Allomorphic variation in Arabic: Implications for lexical processing and representation

Sami Boudelaa* and William D. Marslen-Wilson

MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 2EF, UK

Accepted 2 December 2003
Available online 4 February 2004

Abstract

This study probes the effects of allomorphy on access to Arabic roots and word patterns in two cross-modal priming experiments. Experiment 1 used strong roots which undergo no allomorphy, and weak roots which undergo allomorphy and surface with only two of their three consonants in some derivations. Word pairs sharing a root morpheme prime each other reliably not only when the root was strong (e.g., [muUaarakun]/[Uaaraka] participant/participate), but also when it was weak (e.g., [Uitifaaqun]/[waafaqa] agreement–agree, where the weak root {wfq} surfaces fully in the target but not the prime). This facilitation occurred even when the weak root surfaced with different semantic meanings across prime and target (e.g., [UittiZaahun]/[UawaZaha] destination/confront). Experiment 2 assessed the effects of allomorphy on word pattern processing, comparing word pairs where the word pattern is transparently realised in both prime and target (e.g., [Uinta]/[Utamala] spread/bear), with pairs which share the same underlying word pattern but where a weak root triggers an assimilation process in the prime (e.g., [Uitta]/[Uibtasama] unite/smile). This assimilation process does not disrupt the CV-structure of the word pattern, in contrast to a third condition where this is disrupted in both prime and target (e.g., [Udaara]/[Uqaala] turn around/say). Strong priming effects were observed in the first two conditions but not in the third. The bearing of these findings on models of lexical processing and representation is discussed.

0093-934X/$ - see front matter © 2003 Elsevier Inc. All rights reserved.

1. Introduction

The component phonemes of a morphologically simple word often assume different phonetic shapes depending on their linguistic environment. A phoneme like /æ/ in English is realised as an oral vowel in hat but as nasalised one in ham; and a coronal consonant like /t/ in sweet, may be produced as a labial /p/ in the context of sweet boy (Gaskell & Marslen-Wilson, 1998). This kind of surface form variation, or allophony, is further compounded by another type of variation, allomorphy, which arises when different morphemes are combined in order to create complex forms. Prominent examples of allomorphic variation include pairs like divine–divinity, and elude–elusive. In the former pair, the diphthong /ai/ becomes a lax vowel /a/ following the addition of the suffix ～ity to the stem divine, and in the latter the voiced alveolar stop /d/ in the stem elude surfaces as a voiceless alveolar fricative /s/ in the context of the suffix ～ive (Aronoff, 1976; Scalise, 1986). For theories of lexical processing such changes in the surface realisation of words raise an important question regarding the nature of internal representations. How do the different surface phonetic forms of such words relate to their respective internal representations?

Two general classes of answers have been proposed to this question. According to the first, all predictable phonological changes would be omitted from an abstract underspecified lexicon (e.g., Lahiri & Marslen-Wilson, 1991; Marslen-Wilson, Tyler, Waksler, & Older, 1994; Marslen-Wilson & Zhou, 1999). According to the second, phonological alternants of the same morpheme would correspond to different entities at some level of representation (Baayen, Dijkstra, & Schreuder, 1997; Järvikivi & Niemi, 2002; Schreuder & Baayen, 1995; Tsapkini, Kehayia, & Jarema, 1999). Currently it is difficult to make an empirically based choice between
the two views since both of them are able to accommodate most of the existing data on allomorphy. What we need then, is to build a broader basis for cross-linguistic comparisons with the view to constrain competing accounts of lexical processing and representation.

The present study is a preliminary step in this direction, focussing on allomorphic variation in Modern Standard Arabic (hereafter Arabic for short), a Semitic language where surface forms (e.g., [katama] hide) are usually analysed into a three consonantal root carrying semantic information (e.g., [ktm] hiding), and a word pattern conveying phonological and morpho-syntactic information (e.g., [faṣāla]1 perfective, active). There are a number of reasons for choosing to address this issue in the context of this language. First, it is on the basis of data from Semitic languages like Arabic (Boudelaa & Marslen-Wilson, 2000, 2001a), and Hebrew (Frost, Forster, & Deutsch, 1997) that the strongest case for abstract underlying morphological units has been secured. The consonantal root, which is an abstract bound morpheme that only surfaces as a discontinuous element interleaved with the word pattern, produces significant priming even when the prime and the target have an opaque semantic relationship (e.g., [katiibah]/[maktab] squadronoffice). Similarly, the word pattern, which is both abstract, bound, and discontinuous, gives rise to significant priming effects between the forms sharing it (e.g., [qaabala]/[saahama] meet–contribute). The question we begin to ask here is how abstract are the representations of these entities, and, specifically, whether they are able to abstract away from surface form variation. Do the internal representations of roots and word patterns abstract away from regular and predictable phonological changes, or do they feature a detailed and idiosyncratic description of each and every variant? Second, roots and word patterns play differential functions, with the former conveying semantic information, and the latter conveying phonological and morpho-syntactic information. This functional difference may be reflected in the way allomorphy affects the role played by the two units. Before describing the two cross-modal experiments we ran to address these issues, we present some key characteristics of Arabic morphology.

