

# Morphological Structure in the Arabic Mental Lexicon: Productivity and Priming Effects in Nominal and Verbal Patterns

Abeer Ali Abbas<sup>1</sup>, Lily Xu<sup>2</sup>, Megha Sundara<sup>2</sup>

<sup>1</sup> Department of Foreign Languages, Jazan University

<sup>2</sup> UCLA Department of Linguistics

## Abstract

Semitic languages are characterized as having two types of discontinuous morphemes: roots and word patterns. The role of these morphemes in lexical access and representation remains debated, especially in the case of word patterns. Roots exhibit robust priming for both nouns and verbs, while word patterns yield mixed results—verbal patterns tend to show stronger priming effects than nominal ones. While previous research (e.g., Deutsch et al., 1998) suggested that differences in productivity might explain these word class effects on word pattern priming, no study has directly investigated this hypothesis. To isolate the contributions of productivity and word class, we used a 2×2 factorial design crossing productivity (high vs. low) with word class (verb vs. noun). This design allows us to disentangle the two variables that were previously confounded in studies such as Boudelaa & Marslen-Wilson (2015). We did this using a masked visual priming experiment in Arabic. We found that, regardless of word class, high-productivity patterns showed robust priming, whereas low-productivity patterns did not. Additionally, priming from high-productivity patterns was distinct from semantic and orthographic effects, confirming the independent role of word patterns in morphological decomposition. These results support the dual-route model of lexical access (Baayen et al., 1997).

**Keywords** Psycholinguistics; Mental Lexicon; Morphological priming; Productivity; Arabic; Word Patterns; Lexical Access

## Background

Semitic languages exhibit non-concatenative morphology, which consists of two abstract morphemes: roots and word patterns. Roots are typically sequences of three consonants that carry core semantic meaning. Word patterns, on the other hand, consist mainly of vowels but may also include consonants. These patterns determine a word's category and morphosyntactic structure (Wright, 1995). The role of roots and patterns in lexical access and word recognition has been extensively investigated (e.g., Boudelaa & Marslen-Wilson, 2000, 2005, 2011, 2015; Deutsch et al., 1998; Frost et al., 1997).

Previous research has consistently shown that roots exhibit robust priming effects in both nouns and verbs (Boudelaa & Marslen-Wilson, 2000, 2015; Deutsch et al., 1998; Frost et al., 1997). However, the priming effects of word patterns have been more variable, with verbal pattern more likely to demonstrate priming compared to nominal patterns (Deutsch et al., 1998; Boudelaa & Marslen-Wilson, 2005, 2013, 2015). Priming by nominal patterns has been found in some studies (Boudelaa & Marslen-Wilson, 2005, 2015), but not others (Frost et al., 1997; Abu-Rabia &

Awwad, 2004; Deutsch et al., 2005; Shalhoub-Awwad, 2020).

One possible explanation for the differences in priming effects between nouns and verbs is the productivity of patterns associated with verbs and nouns. We define productivity here as the type frequency, or family size, of a given morpheme—the number of word forms generated by a root or pattern (Deutsch et al., 1998). In both Hebrew and Arabic, verb patterns tend to be more productive than noun patterns, potentially accounting for the stronger priming effects observed for verbal patterns in both languages. However, no study has directly compared the effects of productivity and word class in morphological priming of word patterns. We addressed this gap by investigating whether priming differences between nominal and verbal word patterns in Arabic are due to differences in productivity.

We used a masked visual priming task to test whether productivity or word class is the primary factor affecting word pattern priming in Arabic. This method is ideal because it targets morphological priming before higher-level semantic processing occurs. We used a stimulus onset asynchrony (SOA) of 48 ms, because it has been previously identified as an optimal window for early morphological activation involving word patterns in Arabic (Boudelaa & Marslen-Wilson, 2005; Xu, Sola-Llonch, Wang & Sundara, 2023).

If productivity is the primary factor, high-productivity patterns should exhibit stronger priming than low-productivity patterns regardless of word class. If word class is more influential, verbal patterns should prime regardless of productivity. Furthermore, we compared word pattern priming to semantic and orthographic priming to determine if it is morphological in nature. These findings will clarify whether differences between verbal and nominal patterns stem from productivity or word class distinctions, contributing to models of morphological processing in Arabic and other non-concatenative languages.

## Methods

### Participants

A total of 132 native Arabic speakers aged 20 to 45 years (average age 34) participated in the experiment.

