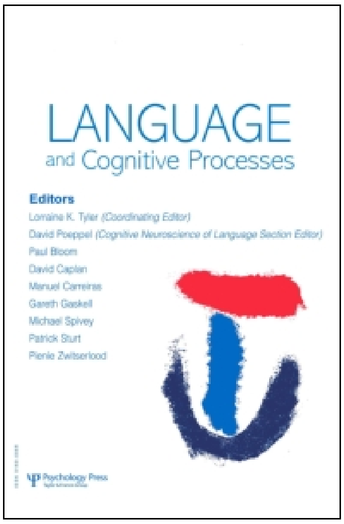

Language and Cognitive Processes
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713683153

Productivity and priming: Morphemic decomposition in Arabic

Sami Boudelaa[ab]; William D. Marslen-Wilson[b]
[a] Department of Linguistics, United Arab Emirates University, Al-Ain, UAE [b] MRC Cognition and Brain Sciences Unit, Cambridge, UK

Online publication date: 12 May 2011

LANGUAGE AND COGNITIVE PROCESSES
2011, 26 (4/5/6), 624–652

Psychology Press
Taylor & Francis Group

Productivity and priming: Morphemic decomposition in Arabic

Sami Boudelaa[1,2] and William D. Marslen-Wilson[2]
[1]*Department of Linguistics, United Arab Emirates University, Al-Ain, UAE*
[2]*MRC Cognition and Brain Sciences Unit, Cambridge, UK*

Word formation in Arabic involves the interleaving of two abstract morphemes—a root consisting exclusively of consonants and conveying semantic meaning, and a word pattern comprised primarily of vowels and conveying phonological and morpho-syntactic information. In masked and cross-modal priming experiments, we probed the processing relationship between these two morphemes during word recognition by examining the roles of word pattern and root productivity (family size) in producing word-pattern priming in Arabic deverbal nouns. Co-varying word pattern and root productivity in a 2 × 2 design, we found that priming was determined entirely by the productivity of the root. Even very productive word patterns did not prime if they appeared in the context of an unproductive root. This pattern of results, which is identical in cross-modal and masked priming, indicates the importance of the root in driving the on-line decomposition of Arabic surface forms into their constituent morphemes.

Keywords: Morphological productivity; Arabic patterns; Roots; Priming; Obligatory decomposition.

A key feature of lexical systems across the world's languages is the role of morphological structure. Most surface words in most languages are

Correspondence should be addressed to Sami Boudelaa, Linguistics Department, Faculty of Humanity & Social Sciences, United Arab Emirates University, PO Box 71117, A1 Ain, UAE. E-mail: s.boudelaa@uaeu.ac.ae

This work was supported by the Medical Research Council UK (U.1055.04.002.00001.01). The authors would like to thank Abdallah Megbli, Headmaster of the High School of Tataouine, Tunisia, for providing access to the participants who took part in the study. We thank Ian Nimmo-Smith, and other members of the Language Group at the MRC Cognition and Brain Sciences Unit, for their help. We are also thankful to the anonymous reviewers for helpful comments on earlier versions of this work.

http://www.psypress.com/lcp DOI: 10.1080/01690965.2010.521022

morphologically complex, being made up of two or more underlying morphemes—typically a stem and one or more inflectional or derivational morphemes, as in the English forms *walks* ({walk} + {-s}), *politeness* ({polite} + {ness}), and *disagreeable* ({dis-} + {agree} + {-able}). The existence of this complexity poses basic questions about the nature of lexical representation in the cognitive and neural systems underlying human language function, and about the manner in which such complex forms are analysed in processes of language comprehension and production. In recent research, we have focused on Arabic, which has generated particularly clear-cut evidence for the psychological reality of morphological structure in the lexical representation of complex forms (e.g., Boudelaa & Marslen-Wilson, 2001, 2004, 2005). In the research reported here, we extend this research to investigate the process of access to these representations, so that we can begin to specify the mechanisms whereby a complex linguistic form can be analysed on-line in order to unpack its underlying morphemic organisation.

It is widely accepted that the phonetic, or surface, word in Semitic languages such as Arabic and Hebrew is made up of at least two abstract underlying morphemes, the word pattern and the root (e.g., Glinert, 1996; Hilaal, 1990; Holes, 1995; Idrissi, Prunet, & Béland, 2008; Versteegh, 1997; Wright, 1995; though see Berent, Vaknin, & Marcus, 2007; Ratcliffe, 2004; Ussishkin, 1999, 2005, for alternative stem-based views). The root is exclusively consonantal (e.g., in Arabic {frq}, {dðr}, and in Hebrew {drx}, {khl}), while the word pattern consists of vowels and a subset of consonants (e.g., {faʕlatun}, {mafʕalun} in Arabic and {Hifʕil} in Hebrew), where the letters "fʕl" are place holders for the first, second, and third root consonants. Functionally, the root conveys the general semantic meaning, which will be more or less consistent across the surface forms featuring that root. Thus the general meaning of entering, inherent in the Arabic root {dxl}, is expressed to various degrees in most of the derivatives containing this root (e.g., [daxala] enter, [ʔadxala] insert, [ʔidxaalun] insertion). Similarly, the place noun, singular, and masculine reading characteristic of the Arabic nominal word pattern {mafʕalun} is present in the forms involving this pattern (e.g., [marqadun], place where one sleeps, place noun, singular, masculine; [maʃrabun] drinking place, place noun, singular, masculine). Verbal word patterns, which combine with roots to form verbal rather than nominal forms, operate in a similar manner to create surface forms—thus the Arabic verbal word pattern [ʔafʕal], with the meaning active, perfective, causative, will combine with the same root {dxl} to give the form [ʔadxala], meaning insert.

Although both roots and word patterns are linguistically abstract underlying morphemes, in the sense that they never appear individually as surface phonetic forms but only in combination with each other, there is substantial evidence for their active role in the perception and production of words in Arabic and Hebrew. This evidence, we should note, presents apparently insoluble problems for alternative stem-based approaches which dispense with the root and the word pattern as the building blocks of Semitic words (Berent et al., 2007; Ratcliffe, 2004; Ussishkin, 1999, 2005). On this view, Hebrew words like "gidel" he raised, and "gadal" he grew are related, not because they share a root, but because "gidel" is productively formed by modifying the stem "gadal" (Ussishkin, 2005). Forms like "gadal" and "gamal", however, which are morphologically related by virtue of sharing the word pattern {paℵal}, would nonetheless be classed as unrelated because they are not derived from each other. This claim is in conflict with the strong psycholinguistic evidence, summarised below, that words sharing word patterns, just as much as those sharing roots, can be treated by Arabic and Hebrew listeners as being strongly related.

Words sharing a root will prime each other effectively regardless of whether their semantic relationship is transparent (e.g., [kitaabatun]/[kaatibun] writing/writer) or opaque (e.g., [kitaabatun]/[kaatibun] squadron/writer), and they do so in a variety of priming paradigms including masked priming, cross-modal priming, and auditory–auditory priming (Boudelaa & Marslen-Wilson, 2000, 2001). In Arabic, root priming occurs even when the prime and target belong to differential syntactic categories, suggesting that this morpheme functions as an abstract organising unit of the lexicon. Similar effects are found in Hebrew (Frost, Deutsch, Gilboa, Tannenbaum, & Marslen-Wilson, 2000; Frost, Forster, & Deutsch, 1997). This apparent salience and independence of the root is buttressed by neuropsychological data showing selective impairment of production processes involving root consonants (Berg & Abd-Al-Jawad, 1996; Prunet, Béland, & Idrissi, 2000).

Where the word pattern is concerned, both Arabic and Hebrew exhibit similar properties for verbal word patterns—that is, word patterns that combine with a consonantal root to produce a verb form. Arabic verbs sharing a word pattern (e.g., [ʔaʕlama]/[ʔaxraʒa] cause to know/cause to go out, sharing the causative word pattern [ʔafʕal]) prime each other significantly across a variety of priming paradigms (Boudelaa & Marslen-Wilson, 2000, 2001, 2005). Similar masked and cross-modal facilitation is found among Hebrew verbs sharing a word pattern (e.g., [huklat]/[hugdar] was recorded/was defined) (Deutsch, Frost, & Forster, 1998; Frost et al., 2000). This behavioural evidence is corroborated by Hebrew neuropsychological data suggesting selective impairment of processes involving word patterns (Barkai, 1980).