2. Allomorphic variation in Arabic: Background

There are two types of roots in Arabic (Holes, 1995; Wright, 1995). First, strong roots, like {qrb} getting closer, {ktm} hiding, {xtm} ending, which are so called because their three consonants systematically surface in every derived or inflected form. For example, the three consonants of the root {qrb} feature in every complex form in which this root is involved (e.g., [qarib] close, [qaraabah] closeness, [qaribah]2 relatives). The second type of Arabic root is known as the weak root. These make up about 10% of all roots (Mrayati, 1987), and typically involve the glides [w] or [y] or both as part of their component consonants (e.g., [nwm] sleeping, [wbd] uniting, [wfa] being successful). In some cases the glide appears in the surface phonetic form (e.g., [nawwama] cause to sleep, or [tāwahhada] unite), and no allomorphy occurs. In others, such as when the weak root is combined with a word pattern like [qitaṣala], the glide undergoes regressive assimilation from the word pattern consonant immediately following it, and an allomorphic form obtains.

Consider for instance the weak root {wfq} in the context of the word pattern [qitaṣala]. Here the glide [w] surfaces as [l] giving rise to the phonetic form [qitaṣala] agree. This kind of allomorphy preserves the CV-structure of the surface form; the underlying form [ḥwtafaqa] has the same number of consonants and vowels as the surface form [qitaṣala]. This contrasts with the disruption of the CV-structure when the glide of the root is allomorphically deleted. To exemplify, when the weak root {hyn} drawing closer, is interleaved with the word pattern [faṣāla], an intermediate form [hayqa]a obtain, and this will undergo a regular glide deletion process that will generate the surface form [haana] draw closer. Here the CV-structure of the surface form [haana] is disrupted in the sense that it has a different number of vowels and consonants from the underlying form [hayqa]a (Brame, 1970; Holes, 1995; McCarthy, 1979, 1981, 1982; Verssteegh, 1997; Wright, 1995).

Glide elision, which sparks off root and word pattern allomorphy, is a highly productive process in Arabic. It systematically applies when the glide is flanked by two homorganic vowels, as in [qawama], which surfaces as [qaama] stand up, or by a low vowel [a] and a high vowel /i/ as in [xawafa], which surfaces as [xaafa]a be frightened (Brame, 1970; Holes, 1995; McCarthy, 1981; Wright, 1995). In fact, it is so productive that the speakers of the language have come to use glides as default consonants when the Consonantal-Slots (C-slots) of the word pattern outnumber the consonants of the root (McCarthy & Prince, 1990). For example, in order to form the plural of the singular noun [jaamuus] buffalo, the three root consonants of this form must be mapped onto the C-slots of the word pattern [faṣāl], which contains four C-slots. The correct plural form of [jaamuus] is [jawaamiis] buffaloes, where the fourth consonantal-slot

---

1 The letters "f, q, l" are place holders for the first, second, and third consonant of the root. Thus in a verb form like [qariba] draw closer, the /l/ is usually referred to as the "l" of the verb, the /f/ as the "f" of the verb, and the /q/ as the "q" of the verb.

2 We underline the letters of the root where appropriate to highlight the difference between strong and weak roots.
is filled by the glide /w/. Even recent loan words undergo processes of glide elision or glide ephenthesis. The French loan word *cadre*, for example is pluralised as *[kwaadīr]* with an epenthesis glide.

Being productive and applicable to two functionally distinct morphemes, Arabic allomorphic variation poses an interesting range of questions about how the human language processor handles variability of linguistic input. Here we focus on allomorphic processes where an underlying glide surfaces in a different phonetic form depending on the local phonetic environment, as in cases such as *[ittīyaahun]*, *destination*, where the weak root {w3h} surfaces in the phonetic form {t3h}, in contrast to cases like *[waayaaha]* *confront* where there is no allomorphic change in the phonetic expression of the root. For both roots and word patterns, we ask whether these types of variation will affect the speed and effectiveness of access to the morphemic components of the word, as measured in the cross-modal priming task. We begin with the effects for root priming.

### 3. Experiment 1: Root priming and allomorphic variation

Experiment 1 draws on the cross-modal priming paradigm where participants hear a spoken prime immediately followed by a visual target to which they have to make a lexical decision (Marslen-Wilson et al., 1994). Previous research has established that cross-modal priming is sensitive to genuine morphological effects based on repeated access to the same morphemic entity. Moreover, in this paradigm pure phonological relatives like *[bulletin]*/bullet, *[tinseltin]* do not prime, and semantically highly related pairs such as *gavel/give* or *hid/hide* can also fail to prime (Boudelaa & Marslen-Wilson, 2001a; Marslen-Wilson, Hare, Older, & Ford, 1995; Marslen-Wilson et al., 1994).

In the current experiment we examined root priming under different conditions of allomorphic variation as outlined in Table 1.

