety used for education and administration. It is the variety shared, with some differences, by all Arab countries. The spoken varieties are highly diverse and can be very different from the standard variety.

+Orthography/+Phonology, and Unrelated (see Figure 1). In the first condition, primes and targets shared the same root but very little meaning. The primes in the second condition shared the same number of letters with the targets, as did the morphological primes. In the third condition, the primes were unrelated in form and meaning to the targets. The unrelated condition served as the baseline against which the priming effects were measured. If there was an effect of the root we would find significant priming in the first condition in both sets that is different from the orthographic/phonological controls in the second condition in both sets.

Figure 1

Examples of prime-target pairs used in study 1, with Arabic script, phonetic transcription, and gloss

	Prime	**Target**
Set 1		
	تـقـسّـم	قـاسـم
1. +Root+Sem	/ta**q**a**s**sa**m**a/	/**q**aasa**m**a/
	‘was divided’	‘shared’
	تـقـاعـس	قـاسـم
2. +Orthog/+Phono	/ta**q**aa**ʕ**asa/	/**q**aasa**m**a
	‘was uninterested’	‘shared’
	تـصـدّر	قـاسـم
3. Unrelated	/ta**sˤ**a**d**da**r**a/	/**q**aasa**m**a/
	‘occupied the leading position’	‘shared’
Set 2		
	احـتـرم	حـرّ م
1. +Root–Sem	/ʔi**ħ**ta**r**a**m**a/	/**ħ**a**r**ra**m**a/
	‘respected’	‘forbid’
	تـكـرّ م	حـرّ م
2.+Orthog/+Phono	/ta**k**a**r**ra**m**a/	/**ħ**a**r**ra**m**a/
	‘showed one’s generosity’	‘forbid’
	تـوطّـد	حـرّ م
3. Unrelated	/ta**w**a**t**ˤ**t**ˁa**d**a/	/**ħ**a**r**ra**m**a/
	‘was strengthened’	‘forbid’

4.4 Procedure and Apparatus

One third of the 36 participants were arbitrarily assigned to one of three lists. The participants were instructed to respond as quickly and as accurately as possible. This experiment and all the following experiments were conducted on an HP portable Computer running the display system DMDX.[2] Each trial consisted of three events. The first event was a mask of 28 vertical lines (following Boudelaa

[2] The DMDX software was developed by J. C. Forster at the University of Arizona.

and Marslen-Wilson 2001) that was displayed for 500 ms. The second event that immediately followed was a prime word that appeared for 50.25 ms (three ticks). The prime duration was short enough to avoid any episodic memory traces. The last event that immediately followed the prime was a target word, which remained on the screen for 2000 ms. or until a response was provided.

4.5 Results

The averages of correct response times and mean error frequencies were obtained for both participants and items and were analyzed using separate analyses of variance. The two independent variables were prime condition and list. However, the effect of list will not be reported because it was introduced to reduce variance.

Table 1

Lexical decision reaction times (RTs), standard deviations (SD), and percentage error rates (% error) in study 1

Condition	RT(ms)	Mean difference	SD	% error
1. +Root+/-Sem	716	-15.66	86	6.5
2. +Orthog/+Phono	735	+3.34	94	10.9
3. Unrelated	744	+12.34	92	9.7
1a.+Root+Sem	761	-3	113	5.6
2a. +Orthog/+Phono	766	+2	101	7.9
3a. Unrelated	765	+1	95	6.5
1b.+Root-Sem	700	-21.33	79	6.4
2b. +Orthog/+Phono	721	+.33	98	11.1
3b. Unrelated	743	+21.67	110	8.5

To check whether the root had a special priming effect, I ran a set of ANOVAs on the first three conditions (see means in Table 1). Overall, prime condition was significant only in subject analysis, $F_1(2, 66)=5.82$, $p<.005$.[3] Planned comparisons revealed a significant difference between the morphological condition and the unrelated condition: $F_1(1, 33)=10.65$, $p<.005$ and between the morphological condition and the form condition, $F_1(1, 33)=7.24$, $p<.05$.

To test whether semantics had a priming effect in the morphological condition, I further ran two other sets of ANOVAs. The first set included these three conditions: +Root, +Semantics, +Orthography/+Phonology, and Unrelated.

[3] Item analysis (F2) is reported only when significant. The lack of significant effect in item analysis, when there is an effect with subjects, is due to a large variance within items.

The analysis did not yield a main effect of prime condition. The second set of ANOVAs included the (i) +Root-Semantics, (ii)+Orthography/+Phonology, and (iii) Unrelated conditions. The prime condition was significant, $F_1(2, 60)= 6.73$, $p <.005$. A deviation planned comparison test revealed that only the contribution of the morphological condition was significant, $F_1(2, 60)= 8.62$, $p <.01$.