Nominal word patterns, where the word pattern combines with a root to form a noun, seem to behave quite differently in the two languages. Research in Arabic, using a similar variety of priming techniques, shows that priming by nominal word patterns is as robust as priming by verbal word patterns (Boudelaa & Marslen-Wilson, 2000, 2001, 2005). The Hebrew data on the use of noun word patterns are quite different. Earlier results from segment shifting experiments, where participants are instructed to segment and shift a designated word pattern from a source word onto a target word and to name the new result as quickly as possible, indicated that nominal word patterns may be functional units in Hebrew processing (Feldman, Frost, & Pnini, 1995). However, more recent priming studies by Frost and collaborators (1997, 2000) show a consistent lack of priming between Hebrew nouns sharing a word pattern. In masked as well as cross-modal priming, Hebrew pairs sharing a nominal word pattern (e.g., [taklit]/[targil] record/exercise) do not prime. This demonstrates a substantial cross-linguistic difference in the organisation of the Arabic and Hebrew mental lexicons and suggests that caution is needed in linking data across the two languages.

Focusing here on Arabic, the clearest evidence for the processing separability of root and word-pattern morphemes comes from a further masked priming experiment examining effects at different Stimulus Onset Asynchronies (SOAs) (Boudelaa & Marslen-Wilson, 2005), bearing in mind (see Marslen-Wilson, Bozic, & Randall, 2008) that this is a task that taps primarily into an early but morphologically sensitive process that decomposes visual inputs into potential constituent morphemes. It does not—at least as far data from English and related languages indicate (Marslen-Wilson et al., 2008)—reflect the properties of higher level lexical and semantic representations. In the Boudelaa and Marslen-Wilson (2005) experiment, SOA (the delay between the onset of the masked prime and the onset of the target) was varied between 32 and 80 ms, with strikingly different effects over SOAs for root and word-pattern priming.[1] Root priming—where prime and target share the same root—was present at the shortest SOA and continued at much the same level across SOAs. This held both for pairs like [duxuul] entering and [daxal] enter, where the meaning of the root is transparent and shared by both prime and target (+ R + S), and for pairs like [miℵt^{ℵ}af] coat and [ℵat^{ℵ}f] compassion, which also share the same root, but where the meaning of the root is different in the two cases (+ R − S). This is consistent with earlier findings for Semitic languages that effects of morphological relatedness are obtained irrespective of the semantic relatedness of prime and target, across a variety of priming tasks.

Word-pattern priming showed a quite different pattern across SOAs with neither verbal nor nominal word patterns showing significant priming at the shortest SOA (32 ms). Both types of word pattern showed significant priming at the next SOA (48 ms), but this was only transient. For verbal word patterns, no priming was detected at either of the later SOAs (64 and 80 ms), while for nominal word patterns, priming was still present at 64 ms, but

[1] Note that our experience of masked priming using Arabic script shows that even at SOAs of 80 ms the prime still seems strongly masked and not detectable by the participants (Boudelaa & Marslen-Wilson, 2005).

not at 80 ms. This is quite different from the root-related pairs, where priming is both more immediate and is maintained at a consistent level across SOAs.

This contrast in the temporal signature of root and word-pattern priming is not only persuasive evidence for their separability as effective cognitive entities, but also raises basic questions about the nature of the analysis processes underlying their identification and segmentation during visual word recognition. What are the mechanisms whereby the reader extracts information about the root, primarily carried by consonants? How does this relate to the extraction of information about the word pattern, primarily carried by vowels (though also by consonants in some cases)?

Clearly, one important factor in the experiments reported above, which almost all use tasks like masked or cross-modal priming with visually presented targets, is the nature of Arabic orthography. Arabic has a primarily consonantal script, where consonants are always specified in the written form, but where most vowels are only specified by the use of special diacritics, and where these diacritics are omitted in almost all forms of Arabic encountered in daily life—they only occur in materials for children learning to read and in religious texts. This is equivalent to the pointed/unpointed contrast seen for the Hebrew script.

The consequence of this for Arabic morphology is that the root is always fully specified in the orthographic form, while word patterns are at most only partially specified—as when the word pattern contains a consonant or when it contains the type of long vowel that is represented as a full grapheme in the orthography and not as a diacritic (for example, the word pattern [fuℵuul] containing the long vowel/u:/, expressed as the letter "uu" in forms like [duxuul] entering). This absence of vowel information in the surface form leads to the ambiguity of many written words in Arabic, where the same form (e.g., سلم specifying the root {slm}) is consistent with several different words in which this root is combined with different word patterns (as in [silm] peace, [sullam] ladder, [salim] be safe).

In the experiments we have conducted in Arabic, these ambiguities were avoided by only using word patterns containing one or more long vowels. Since these were marked in the default form of the script, this meant that although the word pattern was never fully specified in the visual prime or target, the word was still rendered unambiguous. This did not, however, remove the intrinsic difference between the root and the word pattern in Arabic orthography—that the root is fully and immediately given (even if other, word-pattern-related consonants may also be present), while the appropriate word pattern must in some way be inferred from the constraints provided by the root and by whatever components of the word pattern are actually present in the surface orthography.

This differential dependency may be reflected in the Arabic incremental masked priming experiment described above (Boudelaa & Marslen-Wilson, 2005), where root priming is detectable at a shorter SOA than word-pattern priming. It also suggests that the properties of the root—in particular, its extractability during processing—will interact with the properties of the word pattern to determine the timing with which the word pattern is itself extracted. By exploring such interactions, we can obtain a better purchase on the nature of the operations involved in the access of morphological information from written Arabic. As a first step in this direction, we focus here on the productivity of roots and word patterns to examine how variations in the properties of each affect the timing of access to word patterns.

PRODUCTIVITY AND MORPHOLOGICAL PROCESSING

The term "productivity" is widely used in linguistics and psycholinguistics to refer to a variety of different distributional properties of words and morphemes. Here we focus on a widely used definition of productivity in quantitative terms, as the frequency with which a particular morpheme is involved in the processes of word formation. This type of definition has become prominent under the label "family size" (Baayen, Lieber, & Schreuder, 1997; Schreuder & Baayen, 1997), a type count of the number of word forms that incorporate a particular stem (such as *dark*), either by derivation (as in *darkness*) or by compounding (as in *darkroom*).

This stem-based measure of productivity has been shown to affect lexical processing—chiefly in single word visual lexical decision tasks—across several languages.[2] For example, Dutch noun stems with a high family size (e.g., *rijst* rice)—appearing as constituents in a large number of complex forms—elicit higher subjective frequency ratings and shorter latencies in visual lexical decision than noun stems with smaller family sizes (e.g., *rund* cow) (Baayen et al., 1997). Turning to more linguistically complex words, a Dutch form like *rookte* smoked, made up of the productive stem *rook* and the past tense suffix -*te*, is responded to more quickly and more accurately than an otherwise matched complex form like *krijste* shrieked, with the less productive stem *krijs* but the same inflectional suffix -*te* (Bertram, Baayen, & Schreuder, 2000; De Jong, Schreuder, & Baayen, 2000). Similar results are reported for other languages, including English (e.g., Ford, Marslen-Wilson, & Davis, 2003) and Hebrew (Moscoso del Prado Martin, Deutsch, Frost,

[2] Family size will normally correlate with both lemma and word frequency. In all of the studies reporting family size effects, this is in the context of regression analyses where the effects of frequency have been partialled out.

Schreuder, De Jong, & Baayen, 2005). In the latter study, family size computations are based on the consonantal root.

Grammatical morphemes, such as affixes, have generally not been studied in the family size context, although they also vary along the same dimension—the agentive {-*er*} in *driver*, for example, is highly productive, occurring in more than 1,100 complex forms, contrasting with a suffix like {-*ent*}, as in *correspondent*, which occurs in less than 80 forms in a standard lexical database (Baayen, Piepenbrock, & van Rinj, 1993). Although family size per se has not been looked at for affixes, there is some evidence for processing effects of their distributional properties. In Italian, for example, the number of word types in which a letter sequence (such as *ri*- or *co*-) occurs as a real prefix as opposed to a pseudo-prefix affects the speed and accuracy with which nonwords are responded to in a lexical decision task (Laudanna, Burani, & Cermele, 1994). A complementary pattern of results is found in English, where word pairs sharing a productive bound affix (e.g., happiness/darkness) facilitate each other reliably in cross-modal priming, while those sharing a nonproductive affix do not (Marslen-Wilson, Ford, Older, & Zhou, 1996).[3]

In addition to this general evidence for effects of productivity on lexical processing, there are also some indications that productivity in Semitic languages does affect whether or not word-pattern priming is observed. One argument for this comes from the absence of word-pattern priming in the subset of the Arabic vocabulary known as the "primitive nouns" (Boudelaa & Marslen-Wilson, 2001). These are a closed set of around 100 concrete nouns, referring to natural categories such as body parts and animal names, that stand apart from the much more numerous and dominant verbal and deverbal morphology, and which are notable for the low productivity of their roots and word patterns. Most primitive noun roots participate in at most two or three different forms and the word patterns are equally unproductive. This contrasts with the studies in Arabic where nominal word-pattern priming has been observed (e.g., Boudelaa & Marslen-Wilson, 2001, 2005), in which only productive word patterns (PWPs) were used together with productive roots (PRs).