In Condition 1, labeled [+Root +S – Allom], the prime and target share a root, are semantically related, and are not allomorphic. The three consonants of the root surface both in the prime and target (e.g., *[muʃaarikun]*/[ʃaaraaka]*

<table>
<thead>
<tr>
<th>1. [+Root +S – Allom]</th>
<th>Test</th>
<th>Baseline</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>[muʃaarikun]</td>
<td>[ʃaaraaka]</td>
<td>participant</td>
<td>participant</td>
</tr>
<tr>
<td>[muʃaqaubun]</td>
<td>[ʃaaraaka]</td>
<td>punished</td>
<td>participate</td>
</tr>
</tbody>
</table>

3. Note that we are glossing over the thorny issue of the directionality of the association between the radical consonants and the C-slots here. Does the association proceed from left to right such that the first consonant of the root {zms} is mapped onto the first C-slot of the pattern [ʃaʃalil]? Or does it proceed from right to left with the /s/ of {zms} associated with the last /l/ of [ʃaʃalil]? Alternatively, does it proceed by associating the two consonants at the edge of the root, that is the /s/ and the /l/, first and then by filling the rightmost C-slot with the remaining /m/, and finally by adding the default glide. The choice between these three possibilities is far from trivial (Bohas, 1997; McCarthy, 1981; Yip, 1988). Nonetheless it is orthogonal to the point we would like to make here, which concerns the systematic recourse to a glide whenever the consonantal material of the root is insufficient.

### Table 1

<table>
<thead>
<tr>
<th>Sample stimuli for Experiment 1, with Arabic script, IPA transcription, and English glosses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test</strong></td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>1. [+Root +S – Allom]</td>
</tr>
<tr>
<td>2. [Root +S + Allom]</td>
</tr>
<tr>
<td>3. [Root –S + Allom]</td>
</tr>
<tr>
<td>4. [Root –S + Phon]</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Participant/Participate, [muʃtaɾaman]/[ʔaʃtaɾama]* respectable/respect*. This condition functions as a priming baseline. In most of our previous investigation we have found significant priming among word pairs like these, and we expect to observe facilitation here. In Condition 2, [+Root +S + Allom], primes and probes also share a root, and a transparent semantic relationship. However the prime is allomorphic, surfacing with only two of the three root consonants intact, and with the third consonant undergoing phonological change. Thus, in the pair *[ittīfaaqun]*/ittiqaqabun, *agreement* and *[waafaqa]*/waafaqabun, *agree* based facilitatory effects are not seen in priming experiments, even though earlier experiments suggest that purely form-based facilitatory effects are not seen in priming experi-
ments in Arabic or Hebrew (Boudelaa & Marslen-Wilson, 2001b; Frost, Kugler, Deutsch, & Forster, 2001).

3.1. Method

3.1.1. Participants

We tested 32 volunteers aged 16–20. They were pupils at the High School of Tataouine in South Tunisia, and used Modern Standard Arabic on regular basis.

3.1.2. Material and design

Ninety-six familiar verbs which were 3–6 letters long and 3.5 syllables on average were chosen to serve as targets. They were divided into the four conditions described above with 24 probes per condition (see Table 1). The target words were orthographically unambiguous. The number of segments shared between primes and targets was 3.04 in Condition 1, 1.96 in Condition 2, 2.17 in Condition 3, and 2.21 in Condition 4.

In Condition 1, the prime and target shared a full root and a transparent semantic relationship. This contrasts with Condition 2 where the root was allomorphic in the prime and with Condition 3, where prime allomorphy was further compounded by an opaque semantic relationship between prime and probe. The transparency/opacity of the semantic relationship between pairs of critical materials was established in a semantic judgement pretest. Fifteen judges, who were native speakers of MSA, were asked to determine on a 1–9 scale (with 1 being unrelated and 9 highly related) how semantically related a set of 1200 word pairs were. Semantically transparent pairs as used in Conditions 1 and 2 received a rating of 6.5 on average while the semantically opaque pairs used in Condition 3 received an average of 3. Finally, Condition 4 provides a phonological baseline where prime and target share comparable vocalic and consonantal overlap as in the first three conditions. For each of the 96 test primes an unrelated baseline word matched as closely as possible on familiarity, number of letters, and syllables, and form class was selected. Familiarity was also determined in a pretest where another 15 judges rated words on a 1–5 scale (1 unfamiliar and 5 very familiar). All the words used in this and the next experiment had an average familiarity rating of 3.

A further 96 words were selected and paired with pseudo-word targets. Of these, 48 word–pseudo-word pairs shared three letters (e.g., [gabaadatun]*[ab3ada] worship), and 48 shared only two letters (e.g., [bukaα2u]*[bakat3a] weeping). These ensured that not all prime–target pairs which overlapped phonologically had a “YES” response. In order to increase the proportion of unrelated pairs in the experiment 48 word–word filler pairs that did not share any relation (e.g., [xuluudun]/[zahraqa] eternity/burn) were included along with a further 48 word–pseudo-word fillers (e.g., [nuhuudun]/ *[πnkafada] standing up). Additionally, 40 practice trial consisting of 20 word/word pairs and 20 word/non-word pairs were constructed in such a way as to mimic the experimental trial pairs. Two experimental list were constructed each containing 328 pairs of which 164 were yes responses and 164 no responses. Participants were assigned randomly to one of the two lists and did not encounter the same prime or target more than once.

3.1.3. Procedure

All the prime words were recorded by a native speaker of Arabic and digitised with a sampling rate of 44 kHz, then downsampled to 22 kHz using the CoolEdit program and stored on a portable PC. Three portable PC monitors were used to test subjects in threes in a quiet room. They heard the stimuli at a comfortable level through HD 250 Sennheiser headphones. The sequence of stimulus events within each trial started with a 1000 ms silence followed by an auditory prime. Immediately at the offset of the prime a visual target was displayed on the screen for 2000 ms. A new trial would start at the end of this period unless the subject responded within the time-out. Timing and response collection were controlled by a laptop PC running the DMDX package.4 Participants were instructed to make a lexical decision as quickly and as accurately as possible by pressing a “YES” or “NO” key. The “YES” response was always made by the dominant hand. To make sure that participants attended to the auditory prime they were asked at intervals to write down the prime word of one of the fillers trials. The experiment, which lasted about 20 min, started with the 40 practice trials followed by the rest of the stimuli.