4.6 Discussion

The results of this experiment support the hypothesis that the root has a priming effect and therefore has a cognitive validity in the Arabic mental lexicon that is different from the formal and the semantic effect.

5. Experiment 2: Etymons

5.1 Objectives

This study aimed first to revisit the question of whether the etymon has a role in Arabic lexical processing. It also examined whether the order of the consonants of the etymon (i.e. the allomorphy) had an effect.

5.2 Participants

Thirty-six Arabic-speaking students participated as volunteers in the experiment. None had participated in Experiment 1.

5.3 Stimuli and Design

The targets were 48 triliteral verbs in the past tense third person singular with similar familiarity scores. Each of these words was paired with three primes, one from each of the three conditions: (i) +Etymon (ordered or not), (ii) +Orthography/Phonology, and (iii) Unrelated. But to allow for an evaluation of the priming effect of the order of etymon consonants, the target words were divided into two sets of three conditions (see Figure 2). The first set included three conditions: (i) +Ordered Etymon, (ii) +Orthography/+Phonology, and (iii) Unrelated. In the first condition, primes and targets shared an ordered etymon. In the second condition, primes and targets shared two consonants. In the third condition, primes and targets had roughly the same number of different letters. The second set included, in addition to the form condition and the unrelated condition, the +Non-ordered Etymon condition.

Figure 2

Examples of prime-target pairs used in study 2, with Arabic script, phonetic transcription, and gloss

	Prime	**Target**
Set 1		
	قـضـم	انـقـرض
1. +Ordered Etymon, -Sem	/**q**a**ðˁ**ama/	/ʔin-**q**ara**ðˁ**a/
	'gnawed'	'became extinct'
	رفـض	انـقـرض
2. +Orthog/+Phono	/rafadˁa/	/ʔin-**q**ara**ðˁ**a/
	'refused'	'became extinct'
	سـلـك	انـقـرض
3. Unrelated	/salaka/	/ʔin-**q**ara**ðˁ**a/
	'followed'	'became extinct'
Set 2		
	مـوّه	اهـتـمّ
1. +Non-ordered Etymon, -Sem	/**m**awwa**h**a/	/ʔi**h**ta**m**ma/
	'feigned'	'was interested'
	مـتّـع	اهـتـمّ
2. +Orthog/+Phono	/**m**a**t**taʕa/	/ʔi**h**ta**m**ma/
	'made enjoy'	'was interested'
	عـطّـل	اهـتـمّ
3. Unrelated	/ʕatˁtˁala/	/ʔi**h**ta**m**ma/
	'obstructed'	'was interested'

5.4 Procedure and Apparatus

These were the same as in the previous experiment.

5.5 Results

I ran a first set of ANOVAs in which the morphologically-related condition included a balanced number of ordered and non-ordered etymons, i.e., the first three conditions in Table 2. The prime condition was significant by subjects, F_1 (2, 66)=8.66, p <.001. Planned comparisons revealed a significant difference between the +Etymon condition and the unrelated condition, F_1 (1, 33)=18.98, p <.001 as well as between the +Etymon condition and +Orthography/Phonology condition, F_1 (1, 33)=7.74, p <.01. The form condition did not differ from the unrelated condition, F_1 (1, 33)=2.45, p >.05.

Table 2

Lexical decision reaction times (RTs), standard deviations (SD), and percentage error rates (% error) in study 2

Condition	RT(ms)	Mean difference	SD	% error
1.+Etymon+/-ordered	800	-14.66	113	10.9
2. +Orthog/+Phono	816	+2.66	118	12.4
3. Unrelated	828	+13.34	120	14.9
1a.+Ordered etymon	774	-11	115	12.4
2a. +Orthog/+Phono	781	-4	118	14.2
3a. Unrelated	800	+15	123	12.4
1b.+Non-ordered etym	832	-6	134	10.7
2b. +Orthog/+Phono	842	+4	132	9.1
3b. Unrelated	840	+2	130	14.3

To see whether the order of the etymon consonants had an effect on priming, I further ran two sets of ANOVAs. The first included the +Ordered Etymon, +Orthography, and Unrelated conditions. Prime condition was significant, F_1 (2, 66)=4.56, p <.01. A deviation contrast test showed that only the morphological condition had a significant effect, F_1 (1, 33)=6.92, p<.01. In the other set of ANOVAs the +Non-ordered Etymon, the form, and the Unrelated conditions were compared, but there was no significant difference between them.

5.6 Discussion

The results partly replicate what has previously been reported about the priming effect of etymons in Arabic by Boudelaa and Marslen-Wilson (2001). What is new in the present experiment is the fact that priming occurs only when primes and targets share an ordered etymon.