Taken together, these various strands of research suggest that variations in morphemic productivity do affect lexical access processes, and that this is a variable that is effective in nonconcatenative morphologies such as Arabic as well as in concatenative morphologies such as English or Dutch. In the experiments reported here, we aim to systematise these potential effects during lexical access and decomposition by co-varying the productivity of

[3] Note that this study used a different definition of productivity, defined as the continuing use of the affix in novel word formation. This is not the same as, though correlated with, a simple type count.

the Arabic root with the productivity of the word pattern, and focusing on the effects of these on word-pattern priming, examined initially in the same masked priming task as used in other recent studies. Of particular relevance here is a study by Forster and Azuma (2000), showing that masked priming between pairs sharing a bound stem (e.g., *permit/submit*, *reflect/inflect*) depends on the productivity of that stem. Low family size stems such as {-*vive*}, which only occur in the two words *survive* and *revive*, did not show priming. Although bound stems in English are linguistically quite different to roots and word patterns in Arabic, this result nonetheless suggests that variations in productivity do play a role in the decompositional analysis of morphologically complex words.

For Arabic our incremental-masked priming results (Boudelaa & Marslen-Wilson, 2005), with word-pattern priming emerging later than root priming (at 48 rather than 32 ms SOAs), are consistent with a processing dependence between the extraction of the Arabic root and the extraction of the Arabic word pattern. One possibility is that the shortest SOA does not allow the decompositional processing of the prime to progress sufficiently for the extraction of the word pattern to have reached a point at which it becomes available to affect parallel extraction processes in the target. Extraction of the root, in contrast, seems to progress sufficiently rapidly in both prime and target to allow root-priming effects to be observed at the earliest SOA. Note that the transience of word-pattern priming across SOAs in Boudelaa and Marslen-Wilson (2005) suggests that these effects are highly sensitive to the relative timing of different access processes. This in turn means that if we vary root productivity, and test at the 48 ms SOA at which word-pattern priming was first elicited in the earlier study, then priming should be reduced for low productivity roots—in the SOA study, we used generally PRs (mean family size 18.3, range 7–35) and word patterns (mean family size 452, range 79–967). By the same token, since the early parsing processes tapped into by masked priming must operate on information simultaneously provided about potential roots and word patterns, we should also expect effects of word-pattern productivity, on the assumption that the access and representation of these grammatical morphemes is also sensitive to type frequency.

EXPERIMENT 1: MASKED PRIMING OF NOMINAL WORD PATTERNS

This experiment seeks to determine whether priming between deverbal noun word patterns is modulated by their own productivity, the productivity of the root with which they co-occur, or by the productivity of both these units. To do this, we used the masked priming task (Forster & Davis, 1984; though see

Norris & Kinoshita, 2008), which, as indicated above, has been used extensively to probe early stages in the recognition of visually presented morphologically complex words in Arabic and Hebrew as well as in many Indo-European languages (Boudelaa & Marslen-Wilson, 2005; Forster & Azuma, 2000; Frost et al., 1997; Longtin, Segui, & Halle, 2003; Rastle, Davis, Marslen-Wilson, & Tyler, 2000).

To examine these questions, we will pursue two investigative strategies. The first, using standard ANOVA-based analyses, will orthogonally co-vary root and word-pattern productivity over a set of four conditions, together with a fifth orthographic control condition. In each of the first four conditions, the prime and target systematically share the phonological structure and the morpho-syntactic meaning of a deverbal nominal word pattern without sharing a root. This shared word pattern is itself either productive or unproductive and occurs in the context of either a productive or an unproductive root (UR). The contrasts in overall priming effects between these conditions will provide general information about the dependencies between roots and word patterns in the early stages of visual word recognition.

The second, more exploratory strategy will use multi-level regression techniques to exploit the variations in lexical properties between primes and targets to build up a more differentiated picture of how the properties of prime and target family sizes interact to drive the lexical access process in a context like masked priming where the two inputs are expected to overlap. Since primes and targets in this experiment will have different roots, with accordingly different distributions of family sizes, we can ask whether the properties of the response (overall RT and amount of priming) are better accounted for by variations in prime root family size or in target root family size (or by both). To the extent, for example, that prime root family size accounts for variance in the word-pattern priming effect observed in responses to the target, then this will provide novel evidence about the manner in which the processing of a prime word (here hypothesised to have complex internal structure) can interface with the processing of a similarly complex target word.

Turning to the specific experimental design (see Table 1), in Condition 1 (labelled [PWP/PR]), the prime and the target share a PWP that occurs in the context of a PR, as in [taħkiimun]/[tatˤbiiqun] arbitration/implementation; [qaaℵidun]/[Dℵaalimun] seated/oppressor). We expect priming here on the basis of previous findings of strong word pattern effects between surface forms based on PRs and word patterns (e.g., Boudelaa & Marslen-Wilson, 2000). Condition 2, [UWP/PR], consists of prime and target pairs sharing an unproductive word pattern (UWP) that occurs in the context of PRs (e.g., [taℵaawuniyyun]/[tafaadℵuliyyun] cooperative/differential; [ʔutˤruuħaatun]/[ʔuʢʒuubatun] dissertation/marvel). Assuming that the properties of the root

TABLE 1
Experimental conditions with example stimulus set

	Prime		
	Test	*Baseline*	*Target*
1. +PWP+PR	[taħkiimun] arbitration	[ʕibratun] needle	[tatˤbiiqun] implementation
2. −PWP+PR	[taʕaawuniyyun] co-operative	[ʒaarun] neighbour	[tafaadˤuliyyun] differential
3. +PWP −PR	[saʒdatun] prostration	[sˤadiiqun] friend	[lakmatun] blow
4. −PWP −PR	[maʔsawiyyun] dramatic	[muʒaahidun] fighter	[maʕdaniyyun] mineral
5. +Phon	[taʕℵbiʔatun] mobilisation	[musˤaffaħatun] armoured vehicle	[taʕabun] fatigue

environment provides the appropriate conditions for priming, as in Condition 1, this tests whether the distributional properties of the word pattern also modulate priming effects. In Condition 3, [PWP/UR], the contrast goes in the opposite direction, with prime and target sharing a PWP but in the context of an UR, as in [saʒdatun]/[lakmatun] prostration/blow; [ħariisˤun]/[ʕamiiqun] eager/deep. Priming among such pairs would indicate that nominal word pattern effects are primarily contingent on their own properties and do not need to be associated with a PR. Condition 4, [UWP/UR], is made up of prime and target pairs sharing an UWP that occurs in the context of an UR, as in [maʔsawiyyun]/[maʕdaniyyun] tragic/mineral; [mutˤribatun]/[muʕʒizatun] female singer/miracle. This should provide the most unfavourable conditions for priming. Finally Condition 5, [+ Form], is an orthographic control condition where the primes and targets share 2–3 letters (e.g., [taʕbiʔatun]/[taʕabun][4] mobilisation/fatigue; [ʒaarun]/[ʒaariyatun] neighbour/maid), but do not share either a root or a word pattern. This is to evaluate the effects of form relatedness per se in the absence of any morphological relationship between prime and target.

A critical issue for this set of contrasts is an appropriate measurement and definition of productivity (type frequency) for the two types of morphemes involved. To determine basic productivity, we used a 12,000-item data base, developed in Cambridge, which lists all the attested surface forms featuring each of the 1,000 most frequent roots of Arabic (Khouloughli, 1992).

[4] Note that the final ∼un at the end of each word is the indefinite article which is part of the phonological representation of the word but has no corresponding graphemes in the written form.

It was immediately apparent from this and other sources (e.g., Abdah, 1979, Boudelaa & Marslen-Wilson, 2009) that there are major distributional differences between roots and word patterns, reminiscent of function/content word differences in languages like English. Although there are more than 6,000 different roots attested in Modern Standard Arabic (MSA), the average nominal family size is relatively low (12 for this sample of 1,000), with maxima in the range 30–40. To achieve satisfactory contrasts between conditions, we chose PRs with an average family size of 21 (overall range 16–37) for target words and of 23.65 (range 16–37.5) for prime words. The UR conditions had an average family size of 9.12 (range 3–14.5) for target words and of 9.45 (range 4–14.5) for prime words (see Table 2). It was not possible to use lower productivity items because of the need to match word patterns across prime and target pairs.