3.2. Results

The data in this and the next experiment were inverse-transformed to reduce the influence of outliers (Ratcliff, 1993). Table 2 gives the percent error rates and the harmonic means of the reaction times in the four conditions. No subjects or items were rejected due to excessive error rates.

Separate three-way ANOVAs with participants (F1), and items (F2) treated as random variables were carried out on the reaction time (RT) and accuracy data. One factor was Condition with 4 levels, [+Root +S–Allom], [+Root +S +Allom], [+Root –S +Allom], and [–Root –S +Phon], the second was Prime Type with two levels Test Prime and Baseline prime. Condition was treated as a

---

4 The DMDX experimental software is programmed by Jonathan Forster at the University of Arizona. DMDX is a member of the DMASTR family of experimental software developed at Monash University and the University of Arizona by K.I. Forster and J.C. Forster. For further information see: http://www.u.arizona.edu/~kfortser/dmdx/dmdxhp.htm.
repeated factor in the participants’ analysis and as an unrepeated factor in the items analysis, while Prime Type was treated as a repeated factor in both analyses. The third variable was a dummy one representing either the participant grouping in the allocation of participants to experimental lists (for the participants analysis), or the test item grouping in the allocation of items to lists (in the items analysis). This variable, which was treated as a between subjects factor in the participants’ analysis and as a between items factor in the items’ analysis, was included only to reduce the estimate of random variation, and its effects will not be reported (Pollatsek & Well, 1995).

The ANOVA revealed a significant main effect of Condition \(F_1(3, 31) = 3.49, p < .003, F_2(3, 95) = 4.21, p < .008\) and Prime Type \(F_1(3, 31) = 5.29, p < .021, F_2(3, 95) = 6.04, p < .01\) in the participants and items analyses. We focus here on the outcome of the planned pairwise comparisons to assess differences between pairs of conditions directly. These comparisons, using Bonferroni-corrected protection levels (Keppel, 1982), showed that there were significant facilitatory effects in Condition 1, [+Root +S – Allom], \(F_1(1, 30) = 10.75, p < .003, F_2(1, 22) = 4.08, p < .05\). No significant effects were found in Condition 4, [–Root – S – Phon] with \(F_1 < 1, F_2 < 1\) despite a numerically sizeable interference effect. There were no differences between the first three conditions in terms of the magnitude of priming they generated (all \(F_1 < 1\), and all \(F_2 < 1\)). Furthermore, Conditions 1–3 each showed significantly more facilitation than Condition 4. For the [+Root +S – Allom] condition and the [–Root – S + Phon] condition; \(F_1(1, 30) = 6.58, p < .016, F_2(1, 47) = 4.40, p < .042\), for the [+Root +S + Allom] condition and the [–Root – S + Phon] condition; \(F_1(1, 30) = 9.09, p < .005, F_2(1, 47) = 6.90, p < .012\); and for the [–Root – S + Allom] and the [–Root – S + Phon] condition; \(F_1(1, 30) = 7.43, p < .011, F_2(1, 47) = 4.93, p < .032\). Similar ANOVA’s conducted on the accuracy data revealed no significant main effects or interactions.

3.3. Discussion

The interesting first result of this experiment is that a probe like [waafaqa] agree, where the root {wfq} sur-

Table 2

<table>
<thead>
<tr>
<th></th>
<th>Test</th>
<th>Control</th>
<th>Priming</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. [+Root +S – Allom]</td>
<td>560 (4.7)</td>
<td>583 (4.9)</td>
<td>23</td>
</tr>
<tr>
<td>2. [+Root +S + Allom]</td>
<td>544 (5.2)</td>
<td>578 (4.9)</td>
<td>34</td>
</tr>
<tr>
<td>3. [+Root – S + Allom]</td>
<td>567 (4.2)</td>
<td>598 (3.9)</td>
<td>31</td>
</tr>
<tr>
<td>4. [–Root – S + Phon]</td>
<td>615 (4.4)</td>
<td>581 (3.9)</td>
<td>–34</td>
</tr>
</tbody>
</table>

faces in full, is reliably facilitated by the prior presentation of a semantically transparent prime like [Zittifaqun] agreement, where the same root surfaces in an allomorphic form as /tfd/. This shows that surface phonological variation in the phonetic form of the root need not modulate the contribution of the consonantal component of Arabic words to the process of lexical access. The amount of priming is equally strong between morphologically related pairs that involve an allomorphic prime (Conditions 2 and 3) and those that do not (Condition 1).