### Materials

Productivity in this study was measured using type frequency—the number of distinct words that can be formed from a morpheme. We used type frequency data from the Aralex database (Boudelaa & Marslen-Wilson, 2010). This metric reflects the number of distinct roots associated with

each pattern. Type frequency offers a simple, corpus-based estimate that has been previously used to evaluate Arabic morphology. The word patterns used in this experiment belonged to one of four conditions, based on word class (verb vs. noun) and productivity (high vs. low).

The experiment included 128 target items across four conditions defined by productivity (high vs. low) and word class (noun vs. verb). Four templates were paired with eight roots, resulting in 32 items per condition. An additional 128 non-words were constructed using real word patterns with pseudo-roots. Each target was paired with a test prime (same word pattern, productivity level, and word class), a semantic control prime (meaning-related but without morphological or orthographic similarity), an orthographic control prime (sharing three phonological segments but unrelated in meaning or morphology), and an unrelated control prime (no relation to the target). Four counterbalanced lists of materials were constructed to avoid repetition of the targets.

Stimuli were matched on root frequency using the Aralex corpus, with all roots having type frequencies between 20–25 to control for lexical effects. This range represents the median productivity of roots in Aralex and was chosen to minimize interference from highly frequent or infrequent roots.

For example, a high-productivity verb with the template [Ca.CaC] (type frequency: 456), such as the verb [sa.nad] ‘supported,’ was paired with [ha.raθ] ‘plowed’ (morphological), [ta.ʕaa.wun] ‘cooperation’ (semantic), [sa.ʕiid] ‘happy’ (orthographic), and [ha.qii.bah] ‘bag’ (unrelated). Table 1 presents the mean, median, and range of type frequencies for the word patterns in each condition. Importantly, the type frequency measures were comparable across word classes.

Table 1: Productivity of word patterns in stimuli

Conditions	Mean	Median	Range
High productivity verbs	464.75	448.5	382-580
Low productivity verbs	143.5	130	18-296
High productivity nouns	481.25	450.5	392-632
Low productivity nouns	151.25	147.5	37-273

Orthographic ambiguity was minimized by using word patterns fully specified with long vowels or diacritics, ensuring consistency across conditions. The only ambiguous template used was [fa.ʕal], a high-productivity verb pattern, disambiguated with diacritics to ensure accurate processing and avoid misinterpretation.

Ten native Arabic speakers rated word familiarity on a 1–7 scale, with all items scoring at least 5. Another group rated semantic relatedness, where semantic control pairs scored 5 or higher, while unrelated, orthographic, and morphological pairs scored below 3.

## Procedure

Participants completed a lexical decision task in which they were asked to identify whether or not the stimulus was a real Arabic word. Each trial began with a 500 ms forward mask of 28 vertical lines in a 30-point font size, following Boudelaa & Marslen-Wilson (2011), who found this mask more effective than standard hash marks. A visual prime was then presented for 48 ms (the SOA) in a 24-point font size, followed by the target in 34-point. The target remained on screen until the participant responded or 2,000 ms elapsed.

Trials were randomized, and each session lasted 15–20 minutes. Reaction times and accuracy were recorded to assess morphological priming effects across productivity levels, word classes, and control conditions (morphological, semantic, orthographic, and unrelated).

## Analysis and Results

Reaction times (RTs) were analyzed using linear mixed-effects models (LME) with the lme4 package (Bates et al., 2015) in R (R Core Team, 2020). Fixed effects included prime type (test, semantic, orthographic, unrelated), productivity (high vs. low), and word class (verb vs. noun), along with their interactions. Random intercepts for subjects and items were included. Estimated marginal means (EMMs) were computed for significant interactions using the emmeans package (Lenth, 2025). Accuracy was analyzed using a generalized linear mixed-effects model (GLMM) and the same fixed and random effects. Treatment contrast coding was applied to all categorical predictors in the linear mixed-effects models.

## Accuracy

Accuracy was consistently high (96–98%) across all conditions. Participants performed at ceiling in the lexical decision task and were unaffected by the experimental manipulations ( $p > 0.05$  for prime type, productivity, and word class). These data were not analyzed further.