Nominal word patterns are differently distributed to root morphemes, reflecting their very different functions in the language. Although there are many fewer word patterns than roots (155 in our sample), their productivity is much higher (averaging 60 in our sample with a range of 1–434), analogous to the much greater productivity of grammatical morphemes in Indo-European languages. In numerical terms, this meant that we could not fully match the differences in productivity for roots and word patterns. Low productivity word patterns—with the same word pattern occurring in both prime and target—could be kept broadly comparable, with an average family size of 17 (range 3–56). High productivity sets had a much larger family size, averaging 331.5 nominal surface forms (range 200–434). This appropriately reflected the distributional properties of word pattern family sizes and ensured that we had a robust contrast along this dimension.

Method

Participants

We tested 54 volunteers aged 16–20. They were pupils at the High School of Tataouine in South Tunisia and used MSA on regular basis. All had normal or corrected to normal vision.

Materials and design

Two counterbalanced lists of materials were constructed. Targets paired with a related prime in the first list were paired with an unrelated prime in the second and vice versa. Twenty-six participants were randomly assigned to receive the first list and 28 to the second. There were 24 targets in each of the five conditions shown in Table 1. All the primes and targets were deverbal nouns. Each was made up of a root and a word pattern that was either productive or unproductive, as described above. On average, 40% of the

TABLE 2
Stimulus properties for the targets, test primes, and baseline primes in the five conditions (standard deviations in parentheses). Semantic relatedness is between prime and target

	N syllables	*N* letters	*Familiarity*	*Total root family size*	*Total WP family size*	*Semantic relatedness*	
1. +PWP+PR	Target	3.00 (0.00)	4.42 (0.50)	3.84 (0.99)	23.0	340.2	
	Test Prime	3.00 (0.00)	4.42 (0.50)	3.93 (0.97)	23.9	340.2	1.54 (0.78)
	Baseline	3.92 (0.83)	5.46 (1.06)	3.23 (1.20)	14.2	124.7	1.04 (0.20)
2. −PWP+PR	Target	4.50 (0.72)	5.88 (0.80)	2.87 (1.39)	20.0	10.7	
	Test Prime	4.50 (0.72)	5.88 (0.80)	3.12 (1.34)	23.4	10.7	1.29 (0.46)
	Baseline	3.58 (0.72)	4.71 (0.69)	3.71 (1.12)	17.7	182.9	1.25 (0.61)
3. +PWP −PR	Target	3.00 (0.00)	4.33 (0.48)	3.07 (1.43)	8.2	322.9	
	Test Prime	3.00 (0.00)	4.33 (0.48)	3.14 (1.36)	8.5	322.9	1.33 (0.76)
	Baseline	3.58 (0.83)	4.79 (0.72)	3.54 (1.09)	17.4	157.3	1.08 (0.28)
4. −PWP-PR	Target	4.08 (0.72)	5.21 (0.72)	2.78 (1.55)	10.1	22.5	
	Test Prime	4.08 (0.72)	5.21 (0.72)	3.28 (1.22)	10.7	22.5	1.04 (0.20)
	Baseline	3.46 (0.83)	4.88 (0.95)	3.27 (1.48)	21.2	244.8	1.08 (0.28)
5. +Phon	Target	3.67 (0.56)	4.75 (0.85)	3.31 (1.24)	14.7	82.0	
	Test Prime	3.63 (0.77)	4.63 (0.88)	3.75 (1.07)	16.0	185.9	1.38 (0.58)
	Baseline	3.54 (0.83)	4.75 (0.85)	3.11 (1.48)	11.9	172.4	1.33 (0.70)

word-pattern phonemes were orthographically present in the primes and targets—for example, for the surface form [taħkiimun] arbitration, containing the word pattern {tafʕiilun}, only the phonemes /t/ and /ii/ were captured orthographically. We used 55 different word patterns in total. The full list of stimuli can be accessed here.[5] To provide a baseline against which to measure priming in the five conditions, each target was paired with a randomly chosen word that was used as a related prime in a different condition. On average, targets and related primes shared 1.5, 2.96, 1.63, 2.33, and 3.13 letters in Conditions 1–5, respectively. The overlap in letters between baseline primes and targets was much lower throughout, averaging 0.38, 0.83, 0.67, 0.96, and 0.5 in Conditions 1–5, respectively. The average length in letters and syllables, as well as the average familiarity of the experimental materials are also given in Table 2.

Corresponding to these 120 word–word prime targets, a further 120 prime words were paired with orthographically legal nonword targets. These nonwords were formed by changes to an existing form (e.g., [ħayaatun] life changed into *[ħaʒaafun]), or [falakun] astronomy transformed into *[falkaʒatun]). The amount of form overlap between the word/nonword pairs matched as closely as possibly the experimental word/word pairs. A set of 20 practice trials, consisting of 10 word/word responses and 10 word/nonword responses, were constructed to have similar properties to the experimental trials.

Procedure

All primes were presented in 24-point traditional Arabic font size for 48 ms. They were preceded by a 500 ms forward mask consisting of 28 vertical lines, in a 30-point size using the same font. This mask was chosen on the basis of pretesting sessions, where it was found to be more effective than the standard hash marks mask. The prime was immediately followed by a target word or nonword written in a 34-point font size without diacritics. Targets were displayed until participants responded or 2,000 ms had elapsed. Participants were advised that they would be seeing a series of letter strings presented one at a time, and that they would be required to decide as quickly and accurately as possible whether each string was an Arabic word or not.

Timing, stimulus display, and data collection were controlled by three laptop PC's running the DMDX package (Forster & Forster, 2003), so that up to three participants could be tested simultaneously. The inter-trial interval was 1,000 ms. There was no reaction time (RT) or accuracy feedback. The experiment started with the practice trials followed by the

[5] http://www.mrc-cbu.cam.ac.uk/people/sami.boudelaa/Boudelaa_Marslen-Wilson_LCP_Annex%281%29.pdf

TABLE 3
Harmonic mean reaction times in millisecond (standard deviations) and percentage of error rates in each condition in Experiment 1

	Test		*Baseline*	
	Mean RT	*% Error*	*RT*	*% Error*
1. +PWP+PR	565 (39.55)	2.52	596 (47.52)	2.20
2. −PWP+PR	597 (52.28)	5.82	630 (58.95)	5.50
3. +PWP −PR	604 (55.62)	3.46	597 (71.38)	3.93
4. −PWP −PR	630 (51.68)	4.40	633 (48.51)	4.01
5. +Phon	607 (47.80)	3.38	612 (54.60)	3.46

experimental items. There were two breaks in the test session; one after the practice session and the other halfway through the main test sequence. The experiment lasted about 15 minutes.

Results

Data from one participant were deleted from the analyses because his response accuracy fell below 50%. No items were rejected. The data were inverse transformed to reduce the influence of outliers (Ratcliff, 1993). Table 3 gives the per cent error rates and the harmonic means of the RTs in the five conditions. Priming effects by condition are shown in Figure 1.

We analysed the data in two ways, starting with standard ANOVAs to evaluate the basic factorial contrasts built into the experiment. The second set of analyses used multi-level regression analyses to separate out the specific roles of variations in prime root family size and in target root family size. In addition, by treating family size as a continuous variable (although

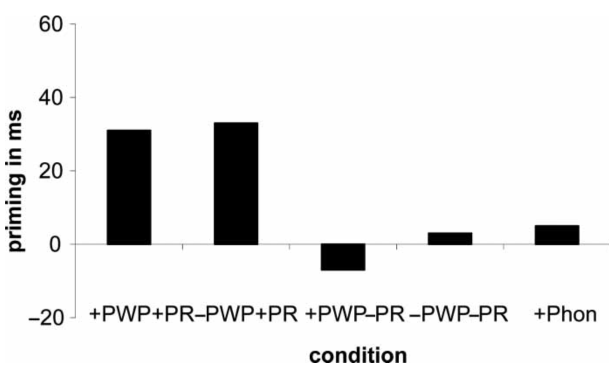

Figure 1. Masked priming effects across conditions in Experiment 1.
[a]Effects significant by subjects and items

the design to some extent dichotomises it) in the context of these powerful analytic techniques, we can provide more rigorous tests of whether or not word pattern and root productivity differ in their processing consequences.

Factorial analyses

In the first set of analyses, separate three-way ANOVAs with participants (F_1) and items (F_2) treated as random variables were carried out on the RT and accuracy data with the factor of Condition (with five levels) and of Prime Type (with the two levels of Test and Baseline). Condition was treated as a 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), and is not reported below (Pollatsek & Well, 1995).

There were significant main effects of Condition [$F_1(4, 51) = 13.98$, $p < .000$; $F_2(4, 110) = 15.96$, $p < .00$] and Prime Type [$F_1(1, 51) = 9.70$, $p < .003$; $F_2(1, 110) = 5.13$, $p < .02$], and an interaction between them [$F_1(4, 51) = 3.80$, $p < .005$; $F_2(4, 110) = 3.71$, $p < .034$]. Planned comparisons using Bonferroni-corrected protection levels (Keppel, 1982) showed that there were significant facilitatory effects only in Conditions 1, [PWP/PR], [$F_1(1, 51) = 10.94$, $p < .002$; $F_2(1, 22) = 18.87$, $p < .000$] and 2, [UWP/PR], [$F_1(1, 51) = 8.43$, $p < .005$; $F_2(1, 22) = 4.90$, $p < .035$]. These two PR conditions each showed significantly more priming than either of the UR conditions or the orthographic overlap condition.