The second important outcome is the significant priming in Condition 3 between word pairs such as [Zittigaalun]/[waaazaah] destination/confront. These have different surface roots, /tfd/ in the prime and /wfd/ in the target; different semantic interpretations, destination in the prime, confront in the target; and different syntactic properties, the prime being a noun and the target a verb. If each phonological variant of a morpheme is assumed to have a separate representation, priming in Condition 3 can be attributed neither to overlapping semantic/syntactic representations, nor to overlapping phonological representations, and is not predicted to occur. Alternatively, if regular morphological alternants are assumed to map onto internal representations that abstract away from surface form variation, priming may be expected. More specifically, if the glides /w/ and /fd/, which trigger allomorphic variation in Arabic and act as default consonants (McCarthy & Prince, 1990), are assumed to be underlyingly specified only for the feature [-Syllabic], then any other consonant of the language may potentially be used to fill their C-slot. From the perceptual point of view, this means that the underlying morpheme [wfd] will be as well matched by the surface realisation /tfd/ as by /wfd/, since all that the C-slot allocated to the glide of the underlying root requires is a segment with the specification [-Syllabic]. In view of the characteristics of the word pairs used in Condition 3, Arabic root priming seems to be driven by repeated access to an underspecified abstract morpho-phonological entity, rather than by repeated access to overlapping semantico-syntactic representations (De Jong, Schreuder, & Baayen, 2000). The importance of the shared underlying morphological unit is reinforced by the results from Condition 4, where phonological overlap per se does not generate any cross-modal facilitation at all.

4. Experiment 2: Word pattern priming and allomorphic variation

The aim of this experiment is to evaluate whether the internal representations of word patterns abstract away from regular surface form changes in the same way as roots. As mentioned earlier, word patterns and roots are functionally different morphemes, the former convey
information about the phonological structure of the word and its morpho-syntactic meaning, while the latter express its semantic meaning. The differential functional properties of the word pattern are captured in the linguistic description of this morpheme as consisting of CV-Skeleton and a Vocalic Melody (McCarthy, 1981, 1982). For example, the pattern {faɦala} is factored into the CV-Skeleton {CVCCCV}, which conveys a causative meaning and determines the phonological structure of the overall surface form, and into the vocalic melody [a-a], with a perfective, active meaning. In an earlier research (Boudelaa & Marslen-Wilson, 2004) we found that the CV-skeleton was the cognitively most salient component of the word-pattern. Thus, although allomorphic variation in the surface phonetic expression of an underlying segment remains the focus of this second experiment, maintaining comparability with Experiment 1, we also include a further contrast allowing a preliminary evaluation of the effects of allomorphic variation which disrupts the CV-skeleton, through the deletion of an underlying segment.

In Condition 1, labeled [+WP−Allom], primes and targets like [pɪntạra]/[pɪhtạmala] spread/endure are used (see Table 3). The word pattern {pɪtaðala} combines with a strong root {ɪnfr} and {ʊml} in the prime and the target given here, and no allomorphy obtains. We have consistently found facilitation between such pairs in a variety of priming tasks (Boudelaa & Marslen-Wilson, 2000, 2001b, 2001c). In Condition 2, [+WP+Allom1], prime and target pairs again share a word pattern. However here the word pattern is combined with a weak root in the prime (e.g., {wðd} in [pɪttạhạda] unite, or {wñm} in [pɪttạhamala] accuse) but with a strong root in the target (e.g., {bsm} in [pɪbtạsama] smile, or {bðl} in [pɪbạtạda] demean). The glide consonant of the root in all the prime words of this condition is assimilated to the consonant of the word pattern resulting in primes that are allomorphic overall but with preserved CV-structures. These are the same manipulations as in Conditions 2 and 3 of Experiment 1, and we expect to see the same results here. The word pattern is an intrinsically abstract entity, that by definition abstracts away from the phonetic properties of whatever consonants fill the specified consonants slots in its CV structure.

Since prime–target pairs that share a word pattern in this condition are typically not semantically related, we do not have the same contrast between pairs that vary in semantic relation. Instead, we include as Condition 3, labelled [+Root+Allom2], a condition where allomorphic processes disrupt CV-structure. The shared word pattern is combined with weak roots both in primes (e.g., {ðn} in [ʔaðnạa] praise, [ʔɪmạhạa] be erased) and targets (e.g., [ʔala]a cancel, [ʔɪnạlaa] become clear). In each case the glide is elided, so that the underlying segment does not appear in the surface phonetic form. This disrupts the CV-structures of both prime and target—e.g. the underlying form [ʔɑl邑a] appears as [ʔaḷɪa]a with a missing consonantal slot. Given the evidence that the CV-structure is salient in lexical processing (Boudelaa & Marslen-Wilson, 2001a), and given the evidence from work in Hebrew, also looking at word pattern priming when the CV-structure is disrupted (Frost, Deutsch, & Forster, 2000), it is possible that priming will be disrupted here. On the other hand, forms like [ʔaðnạa] praise seem perfectly intelligible and unmarked in normal language use, suggesting that root and word pattern information can successfully be extracted, even if there is surface disruption of the CV-structure.

In Condition 4, [+Phon], primes and targets share a form overlap (e.g., [ʔaʋarạtụn]/[ʔaamaːla] raid/be amiable, [ʔa axrạtụn]/[kallạfa] boulder/entrust) that mimics the phonological overlap underlying the test items sharing a word pattern. For example, in the pair [ʔaʋarạtụn]/[ʔaamaːla] raid/be amiable, the vowel sequences [aa]–[a] are shared in prime and target, but without any morphological overlap. This was included in order to evaluate the possible interpretation of word pattern effects in terms of simple form overlap.

### 4.1. Method

#### 4.1.1. Participants

Another group of 32 volunteers aged 16–20 were tested. They were pupils at the High School of Tataouine in South Tunisia, and used Modern Standard Arabic on regular basis.