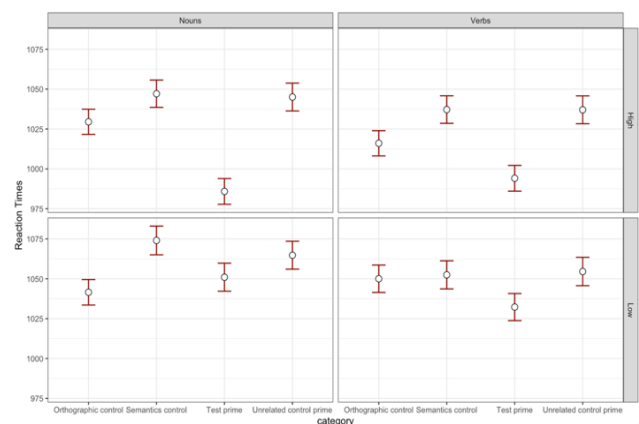


Figure 1: Mean reaction times (RT) for high and low productivity nouns and verbs across test, phonological, semantic, and unrelated control conditions in the masked visual priming task with an SOA of 48ms.

## Reaction times

As shown in Figure 1, morphological priming was observed for both high-productivity nouns and verbs, with significantly faster reaction times in the test prime condition compared to unrelated controls ( $\beta = 43.39$ ,  $SE = 18.15$ ,  $t = 2.39$ ,  $p = .017$ ). Low-productivity nouns and verbs showed no priming. The interaction between test prime and productivity was significant, with only high-productivity items showing priming. However, there was no interaction between test prime and word class ( $\beta = 19.35$ ,  $SE = 18.16$ ,  $t = 1.07$ ,  $p = .286$ ), indicating that both high-productivity nouns and verbs exhibited similar priming effects. These results confirm that productivity, rather than word class, is the primary factor affecting word pattern priming.

Further, No priming was found in the orthographic ( $\beta = -18.26$ ,  $SE = 12.79$ ,  $t = -1.43$ ,  $p = .154$ ) or semantic control conditions ( $\beta = .73$ ,  $SE = 12.81$ ,  $t = .057$ ,  $p = .954$ ). These findings confirm that word pattern priming is morphological and is distinct from orthographic and semantic influences.

## General Discussion

The results of this study demonstrate that productivity, rather than word class distinctions, is the primary factor influencing morphological priming. High-productivity patterns in both nouns and verbs exhibited morphological priming, whereas low-productivity items showed no evidence of priming in either word class. These findings align with research on concatenative languages like English, where high-productivity affixes enhance lexical processing (Marslen-Wilson et al., 2019). Extending these observations to Arabic, a non-concatenative language, underscores productivity as a key factor in morphological priming across different linguistic systems.

Our findings have implications for models of lexical access. A fully decompositional model (Taft & Forster, 1975) where all words are morphologically decomposed can account for the robust priming observed for high-productivity nouns and verbs in this study. However, it fails to account for the absence of priming in low-productivity items, challenging the assumption that all words are represented and accessed via morphemes. What remains to be determined is whether low productivity items fail to prime in this experiment simply because their decomposition takes longer. We are testing whether word patterns with low productivity prime at longer SOAs.

Whole-word models (Seidenberg & Gonnerman, 2000) where words are stored and accessed as unanalyzed units also fail to account for all the findings. The lack of priming in low-productivity items is consistent with the representation and access of unanalyzed units. But this model fails to explain the priming in high-productivity words.

More recent distributed models, such as Stevens and Plaut (2022), extend this view by arguing that priming effects can arise from learned form-meaning associations, without explicit morphological decomposition. These models fall within the broader class of whole-word accounts. Counter to traditional whole word accounts, in this view, the robust

priming for high-productivity patterns may reflect more reliable form-meaning mappings, as predicted by distributed theories. However, priming in non-concatenative systems like Arabic remains a challenge for existing implementations of distributed models such as Baayen et al. (2022) and Chuang et al. (2021) which use bigram or trigram based orthographic representations. Future computational modeling is needed to address this issue.

The dual representation, dual access model (Baayen et al., 1997) offers the most comprehensive explanation of our findings. In these models, words are represented as whole-word and decomposed forms, with productivity determining access routes. Words containing high-productivity morphemes are accessed faster through the activation of constituent morphemes, whereas words with low-productivity morphemes are retrieved faster as whole-words. This aligns with the significant pattern priming observed for high-productivity items at 48 ms SOA, whereas the absence of priming for low-productivity items is consistent with whole-word access.

Overall, these findings support a dual-route model for lexical access, where both decomposition and whole-word retrieval contribute to word recognition, but productivity determines which mechanism dominates. As a result, High-productivity items consistently show morphological priming, while low-productivity items depend on whole-word access, limiting priming effects.

Throughout this paper, we have dichotomized productivity. Future research is needed to determine whether effects of productivity are more gradient as might be expected under a dual representation, dual access model.