To evaluate more directly the relative effects of productivity for different types of morpheme, we ran a further two-way ANOVA just on Conditions 1–4, with the factors Root Productivity (high, low) and Word Pattern Productivity (high, low), based on the priming scores by item and by subject. There was a significant main effect of Root Productivity [$F_1(1, 51) = 12.27$, $p < .001$; $F_2(1, 22) = 8.38$, $p < .008$], neither an effect of Word Pattern Productivity [$F_1 < 1$; $F_2(1, 22) = 1.38$, $p = .44$] nor any interaction between them ($F_1 < 1$; $F_2 < 1$).

Linear mixed-effect analyses

In this second set of analyses,[6] using linear mixed-effect techniques, we fitted different models with log RT as the dependent variable and with

[6] Note that what we are reporting in this and the next experiment are partial effects of significant predictors. This means that the effects of a given predictor are evaluated given (a) that other predictors are in the model and (b) that they are held constant.

participants, prime type, and target items as crossed random effects (Baayen, Davidson, & Bates, 2008; Baayen, Tweedie, & Schreuder, 2002; Bates, 2005; Bates & Sarkar, 2005). The effects of word pattern and root productivity were treated as continuous variables for both the prime word and the target word. Word pattern family size (the same for target and for prime) had a facilitatory effect on overall RTs ($\beta = -0.03460$; $t(4816) = -1.99$; $p < .05$) with faster responses for targets with larger families. However, this factor did not affect the size of the priming effect with no interaction with Prime Type (consistent with the outcome of the earlier factorial analyses).

The effects were different for root family size. The family size of the prime root had no effect on overall RT but did interact with Prime Type ($\beta = 0.04454$; $t(4816) = 2.82$; $p < .005$). The larger the family size of the prime word root, the more facilitation for primed targets. We see the opposite pattern for the root family size of the target word. This did significantly speed overall RT, with faster latencies as morphological family size increased ($\beta = -0.03707$; $t(4816) = -2.96$; $p < .004$). However, it did not modulate the magnitude of priming, showing no interaction with Prime Type ($\beta = 0.1682$; $t(4816) = 1.09$; NS). It is the properties of root processing in the prime, not the target, that determine whether the presence of a shared word pattern in prime and target leads to priming under the conditions of masked priming.[7]

The accuracy data, analysed using ANOVAs similar to those conducted on the latency data, revealed no significant effects.

Discussion

These results confirm, first, that Arabic nominal word patterns do prime when they occur in the context of a PR. This is consistent with earlier findings (e.g., Boudelaa & Marslen-Wilson, 2000) and suggests that the functional role played by nominal word patterns in Arabic is comparable to that played by verbal word patterns. Second, word-pattern priming seems to be contingent not on the productivity of the pattern itself, but on the

[7] Given the results reported by Moscoso del Prado Martin et al., (2005) for Hebrew, in separate analyses we also assessed the effects of root family size broken down into transparent and opaque family members. There were significant facilitatory effects on overall RT of target transparent family size ($\beta = -0.02908$; $t(4816) = -2.77$; $p < .006$) and of the transparent/opaque ratio for the target ($\beta = -0.07273$; $t(4816) = -3.33$; $p < .001$). In both cases, neither there was effect on RT associated with the prime root family size properties, nor was there any sign of Martin et al.'s (2005) finding for Hebrew that opaque family members were associated with an increase in overall RT. In the analyses of the effects of these variables on the amount of priming, evaluated in terms of interactions with Prime Type, we see small facilitatory effects of both prime and target opaque family size (respectively, $\beta = 0.01720$; $t(4816) = 2.22$; $p < .03$; $\beta = 0.01293$; $t(4816) = 1.935$; $p = .05$). There was also an effect for the target ratio ($\beta = -0.07273$; $t(4816) = 3.329$; $p = .001$) but not the prime ratio.

productivity of the root with which it is combined. Strong facilitatory effects are observed between primes and targets sharing a less PWP provided this occurs in the context of a PR. Conversely, when the context root is less productive, no word-pattern priming occurs regardless of the level of productivity of the word pattern involved.

In interpreting this absence of priming when the root is less productive, it is implausible that this is because the root is not extracted under these processing conditions, so that the absence of word-pattern priming simply reflects a more general absence of decompositional processing for these stimuli. In parallel studies using similar sets of stimuli (contrasting root family sizes of 9 and 26), we see no effects of low family size on root priming (Boudelaa & Marslen-Wilson, 2009). For masked priming, we obtain priming effects on lexical decision of 51 ms and 65 ms, respectively, for large and small family size conditions when prime and target share the same semantically transparent root. The results are equally strong when the root is morphologically identical in prime and target but has a different meaning (51 ms and 50 ms for large and small family sizes, respectively). Orthographic control conditions show no effects. We see the same results for cross-modal priming using these stimuli, as well as in a different experiment using primitive nouns, which also have small family sizes (Boudelaa & Marslen-Wilson, 2009). All of these results are consistent with the view that root extraction is a robust process occurring irrespective of variations in family size.

Further insight into the lexical access processes underlying masked priming is provided by the differential effects of prime and target-related variables. The important result is the contrast between root family size measures for the prime word and for the target word, where we see effects of prime root family size but not of target root family size on the amount of priming. This suggests, consistent with the SOA effects seen in our earlier research (Boudelaa & Marslen-Wilson, 2005), that the facilitatory effects on RT of having the same word pattern in prime and target depend on the timing with which information about the word pattern in the prime becomes available.

If this information is not available within a specific time window during the processing of the target word, then priming does not seem to occur. In the earlier experiment (Boudelaa & Marslen-Wilson, 2005), we observed this effect at very short SOAs (32 ms), where prime word pattern information was not yet available at the critical point in the processing of the target word, despite the productivity of the roots involved. In the current experiment, we seem to achieve the same result at a longer SOA (48 ms), by using prime roots with small family sizes. This slows down the decomposition of the prime word (carrying mainly orthographic information about the root), so that prime word pattern information does not arrive in time to affect the relevant aspects of processing of the target. The fact that these effects are strictly limited to the properties of the root in the prime word (and not in the target word), and that the family size of the word pattern seems to play no role, is strong evidence for the primacy of root extraction in orthographic processing of the Arabic words used here, and of the dependency of word pattern extraction on root analysis, at least in the early and highly automatised stages of visual word recognition that the masked priming task taps into (Marslen-Wilson et al., 2008).

In Experiment 2, we ask how general these dependencies are, or whether they are in some way a consequence of the special conditions of Arabic orthography, in the context of the masked priming task. To do this, we present the same primes auditorily, in a cross-modal auditory–visual priming task (Marslen-Wilson, Tyler, Walker, & Older, 1994). This means that not only is the prime no longer masked, but also that the word pattern is fully specified in the prime surface form—in contrast to written forms which contain only partial information about the word pattern. If the effect of root productivity on word-pattern priming in Experiment 1 is because this underspecification of the word pattern requires it to be inferred in the context of processes of root identification, then the full auditory specification of the word pattern should diminish or even eliminate these effects. If, however, the basic processes of morphemic segmentation in Arabic are organised around the root regardless of input modality, then root productivity effects on word-pattern priming should be the same across both experimental situations.

EXPERIMENT 2: CROSS-MODAL PRIMING OF NOMINAL WORD PATTERNS

This experiment asks whether a fully specified word-pattern prime, presented at a longer SOA, will still only generate priming in the context of a PR. Previous studies using auditory–visual cross-modal priming, both in Arabic and Hebrew (e.g., Boudelaa & Marslen-Wilson, 2004; Frost et al., 2000), have reported word-pattern priming—although in Hebrew (Frost et al., 2000) this was only obtained for verbal word patterns. None of this research, however, manipulated productivity variables, either for roots or word patterns.

Method

Participants

We tested 44 participants from the same age group and linguistic background as those in Experiment 1. None of them reported any history of speech or hearing disorder.

Materials and design

These were the same as in Experiment 1, save for adding 80 unrelated word/word pairs and 80 unrelated word/nonword pairs in order to bring the proportion of relatedness down to 30%, and to obscure the relationships in the test items. A further unrelated 20 pairs, of which were 10 word/word responses and 10 word/nonword responses, were also included. The design did not include a semantic control condition because of the absence of a semantic relationship between the test pairs. In a semantic judgement pretest, the pairs sharing a word pattern received an average rating of 1.30 on a scale ranging from 1 (completely unrelated) to 9 (highly related).