#### 4.1.2. Material and design

There were 24 prime/target pairs in each of the four conditions outlined in Table 3, giving a total of 96 test pairs. The targets were 3–6 letters long and 3.5 syllables on average, and were pretested for familiarity. The number of segments shared between primes and targets

### Table 3

<table>
<thead>
<tr>
<th>Test</th>
<th>Baseline</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. [+WP−Allom]</td>
<td>[pɪntạra] spread</td>
<td>[pɪhtạmala] endure</td>
</tr>
<tr>
<td>2. [+WP+Allom1]</td>
<td>[pɪttạhạda] unite</td>
<td>[pɪbtạsama] smile</td>
</tr>
<tr>
<td>3. [+WP+Allom2]</td>
<td>[ʔaʁnạa] praise</td>
<td>[ʔalạa]a dismiss</td>
</tr>
<tr>
<td>4. [Phon]</td>
<td>[ʔaʋarạtụn] raid</td>
<td>[ʔaamaːla] be amiable</td>
</tr>
</tbody>
</table>
was 3.29 in Condition 1, 3.96 in Condition 2, 2.33 in Condition 3, and 1.92 in Condition 4.

Condition 1, [+WP–Allom], can be viewed as a paradigm check, with the prime and target sharing a non-allomorphic word pattern. Condition 2, [+WP +Allom1] combines prime words where the overall surface form is allomorphic but the CV-structure is preserved with non-allomorphic targets. Condition 3, [+WP +Allom2], combines primes and targets where the CV-structure of the word pattern is disrupted. Finally, Condition 4 provides a phonological baseline where prime and target share a form overlap that is comparable to that shared between primes and target in the first three conditions. Primes and targets in the four conditions were semantically unrelated, scoring 2 or less on the same 9-point scale used to evaluate the critical pairs in Experiment 1. For each of the 96 test primes an unrelated baseline word matched as closely as possible on the same pretest devised to construct the materials of Experiment 1. For each of the 96 test primes an unrelated baseline word matched as closely as possible for each condition.

A further 96 words were selected and paired with pseudo-word targets. These word/pseudo-word pairs shared a phonological overlap that mimicked the overlap between word/word pairs (e.g., [sudun]/*[tasada] worship). This was aimed at discouraging participants from responding “YES” to all prime–target pairs which overlapped phonologically. The proportion of related pairs in the experiment was diluted by including a further 48 word–word filler pairs that did not share any relation (e.g., [burruzun]/*[fanfaqa] emergencel spend), along with a further 48 word–pseudo-word fillers (e.g., [sududun]/*[intasada] climbing). Additionally, 40 item practice trials consisting of 20 word/word pairs and 20 word/non-word pairs were constructed in such a way as to mimic the experimental trial pairs. Two experimental lists were constructed each containing 328 pairs of which 164 were yes responses and 164 no responses. Participants were assigned randomly to one of the two lists and did not encounter the same prime or target more than once.

4.1.3. Procedure
This was the same as in Experiment 1.

4.2. Results

No participant or item was dropped as a result of excessive error rates. Harmonic means, and mean error rates for the four categories of test and baseline items are shown in Table 4, along with the amount of priming for each condition.

Similar analyses of variance as in Experiment 1 were conducted on the inverse data with Condition (4 levels) and Prime Type (2 levels) as main variables, and List (2 levels) as a dummy variable. There was a significant effect of Condition \(F_1(3, 31) = 7.24, p < .001, F_2(3, 95) = 6.02, p < .001\), but only a trend for Prime Type \(F_1(3, 31) = 3.39, p < .075\), \(F_2(3, 95) = 3.20, p < .077\). As in Experiment 1, we ran planned comparisons using Bonferroni-corrected protection levels to directly compare facilitation between pairs of conditions. Priming was significant in the [+WP–Allom] condition \(F_1(1, 30) = 4.14, p < .05, F_2(1, 22) = 4.70, p < .04\), and the [+WP +Allom1] condition \(F_1(1, 30) = 5.19, p < .03, F_2(1, 22) = 7.77, p < .01\), but not in the [+WP +Allom2] condition \(F_1 < 1, F_2 < 1\), or the [+Phon] condition \(F_1(1, 30) = 3.75, p < .06, F_2(1, 22) = 2.08, p < .16\). Furthermore, the amount of priming did not differ significantly between the [+WP–Allom] condition and the [+WP +Allom1] condition showing that CV-structure-preserving allomorphy does not disrupt word pattern priming. By contrast, the 34 ms priming in the [+WP–Allom] condition was significantly different from both the -4 ms of the [+WP +Allom2] condition \(F_1(1, 30) = 4.11, p < .050, F_2(1, 47) = 3.31, p < .05\), and the -21 ms of the [+Phon] condition \(F_1(1, 30) = 10.16, p < .003, F_2(1, 47) = 3.02, p < .05\). The 25 ms facilitation observed in the [+WP +Allom1] condition was also significantly different from the amount of priming observed in the [+WP +Allom2] condition \(F_1(1, 30) = 8.54, p < .007, F_2(1, 47) = 4.54, p < .05\), and the [+Phon] condition \(F_1(1, 30) = 11.73, p < .002, F_2(1, 47) = 8.39, p < .006\). There was no difference in priming between the [+WP + Allom2] and the [+Phon] conditions \(F_1(1, 30) = 1.90, p < .17, F_2(1, 47) = 1.3, p < .28\). The error analysis showed no effects of Condition, Prime Type or their interaction.