## Conclusion

We investigated the role of productivity in morphological priming in Arabic, a language with extensive non-concatenative morphology. We used a masked visual priming experiment to assess priming effects in both verbal and nominal patterns. We asked whether priming differences between verbal and nominal patterns stem from word class distinctions or productivity effects. The results highlight productivity as the key factor in priming for both verbal and nominal patterns. High-productivity patterns in Arabic exhibited strong morphological priming, whereas low-productivity patterns showed no reliable effects, regardless of word class. These findings reinforce the psychological reality of non-concatenative morphemes in Arabic, particularly word patterns, and show that their role in lexical access is moderated by productivity. We suspect similar productivity differences can also explain differences in priming reported for nouns and verbs in Hebrew. Although this study is behavioral, it motivates future neuroimaging research on the role of morphological productivity in typologically rich, non-concatenative systems like Arabic. In doing so, it extends psycholinguistic models to better account for morphologically rich and understudied languages.

## References

- Abu-Rabia, S., & Awwad, J. (2004). Morphological structures in visual word recognition: The case of Arabic. *Journal of Research in Reading, 27*(3), 321-336.
- Baayen, R. H., Lieber, R., & Schreuder, R. (1997). The morphological complexity of simplex nouns. *Linguistics, 35*, 861-877.
- Baayen, R. H., Chuang, Y. Y., Shafaei-Bajestan, E., & Blevins, J. P. (2019). The discriminative lexicon: A unified computational model for the lexicon and lexical processing in comprehension and production grounded not in (de) composition but in linear discriminative learning. *Complexity*, vol. 2019, Article 4895891.
- Bates, D. (2016). lme4: Linear mixed-effects models using Eigen and S4. *R package version, 1*, 1.
- Boudelaa, S., & Marslen-Wilson, W. D. (2000). Non-concatenative morphemes in language processing: Evidence from Modern Standard Arabic. In *Proc. SWAP 2000* (pp. 23-26).
- 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. (2010). Aralex: A lexical database for modern standard Arabic. *Behavior Research Methods, 42*(2), 481-487.
- Boudelaa, S., & Marslen-Wilson, W. D. (2011). Productivity and priming: Morphemic decomposition in Arabic. *Language and Cognitive Processes, 26*(4-6), 624-652.
- Boudelaa, S., & Marslen-Wilson, W. D. (2013). Morphological structure in the Arabic mental lexicon: Parallels between standard and dialectal Arabic. *Language and cognitive processes, 28*(10), 1453-1473.
- Boudelaa, S., & Marslen-Wilson, W. D. (2015). Structure, form, and meaning in the mental lexicon: Evidence from Arabic. *Language, Cognition and Neuroscience, 30*(8), 955-992. doi:10.1080/23273798.2015.1048258
- Chuang, Y. Y., Kang, M., Xuefeng, L., & Harald, B. R. (2021). Vector space morphology with linear discriminative learning. In D. Crepaldi (Ed.), *Linguistic morphology in the mind and brain* (pp. 167-183). Routledge.
- 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.
- Deutsch, A., Frost, R., Pollatsek, A., & Rayner, K. (2000). Early morphological effects in word recognition in Hebrew: Evidence from parafoveal preview benefit. *Language and Cognitive Processes, 15*(4-5), 487-506.
- Deutsch, A., Frost, R., Pollatsek, A., & Rayner, K. (2005). Morphological parafoveal preview benefit effects in reading: Evidence from Hebrew. *Language and Cognitive Processes, 20*(1-2), 341-371.
- 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.
- Lenth R (2025). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.10.7.
- Marslen-Wilson, W. D., Ford, M., Older, L., & Xiaolin, Z. (2019). The combinatorial lexicon: Priming derivational affixes. In *Proceedings of the eighteenth annual conference of the Cognitive Science Society*.
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Seidenberg, M. S. & Gonnerman, L. M. (2000). Explaining derivational morphology as the convergence of codes. *Trends in Cognitive Sciences, 4*(9), 353-361.
- Shalhoub-Awwad, Y. (2020). The role of nominal word pattern in Arabic reading acquisition: Insights from cross-modal priming. *Scientific Studies of Reading, 24*(4), 307-320.
- Stevens, P., & Plaut, D. C. (2022). From decomposition to distributed theories of morphological processing in reading. *Psychonomic Bulletin & Review, 29*(5), 1673-1702.
- Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior, 14*(6), 638-647.
- Wright, W. (1995). *A grammar of the Arabic language*. Cambridge: Cambridge University Press.
- Xu, L., Solá-Llonch, E., Wang, H., & Sundara, M. (2023). A meta-analytic review of morphological priming in Semitic languages. *The Mental Lexicon, 18*(2), 300-337.