Participants were asked at pseudo-randomly distributed intervals during the course of the experiment to write down the spoken prime word of these pairs in order to make sure they were attending to it. The test items were divided into two balanced versions with every prime and target appearing once in each version.

Procedure

All primes were recorded in a sound attenuated booth by a native speaker of Arabic. They were digitised at a sampling rate of 44 kHz, then down-sampled to 22 kHz using the CoolEdit program. Each trial consisted of a 1,000 ms silence followed by an auditory prime. Immediately at its offset a visual target was displayed for 2,000 ms. A new trial started as soon as the subject responded even if 2,000 ms had not elapsed. Stimuli were presented at a comfortable level through HD 250 Sennheiser headphones. Participants were tested in groups of three (each working at a separate laptop), and had to make a speeded lexical decision to the visual target. Every other aspect of the procedure was the same as in Experiment 1.

Results

No participants or items were rejected. The data were inverse transformed to reduce the influence of outliers (Ratcliff, 1993). Table 4 gives the per cent error rates and the harmonic means of the RTs in the five conditions. Priming effects by condition are shown in Figure 2. The data were analysed using both ANOVA and multi-level regression techniques.

Factorial analyses

Harmonic means calculated over participants and items in each condition were entered into three-way by-participants and by-items ANOVA (as in Experiment 1) with the variables Condition (five levels), Prime Type (two levels), and List (two levels). There was a clear main effect of Condition [$F_1(4, 42) = 66.58$, $p < .000$; $F_2(4, 110) = 30.93$, $p < .00$] and Prime Type

TABLE 4
Harmonic mean reaction times in millisecond (standard deviations) and percentage of error rates in each condition in Experiment 2

	Test		*Baseline*	
	Mean RT	*% Error*	*RT*	*% Error*
1. +PWP+PR	542 (45.14)	3.03	575 (76.72)	1.70
2. −PWP+PR	608 (71.75)	4.92	643 (90.12)	4.73
3. +PWP −PR	600 (85.30)	5.30	611 (89.48)	5.02
4. −PWP −PR	627 (86.69)	5.30	636 (82.44)	5.30
5. +Phon	596 (77.45)	3.22	599 (74.01)	2.84

[$F_1(1, 42) = 25.39$, $p < .000$; $F_2(1, 110) = 15.04$, $p < .000$] by subjects and items. However, these two factors did not interact significantly [$F_1(4, 42) = 2.03$, $p = .092$; $F_2(4, 110) = 1.80$, $p = .138$]. A further two-way ANOVA just on Conditions 1–4, using priming scores with the factors Root Productivity (high, low) and Word Pattern Productivity (high, low), showed a significant main effect of Root Productivity [$F_1(1, 42) = 7.28$, $p < .05$; $F_2(1, 22) = 5.35$, $p < .05$]. Word Pattern Productivity did not have a significant effect ($F_1 < 1$; $F_2 < 1$) and did not interact with Root Productivity ($F_1 < 1$; $F_2 < 1$).

Following the same planned comparison approach as before, we also evaluated levels of priming within and between each condition. Priming was significant only in Condition 1, [+PWP+PR], [$F_1(1, 42) = 8.81$, $p < .005$; $F_2(1, 22) = 18.99$, $p < .000$] and in Condition 2, [−PWP+PR], [$F_1(1, 42) = 8.81$, $p < .005$; $F_2(1, 22) = 18.99$, $p < .000$]. Priming in Condition 1 was stronger than in Condition 5, [+PWP+PR], [$F_1(1, 42) = 5.49$, $p < .024$; $F_2(1, 22) = 4.51$, $p < .045$] and in Condition 2 than in Condition 5 [$F_1(1, 42) = 8.64$, $p < .005$; $F_2(1, 22) = 4.44$, $p < .047$]. None of the remaining pairwise comparisons were significant (all Fs < 1). Standard by-subject and

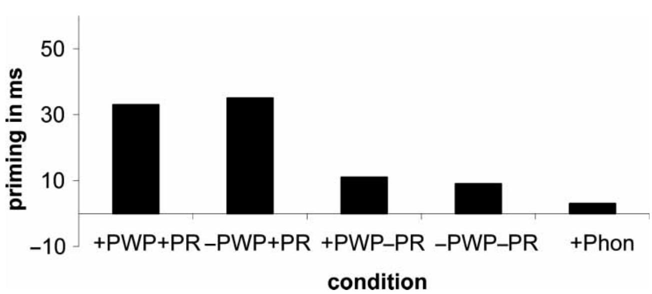

Figure 2. Cross-modal priming effects across conditions in Experiment 2.
[a]Effects significant by subjects and items

by-item analyses of the error scores revealed a main effect only for Condition [$F_1(4, 42) = 4.53$, $p = .002$; $F_2(4, 110) = 3.27$, $p = .014$].

Linear mixed-effect analyses

As before, we also submitted the data to a linear mixed-effects analysis, with log RT as the dependent variable and participants, prime type and target items as crossed random effects, and treating word pattern and root family size as continuous variables. Word pattern family size (the same for prime and target) again had a facilitatory effect on overall RTs ($\beta = -0.02493$; $t(3921) = -2.126$; $p < .05$). As before, this variable did not affect the size of the priming effect, showing no interaction with Prime Type (consistent with the factorial analyses). In contrast with Experiment 1, however, the prime root family size had no effect on performance (either RT or priming). Only the target root family size showed significant effects, facilitating overall RTs ($\beta = -0.06414$; $t(3921) = -3.231$; $p < .001$). In addition, and unlike Experiment 1, we see signs of an interaction between target root family size and Prime Type ($\beta = -0.03747$; $t(3921) = 1.859$; $p = .063$), although the effect is only marginally significant.

In a further set of analyses, we replaced the total root family size measures, covering all forms sharing the same root, with a different family size measure designed to better reflect the temporal dynamics of the spoken word recognition process as it unfolds over time, along the lines suggested by Baayen, Wurm, and Aycock (2007) and Wurm, Ernestus, Schreuder, and Baayen (2006). This measure includes only those morphological family members that match a given auditory prime from onset to offset. We refer to this as the Cohort Morphological Family size (CMF). For instance, if the auditory prime is [kitaab] book, its CMF will include words like [kitaabatun] writing and [kitaabiyyun] written because they feature the same root morpheme {ktb} and match the prime from onset to offset. In contrast, the word [maktab] office is not part of the CMF of the prime [kitaab], although it also contains the root {ktb}, because its onset rules it out as a CMF member.

Re-running the analyses using the CMF measure of prime word family size, we now see reliable facilitatory effects on overall RTs ($\beta = -0.08856$; $t(3921) = -3.485$; $p = .002$) and, critically, a significant interaction with Prime Type ($\beta = -0.07756$; $t(3921) = 2.2412$; $p = .02$), showing that prime word family size, measured in this way, indeed modulates word-pattern priming in responses to the target word. These facilitatory effects of the CMF measure are consistent with earlier reports by Wurm and colleagues (Baayen et al., 2007; Wurm et al., 2006) showing facilitation of spoken word recognition by morphological relatives matching from word onset.

Discussion

The overall pattern of results for Experiment 2, as reflected in the factorial analyses, closely parallel those for Experiment 1. The presence and amount of word-pattern priming is entirely determined by the productivity of the roots with which a word pattern co-occurs. The productivity of the word pattern itself again seems to play no role. The linear mixed-effect analyses also reveal a functionally equivalent set of processing relationships to those seen in Experiment 1, with prime root family size—appropriately redefined for the auditory processing context—being the dominant factor interacting with the word-pattern priming effect.

This suggests that the effects we saw in Experiment 1, of the dependence of word pattern extraction on processes of root identification and extraction, and where word-pattern priming effects depend on the availability of information about the prime word pattern within a relatively narrow time window, also apply in the apparently very different context of an auditorily presented prime containing a fully specified word pattern. This in turn implies that we are looking at more general properties of how lexical access in Arabic is organised. We now turn to a broader discussion of these implications.

GENERAL DISCUSSION

The focus of this research was the nature of the analysis processes underlying the identification and segmentation of root and word-pattern morphemes in Arabic, focusing initially on visual word recognition, but extending this to auditory processing in Experiment 2. We took as a starting point some earlier incremental masked priming research (Boudelaa & Marslen-Wilson, 2005), which suggested that identification of the consonantal Arabic root (from orthographic inputs) was earlier and more robust than the identification of the word-pattern morpheme, and that the timing of word pattern identification was dependent on the timing of root identification. This seemed particularly plausible in the context of Arabic consonantal orthography, where the consonants of the root are fully specified but the short vowels of the word pattern are not.