4.3. Discussion
The main part of Experiment 2 is consistent with the results of Experiment 1, showing that allomorphy that changes the surface realisation of a given underlying segment (but without altering its CV-structure) does not affect the processing role played by this unit, at least as reflected in the priming effects obtained. In Condition 2, just as in Conditions 2 and 3 in Experiment 1, significant priming is obtained. However, when the CV-structure of the word pattern is disrupted (Condition 3) following glide deletion, prime and target pairs sharing a word pattern fail to facilitate each other. The absence of priming among phonologically related words (Condi-
tion 4), rules out an explanation of word pattern priming in terms of form overlap. In fact such an overlap, whether defined in terms of vowels as in Experiment 2, or in terms of consonants as in Experiment 1, generates interference in cross-modal priming.5

Priming by word pattern in Conditions 1 and 2, and the lack of facilitation among words sharing a nonstructural form overlap (Condition 4) is good evidence that the task used here, cross-modal priming, involves access to underlying morphologically decomposable levels of representation. Hearing a prime word where the phonological structure of the word pattern is preserved results in access to the underlying word pattern morpheme and to the phonological and morpho-syntactic information associated with it. When a target containing the same word pattern morpheme is subsequently presented, the decision about the lexical status of the surface form in which it partakes is speeded up because some aspects of its underlying representation have already been accessed during the processing of the prime. The failure of priming in Condition 3 is of interest in this respect, since it suggests that access to components of internal representations of primes and targets has been disrupted. This is consistent with data from Hebrew where word pattern priming was also disrupted when CV-structure was disrupted (Frost et al., 2000). However, these results for Arabic can only be treated as preliminary, and further experiments are needed to probe in more detail the source of the effects in Condition 3. For the remainder of this paper we will focus on the case where allomorphy changes the surface phonetic form of a segment, but where CV-structure is preserved.

5. General discussion

This research addresses the question of how morphemic alternants are represented and accessed in the lexicon. In Experiment 1, we compared priming among words sharing a non-allomorphic root and a transparent semantic relationship with priming among words sharing an allomorphic root and either a transparent (Condition 2) or an opaque (Condition 3) semantics. Facilitation was equally strong in the three conditions compared to a fourth condition where primes and targets overlapped in form. In Experiment 2, we evaluated priming among prime/probe pairs that shared either a non-allomorphic word pattern, an allomorphic word pattern where the CV-structure is preserved, or an allomorphic word pattern where the CV-structure was disrupted. Only when the underlying CV-structure was preserved did word pattern priming emerge. Words sharing simple form overlap again failed to prime. Taken together these results have some interesting implications for existing theoretical accounts of lexical processing and representation, to which we now turn.

5.1. Theoretical implications

We consider the implications of our results for the Direct Access Model (DAM hereafter) (Marslen-Wilson et al., 1994; Marslen-Wilson & Zhou, 1999) and the Morphological Race Model (MRM hereafter) (Baayen et al., 1997; Schreuder & Baayen, 1995).

According to DAM, stems are stored in the same lexical entry regardless of surface form variation, and phonologically regular alternants map directly onto the same lexical entry, which is a unitary triplet of abstract semantic, phonological, and syntactic information. For example, the underlying phonological representation of a morpheme like {sane} is assumed to be something like /sŒn/ where the symbol /Œ/ stands for a vowel unspecified for tenseness. Being unspecified for this feature, the vowel /Œ/ will be a good match not only for the tense vowel /ei/ in sane, but also for the lax vowel /æ/ in sanity. In the context of Arabic root allomorphy, a DAM-like view would imply that allomorphic variants such as /tfq/ and /wfq/ in [ittifaqun]/[waafaqa] agreement/agree, should prime each other because both forms map onto the same underlying lexical entry of {wfq} as illustrated in Fig. 1.

However, the priming pattern among surface forms where root allomorphy is compounded by semantic opacity as in [ittijaahun]/[waajaha] destination/confront is less straightforward to accommodate on a DAM-like account. On this view morphological priming is the result of repeated access to the same lexical entry, defined as a triplet of semantic, syntactic, and phonological information. Accordingly items with different semantic or syntactic attributes should have distinct lexical entries.
and so should items whose form differences are not the result of regular phonological processes. Since the root \( \{w3\} \) in the prime [\( \texttt{ittifaaqun} \)] has a semantic interpretation that is markedly different from what it has in the context of the target [\( \texttt{waafaqa} \)], it should belong to two distinct lexical entries. Depending on further assumptions about the properties of this architecture, cases like [\( \texttt{ittifaaqun}/\texttt{waafaqa} \)], should either show no priming at all, or significantly reduced priming compared to semantically transparent pairs like [\( \texttt{ittifaaqun}/\texttt{waafaqa} \)] agreement/agree. Even if both lexical entries were activated by the root—perhaps analogous to the activation of the two sets of meanings of a homophone in English like bark—one would expect activation of the entry corresponding to the target to be weaker, therefore reducing the priming effect. This is not what we see here, but the implications of this are hard to evaluate without additional research.