6. Experiment 3: Phonetic Matrices

6.1 Objectives

The purpose of this experiment was to investigate whether a prime would facilitate the recognition of a target with which it shared a phonetic matrix. Since I did not find any priming with non-ordered etymons, I did not expect to find priming with the phonetic matrix.

6.2 Participants

The participants were another 36 Arabic-speaking students from the same population as in the previous experiments.

6.3 Stimuli and Design

Three conditions were studied. In the morphological condition, primes and targets shared a phonetic matrix (see Figure 3). Primes in the phonological condition, as in the morphological condition, were related to the target by sharing roughly the same number and order of features/phonemes. The only difference was that in the phonological condition the primes shared only one of the matrix features with the targets. In the unrelated condition, primes and targets shared at most one non-matrix feature.

Figure 3

Examples of prime-target pairs used in study 3, with Arabic script, phonetic transcription, and gloss

	Prime	**Target**
	لـفـح	نـفّـس
1. +Matrix, -Semantics	/la**f**a**ħ**a/	/na**f**fas**a**/
	‘(a hot wind) blew’	‘relieved’
	فـرض	نـفّـس
2. +Phonology	/**f**ara**ðˤ**a/	/na**f**fas**a**/
	‘imposed’	‘relieved'
	عـلـق	نـفّـس
3. Unrelated	/**ʕ**a**l**i**q**a/	/na**f**fas**a**/
	‘stuck’	‘relieved’

6.4 Procedure and Apparatus

These were the same as in the previous experiments.

6.5 Results

The effect of priming in the related conditions was compared to the phonological condition (see Table 3). The prime condition variable was not significant.

Table 3

Lexical decision reaction times (RTs), standard deviations (SD), and percentage error rates (% error) in study 3

Condition	RT(ms)	Mean difference	SD	% error
1.+Matrix	791	+8.67	131	13.1
2. +Phono	776	-6.33	128	9.6
3. Unrelated	780	-2.33	128	13.7

6.6 Discussion

The results show that the phonetic matrix has no priming effect. The lack of priming by the phonetic matrix was expected, especially that no priming was found with the non-ordered etymon, which is much less abstract.

7. General Discussion

The main result of this study that corroborates what has been found in previous studies on Arabic in lexical processing (Boudelaa & Marslen-Wilson, 2005) is the role of the root in Arabic. The results concerning the etymon are less straightforward. This study has replicated the findings of Boudelaa and Marslen-Wilson (2001) about the priming effect of the etymon. If we accept that the etymon is a morpheme, this result, like the one related to the root, support a morpheme-based theory of Arabic morphology and a theory of lexical processing that proposes a morphemic level of representation such as the dual access hypothesis (e.g., Caramazza, Laudanna & Romani, 1988). Yet, the new finding in this study is that priming does not occur when the shared etymon is non-ordered. While this result could be handled by a connectionist model (e.g., Plaut & Gonnerman, 2000), theories that adopt a lexical-entry-opening metaphor in word recognition cannot easily cope with it. The latter localist theories assume an all-or-none priming effect of morphemes. If the etymon is a morpheme it should facilitate access regardless of whether it is ordered or non-ordered. Still, it could be argued from a localist standpoint that the absence of priming with nonordered etymons is due to a constraint on abstractness in lexical processing. It could be argued that lexical processing is not abstract enough to pick on the nonordred consonants of etymons or the phonetic matrices in word recognition. The

implication of the results of on etymons and the related more abstract notion of the phonetic matrix for theories of morphology/competence and theories of lexical processing/performance is that these theories do not and should not necessarily perfectly match.

References

Bentin, S. and Frost, Ram. (2001) Linguistic theory and psychological reality: A reply to Boudelaa and Marslen-Wilson. *Cognition* 81: 113-118.

Bohas, Georges. (2000) *Matrices et étymons: développements de la théorie*. Lausanne: Éditions du Zèbre.

Boudelaa, Sami and Marslen-Wilson, William D. (2001) Morphological units in the Arabic mental lexicon. *Cognition* 81: 65-92.

Boudelaa, Sami and William D. Marslen-Wilson. (2005) Discontinuous morphology in time: Incremental masked priming in Arabic. *Language and Cognitive Processes* 20(1-2): 207-260.

Cantineau, Jean. (1950) Racines et schèmes. In W. Marcais (Ed.), *Mélanges* (119-124). Paris: Maisonneuve et Cie.

Caramazza, Alfonso, Laudanna, Alessandro and Romani, Cristina. (1988) Lexical access and inflectional morphology. *Cognition* 28: 297-332.

McCarthy, John J. (1981) A prosodic theory of non-concatenative morphology. *Linguistic Inquiry* 12: 373-418.

Plaut, David and Gonnerman, Laura M. (2000) Are non-semantic morphological effects incompatible with a distributed connectionist approach to lexical processing? *Language and Cognitive Processes* 15(4-5): 445-485.