We investigated this hypothesis about the basic organisation of Arabic lexical access in a further masked priming experiment where we attempted to directly vary the processing relations between root and word pattern by co-varying their productivity in a factorial design. Productivity, operationalised here as morphological family size, is known to affect the speed of lexical access in behavioural word recognition tasks, such as lexical decision. If word-pattern processing is tied to root processing, then slower root processing should lead to delayed identification of the word pattern—which the earlier SOA experiment suggested should disrupt word-pattern priming.

This is exactly what we found in Experiment 1. The pattern of effects was entirely unidirectional, in the sense that root productivity determined whether or not word-pattern priming would be observed, while word pattern productivity seemed to play no role at all. This is despite the fact that root and word pattern extraction must be complementary processes. The reader or listener needs to assign each letter or phoneme as it is processed to either the root or the word pattern, with a subset of consonants being potentially ambiguous between either type of morpheme. Nonetheless, in the functional context created by Arabic orthography, it is clear that the initial priority of the system is to establish the identity of the consonantal root. The more detailed breakdown provided by the linear mixed-effect analyses corroborates and refines this picture, showing that it is the properties of the prime root alone that determine priming—i.e., whether information about the prime word pattern is made available early enough to affect word pattern identification in the target.

This is not to say that we see no effects at all of word pattern productivity—in both experiments this factor significantly modulates overall RT to the target words, despite the absence of any interaction with the prime type variable. Since word pattern family size is the same for prime and target, we cannot definitively allocate this effect to prime- or target-based processes, but it is likely to be similar in nature to the effects of target root family size, which also for both experiments affected overall RT but did not modulate the priming effect. These are likely to be the same facilitatory effects of family size on lexical decision to visually presented single words that are widely reported in the current literature, though extended here to cover grammatical morphemes (the word pattern) as well as the more standardly used content words and stems. Further insight into the process of morphological analysis in Arabic is provided by the contrasting effects of root productivity on priming effects for each type of morpheme. Where word-pattern priming is concerned, as the current research demonstrates, root productivity strongly modulates the amount of priming observed. For root priming, however, as we noted in Section "Discussion" of Experiment 1, this does not hold. Roots prime each other equally effectively, in numerical terms, irrespective of the productivity of the roots involved. This raises the interesting question of why, if variations in the timing of root extraction associated with root family size have such strong effects on word-pattern priming, this is not seen for root priming as well.

In fact, this contrast seems to follow from the differences we propose in the nature of the processing relationship that holds between prime and target in the two types of morpheme. Our claim for word-pattern priming, based on the earlier SOA research and on the results here, is that this is a mediated time-sensitive process that depends on the timing with which information about the word pattern in the prime is made available to the parallel (possibly overlapping) processes that extract root and word pattern information related to the target. This is a mediated process because of the apparent dependence of word pattern identification on root identification—as reflected in the statistical relationship of word-pattern priming to prime root family size, where this is the only variable that correlates with priming. It is this sensitivity to the timing of internal analysis processes that explains why word-pattern priming is so strongly affected by variations in the timing with which root extraction processes can be conducted.

The same kind of temporal sensitivity does not seem to hold for root priming. As the incremental SOA experiment shows (Boudelaa & Marslen-Wilson, 2005), the presence of a morphological root–root relationship generates robust priming very rapidly (even at a SOA of 32 ms), and remains robustly active over a wide range of SOAs. This suggests a more direct and unmediated link between evidence for a root in the prime and boosted activation levels for the same root in the target, with no critical time window for successful priming and no dependence on some other process (such as word pattern identification). Such a relationship between activation levels should deliver equivalent amounts of priming (speeded responses relative to baseline) irrespective of variations in root family size. There is no time window within which the input from the prime must fall, and since it is the same root in prime and target there should be no mismatch between rise times of activation for prime- and target-based inputs. This would allow overall RTs to rise for low family size target words (as we observe), but for the priming effect to remain intact.

The goal of Experiment 2, in this general context, was to determine how far the evidence for the dominance of the root in the early stages of lexical access and morphemic decomposition was just a consequence of Arabic orthography and the incomplete information it provides about word patterns relative to roots. Strikingly, the use of an auditory prime, where complete information was available about the phonological properties of the word pattern as well as those of the root, did not change the pattern of effects at all, either globally or in the detail of the specific effects. There was still no word-pattern priming when its accompanying root was not productive, and the primary determinant of this effect was still the family size of the prime root. Even under the very different processing circumstances of a cross-modal priming task, with an overt spoken prime, information about the properties of the prime word pattern is apparently still not being made available early enough to affect the extraction of the same word pattern from the target word (itself, of course, still visually presented).

In fact, these parallels between the masked priming effects and the Cohort Family Size effects in the cross-modal task suggest that the general process of mapping morphemes onto internal representations of form and meaning is structured around the root. This may be why effects involving roots seem generally more robust across modalities, and across sources of noise and variation—strong root priming is seen irrespective of semantic transparency (e.g., Boudelaa & Marslen-Wilson, 2005) of root productivity (Boudelaa & Marslen-Wilson 2010a, 2010b) and of allomorphic variation in the phonetic form of the root (Boudelaa & Marslen-Wilson, 2004). Word pattern effects are more labile, with priming being disrupted in cases of allomorphic variation (Boudelaa & Marslen-Wilson, 2004) and of mismatch between the grammatical meaning of prime and target. Furthermore, the robustness of roots also seems to emerge developmentally. Roots are mastered by the age of 3 and are productively used in forming and interpreting novel forms (Clark & Berman, 1984; Hobberman, 1988), while word patterns are not fully mastered even by the age of 7 (Ravid & Farah, 1999).

These differences may, in turn, reflect both the greater communicative weight conveyed by roots, which establish the basic semantic framework for utterance interpretation, and the greater processing complexity associated with the identification and extraction of word patterns. Arguably, for example, the Arabic root will have fewer competitors during recognition process than a word pattern, irrespective of whether a competitor is defined as another word (or morpheme) with the same phonological onset (Marslen-Wilson, 1990), or as a word that has some degree of global phonetic overlap with the target (Vitevitch, Luce, Pisoni, & Auer, 1999). For example, the root {ktb}, when met in the form [kataba] write, has only the candidate [katama] hide to compete with, assuming a competitor set based on the initial phonemes of the surface form (Marslen-Wilson & Welsh, 1978). On the same assumption, the word pattern {mafʕalun}, as in a form like [maktabun] office would have 32 competitors to contend with for recognition (Baalbaki, 2001).

CONCLUSIONS

The combination of variations in morphemic productivity with different types of processing task, evaluated using both factorial statistical techniques and multi-level regression, helps us to build a more differentiated, though still very preliminary picture of decompositional processes in Arabic. The present results, which focus on word-pattern priming effects in Arabic nouns, are in keeping with an obligatory decomposition process whereby Arabic words are systematically parsed into roots and word patterns, and where the root's distributional properties dominate the lexical access process. There are also potential cross-linguistic parallels with Forster and Azuma (2000) reporting comparable productivity effects in English for masked priming between prefixed words sharing a bound stem. These parallels suggest interesting directions for future research.

In terms of theoretical accounts of the Arabic mental lexicon, the presence of strong word-pattern priming effects here and the ubiquitous root effects reported in the literature call into question recent stem-based approaches which dispense with roots and patterns as units of lexical access and lexical organisation in Semitic languages (Berent et al., 2007). An obligatory decomposition model promises to be more viable and meshes well with recent neuro-imaging research showing the left fronto-temporal brain network, typically involved in dealing with linguistic complexity, to be activated by all morphologically complex words in Arabic (Boudelaa, Bozic, Marslen-Wilson, 2010). The challenge for future research is to develop these cognitive claims about obligatory morphemic decomposition and its properties into more explicit claims about the possible neural infrastructure that can support such operations and to evaluate their possible cross-linguistic validity.

REFERENCES

Abdah, D. A. (1979). *Frequent words in Arabic*. Riyadh, Saudi Arabia: University Press of Riyadh.

Baalbaki, R. (2001). *Al-Mawrid: A modern Arabic-English dictionary* (14th ed.). Beirut, Lebanon: Dare El-Ilm Lilmalayin.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412.

Baayen, R. H., Lieber, R., & Schreuder, R. (1997). The morphological complexity of simplex nouns. *Linguistics*, *35*, 861–877.

Baayen, R. H., Piepenbrock, R., & van Rinj, H. (1993). *The CELEX lexical database [CD-ROM]*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Baayen, R. H., Tweedie, F. J., & Schreuder, R. (2002). The subjects as a simple random effect fallacy: Subject variability and morphological family effects in the mental lexicon. *Brain and Language*, *84*, 55–65.

Baayen, R. H., Wurm, H. L., & Aycock, J. (2007). Lexical dynamics for low frequency complex words: A regression study across tasks and modalities. *The Mental Lexicon*, *2*, 419–463.