In contrast to the results of Experiment 1, the word pattern priming profile found in Experiment 2 is compatible with DAM, which would attribute the significant facilitation in Conditions 1 and 2 to the preserved underlying phonological structure shared by primes and targets. The DAM has to assume that each word pattern has some form of lexical entry, specifying its phonological and morpho-syntactic properties, and that as long as the same word pattern morpheme is evoked in the analysis of prime and target, then priming should occur. Interestingly, our earlier research (Boudelaa & Marslen-Wilson, 2000) shows that homophonous word patterns do not prime (these are pairs where two phonetically identical word patterns correspond to two different morphemes). The lack of priming in Condition 3, for primes and targets such as [\( \texttt{a\donaal}/\texttt{a\lyalaa} \)] praise/cancel whose CV-structure is disrupted, could be attributed to differences in phonological structure between the surface word pattern [\( \texttt{a\donaal} \)] in [\( \texttt{a\donaal} \)] and the underlying word pattern \( \texttt{a\lyalaa} \) in [\( \texttt{a\lyalaa} \)]. Here it may not be possible to compensate for allomorphic variation because of the crucial role of CV-structure in conveying information about the phonological structure of the surface form.

In the class of models represented here by MRM, lexical representations consist of three layers: (a) form-based modality specific access representations, or lexemes, (b) an intermediate modality independent integration level called lemmas, and (c) a central semantico-syntactic level of representation. Each of these levels activates the one immediately above it and receives feedback from it. Lexical access takes place through a full form route where a complex lexeme like sanity maps onto its associated lemma, which in turn maps onto the relevant semantic and syntactic representations. Parallel to this full form access is a parsing route where the component morphemes of a complex form are activated, then checked to ensure that their subcategorisation features are compatible, and finally their meaning is computed. On this account, Arabic allomorphic forms such as in [\( \texttt{ittifaaqun}/\texttt{waafaqa} \)] agreement/agree, would have separate access representations, separate lemmas, and will only share aspects of their semantico-syntactic representations as shown in Fig. 2.

This architecture is in principle compatible with the priming observed in Conditions 1 and 2 of Experiment 1. In Condition 2, an allomorphic prime like [\( \texttt{ittifaaqun} \)] agreement facilitates a related non-allomorphic target like [\( \texttt{waafaqa} \)] agree despite their having different access representations and different lemmas because the two forms would have overlapping syntactic-semantico-semantic representations (Schreuder & Baayen, 1995). However, priming should be significantly reduced, relative to Condition 1, if present at all. The

![Diagram](image-url)

**Fig. 2.** A mediated race model of lexical processing as applied to Arabic forms sharing a transparent semantic relationship.
MRM approach assumes that there is no bottom-up activation of the target [waafaqa] by the prime [ittitfaaqun] as each of them has its own access representation and its own lemma. Since feedback to access and lemma representations is assumed to be proportional to the activation of the access representation involved in the first place, the facilitation of [waafaqa] by the alloformic [ittitfaaqun] in Condition 2 should be less strong than that of [jaaraka] by the non-allomorphic [mujaarikun] in Condition 1. Instead we see robust priming in both conditions and no significant difference between them. The pattern of priming observed in Condition 3 among allomorphic forms which share an opaque semantic relationship like [ittiZaaahun]/[waazaZaha] destination/confront is also problematic for this model. Such forms would have not only different access and lemma representations, but also non-overlapping semantic and syntactic representations as depicted in Fig. 3. Indeed, as currently stated, the model appears to provide no basis for priming, either via whole word or decompositional access routes.

The outcome of the second experiment is less problematic for an MRM approach. The word pairs in Conditions 1 and 2 prime each other because they share a word pattern which has the same access representation, the same lemma, and the same morpho-syntactic meaning. For example, despite surface form allomorphy, the pair [ittiZaaahun]/[waazaZaha] destination/confront is also problematic for this model. Such forms would have not only different access and lemma representations, but also non-overlapping semantic and syntactic representations as depicted in Fig. 3. Indeed, as currently stated, the model appears to provide no basis for priming, either via whole word or decompositional access routes.

The present research provides a new cross-linguistic data point on the way the human language processor handles alternants of the same morpheme. The kind of root and word pattern priming reported here suggest that the language processor can tolerate surface phonological changes that depart considerably from underlying representations provided that these changes are regular and predictable. Furthermore, the results are consistent with an Arabic mental lexicon where the crucial unit governing the lexical access process is a morphemic entity that is combinatorially called upon to build different surface forms. Two alternative traditional approaches to lexical architecture, an underspecification approach (DAM), and a dual mechanism approach (MRM), appear to be at odds with the data. It is not clear how either model can be modified to accommodate the present results without substantial architectural changes—for example by adding another representational layer in the case of MRM, or abandoning the view of the lexical entry as a unitary triplet of abstract semantic, phonological, and syntactic information in the case of DAM. It remains to be seen whether more recent distributed approaches to lexical architecture will prove more adequate in accommodating data from typologically different languages (Gaskell & Marslen-Wilson, 1997; Gaskell, Hare, & Marslen-Wilson, 1995; Plaut & Gonnerman, 2000).

Acknowledgments

We would like to thank the members of the language group at the MRC-Cognition and Brain Sciences Unit for their help, and in particular Matt Davis and Mike...
Ford. We also wish to thank Mr. Abdallah Megbi, Headmaster of the High School of Tataouine, Tunisia for his generous help in allowing us access to participants for testing.

References