Bates, D. M. (2005). Fitting linear mixed models in R. *R News*, *5*, 27–30.

Bates, D. M., & Sarkar, D. (2005). *The lme4 library.* Retrieved from http://www.lib.stat.cmu.edu/R/CRAN

Berent, I., Vaknin, V., & Marcus, G. F. (2007). Roots, stems and the universality of lexical representations: Evidence from Hebrew. *Cognition*, *104*, 254–286.

Berg, T., & Abd-Al-Jawad, H. (1996). The unfolding of suprasegmental representations: A cross-linguistic perspective. *Journal of Linguistics*, *32*, 291–324.

Bertram, R., Baayen, R. H., & Schreuder, R. (2000). Effects of family size for complex words. *Journal of Memory and Language*, *42*, 390–405.

Boudelaa, S., Bozic, M., & Marslen-Wilson, W. D. (2010, April 18–20). *Engaging fronto-temporal brain systems with derivational morphemes: An fMRI study of Arabic*. Poster presented at the 17th annual meeting of the Cognitive Neuroscience Society, Montreal, Canada.

Boudelaa, S., & Marslen-Wilson, W. D. (2000). Non-concatenative morphemes in language processing: Evidence from Modern Standard Arabic. In A. Cutler, J. McQueen & R. Zondervan

(Eds.), *Proceedings of SWAP* (*Workshop on Spoken Word Access Processes*) (Vol. 1, pp. 23–26). Max-Planck Institute for Psycholonguistics, Nijmegen, Netherlands.

Boudelaa, S., & Marslen-Wilson, W. D. (2001). The time-course of morphological, phonological and semantic processes in reading Modern Standard Arabic. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the twenty-third annual meeting of the Cognitive Science Society* (pp. 138–143). Edinburgh, Scotland.

Boudelaa, S., & Marslen-Wilson, W. D. (2004). Allomorphic variation in Arabic: Implications for lexical processing and representation. *Brain and Language*, *90*, 106–116.

Boudelaa, S., & Marslen-Wilson, W. D. (2005). Discontinuous morphology in time: Incremental masked priming in Arabic. *Language and Cognitive Processes*, *20*, 207–260.

Boudelaa, S., & Marslen-Wilson, W. D. (2009, June 14–17). *Do morphological family members inhibit each other? Family size effects in Arabic*. Paper presented at the sixth workshop on morphological processing, University of Turku, Finland.

Boudelaa, S., Pulvermüller, F., Hauk, O., Shtyrov, Y., & Marslen-Wilson, W. D. (2009). Arabic morphology in the neural language system. *Journal of Cognitive Neuroscience*, *22*, 998–1010.

Clark, E. V., & Berman, R. A. (1984). Structure and use in the acquisition of word formation. *Language*, *60*, 542–590.

De Jong, N. H., Schreuder, R., & Baayen, R. H. (2000). The morphological family size effect and morphology. *Language and Cognitive Processes*, *15*, 329–365.

Deutsch, A., Frost, R., & Forster, K. I. (1998). Verbs and nouns are organized and accessed differently in the mental lexicon: Evidence from Hebrew. *Journal of Experimental Psychology: Learning Memory & Cognition*, *24*, 1238–1255.

Feldman, L. B., Frost, R., & Pnini, T. (1995). Decomposing words into their constituent morphemes: Evidence from English and Hebrew. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *21*, 947–960.

Ford, M., Marslen-Wilson, W. D., & Davis, M. H. (2003). Morphology and frequency: Contrasting methodologies. In H. Baayen & R. Schreuder (Eds.), *Morphological structure in language processing*. Berlin: Mouton de Gruyter.

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *10*, 680–698.

Forster, K. I., & Azuma, T. (2000). Masked priming for prefixed words with bound stems: Does submit permit? *Language and Cognitive Processes*, *15*, 539–561.

Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, and Computers*, *35*, 116–124.

Frost, R., Deutsch, A., Gilboa, O., Tannenbaum, M., Marslen-Wilson, W. D. (2000). Morphological priming: Dissociation of phonological, semantic, and morphological factors. *Memory & Cognition*, *28*, 1277–1288.

Frost, R., Forster, K. I., & 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 & Cognition*, *23*, 829–856.

Glinert, L. (1996). *Modern Hebrew: An essential grammar*. London: Routledge.

Hilaal, Y. (1990). Deriving from roots and word patterns. *Linguistica Communicatio*, 1, 77–80. (in Arabic)

Hobberman, R. (1988). Local and long distance spreading in Semitic morphology. *Natural Language and Linguistic Theory*, *6*, 544–549.

Holes, C. (1995). *Modern Arabic*. London and New York: Longman.

Idrissi, A., Prunet, J. F., & Béland, R. (2008). On the mental representation of Arabic roots. *Linguistic Inquiry*, *39*, 221–259.

Keppel, G. (1982). *Design and analysis: A researcher's handbook*. Englewood Cliffs, NJ: Prentice-Hall.

Khouloughli, D.-E. (1992). *Basic lexicon of Modern Standard Arabic*. Paris: L'Harmattan.

Laudanna, A., Burani, C., & Cermele, A. (1994). Prefixes as processing units. *Language and Cognitive Processes*, *9*, 295–316.

Longtin, C. M., Segui, J., & Halle, P. A. (2003). Morphological priming without morphological relationship. *Language and Cognitive Processes*, *18*, 313–334.

Marslen-Wilson, W. D. (1990). Activation, competition, and frequency in lexical access: 9. In G. T. M. Altman (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 148–172). Cambridge, MA: MIT Press.

Marslen-Wilson, W. D., Bozic, M., & Randall, B. (2008). Early decomposition in visual word recognition: Dissociating morphology, form, and meaning. *Language and Cognitive Processes.* In press.

Marslen-Wilson, W. D., Ford, M., Older, L., & Zhou, X. (1996). The combinatorial lexicon: Priming derivational affixes. In G. Cottrell (Ed.), *Proceedings of the 18th annual conference of the Cognitive Science Society* (pp. 223–227). Mahwah: NJ: Lawrence Erlbaum Associates, Inc.

Marslen-Wilson, W. D., Tyler, L. K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review*, *101*, 3–33.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions during word recognition in continuous speech. *Cognitive Psychology*, *10*, 29–63.

Moscoso Del Prado Martin, F., Deutsch, A., Frost, R., Schreuder, R., De Jong, N., & Baayen, H. (2005). Changing places: A cross-language perspective on frequency and family size in Dutch and Hebrew. *Journal of Memory and Language*, *53*, 496–512.

Norris, D., & Kinoshita, S. (2008). Perception as evidence accumulation and Bayesian inference: Insights from masked priming. *Journal of Experimental Psychology: General*, *137*, 434–455.

Pollatsek, A., & Well, A. D. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning Memory & Cognition*, *21*, 758–794.

Prunet, J. F., Béland, R., & Idrissi, A. (2000). The mental representation of Semitic words. *Linguistic Inquiry*, *31*, 609–648.

Rastle, K., Davis, M. H., Marslen-Wilson, W. D., & Tyler, L. K. (2000). Morphological and semantic effects in visual word recognition: A time-course study. *Language and Cognitive Processes*, *15*, 507–537.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*, 510–532.

Ratcliffe, R. R. (2004). Sonority-based parsing at the margins of Arabic morphology: In response to Prunet, Béland and Idrissi (2000) and Davis and Zawaydeh (1999, 2001). *Al-'Arabiyya*, *37*, 73–95.

Ravid, D., & Farah, R. (1999). Learning about noun plurals in early Palestinian Arabic. *First Language*, *19*, 187–206.

Schreuder, R., & Baayen, R. H. (1997). How complex simplex words can be. *Journal of Memory and Language*, *37*, 118–139.

Ussishkin, A. (1999). The inadequacy of the consonantal root: Modern Hebrew denominal verbs and output–output correspondence. *Phonology*, *32*, 441–442.

Ussishkin, A. (2005). A fixed prosodic theory of nonconcatenative templatic morphology. *Natural Language and Linguistic Theory*, *23*, 169–218.

Versteegh, K. (1997). *The Arabic language.* Edinburgh, UK: Edinburgh University Press.

Vitevitch, M. S., Luce, P. A., Pisoni, D. B., & Auer, E. T. (1999). Phonotactics, neighborhood activation, and lexical access for spoken words. *Brain and Language*, *69*, 306–311.

Wright, W. (1995). *A grammar of the Arabic language*. Cambridge, UK: Cambridge University Press.

Wurm, L. H., Ernestus, M., Schreuder, R., & Baayen, R. H. (2006). Dynamics of the auditory comprehension of prefixed words: Cohort entropies and conditional root uniqueness points. *The Mental Lexicon*, *1*, 125–146.