Effects and interactions of orthographic depth and lexicality in Arabic visual word recognition: A lexical decision ERP study

Ali Idrissi[a]*, R. Muralikrishnan[b], Eiman Mustafawi[a], Tariq Khwaileh[a] & John E. Drury[a]

[a] Qatar University, [b] Max Planck Institute for Empirical Aesthetics *aidrissi@qu.edu.qa

Max Planck Institute for Empirical Aesthetics

QATAR UNIVERSITY

Introduction

- **The Arabic script uses independent letters to represent consonants and long vowels, typically leaving short vowels unmarked**.

- **However, short vowels can be marked with diacritics**, which provide studies of visual word recognition with degrees of deep/shallow orthography that can be examined in tightly matched minimal pairs.

- **Bourisly et al. (2013)** show distinct fMRI activations suggestive of additional lexical search when diacritics are absent, and greater involvement of mappings to phonology/semantics when they are present (cf. Weiss et al. 2015).

- **An Arabic semantic priming ERP study** (Mountaj et al. 2015) **found larger early negativities when diacritics were present** (~N1 peak) **but no influence on N400 semantic relatedness effects** (cf. Bar-Kochav & Breznitz 2012).

- <u>Present Study</u>. Here we examine ERP responses in lexical decision involving **Real Roots** (RRs; see examples below) or **Pseudo Roots** (PRs). **Three levels of diacritic density** (FULL/MIN/NON) **were included for both RRs/PRs**. Two classes of mismatch conditions were tested, with unattested diacritic patterns occurring with either RR or PR, or possible but RR-mismatching diacritics.
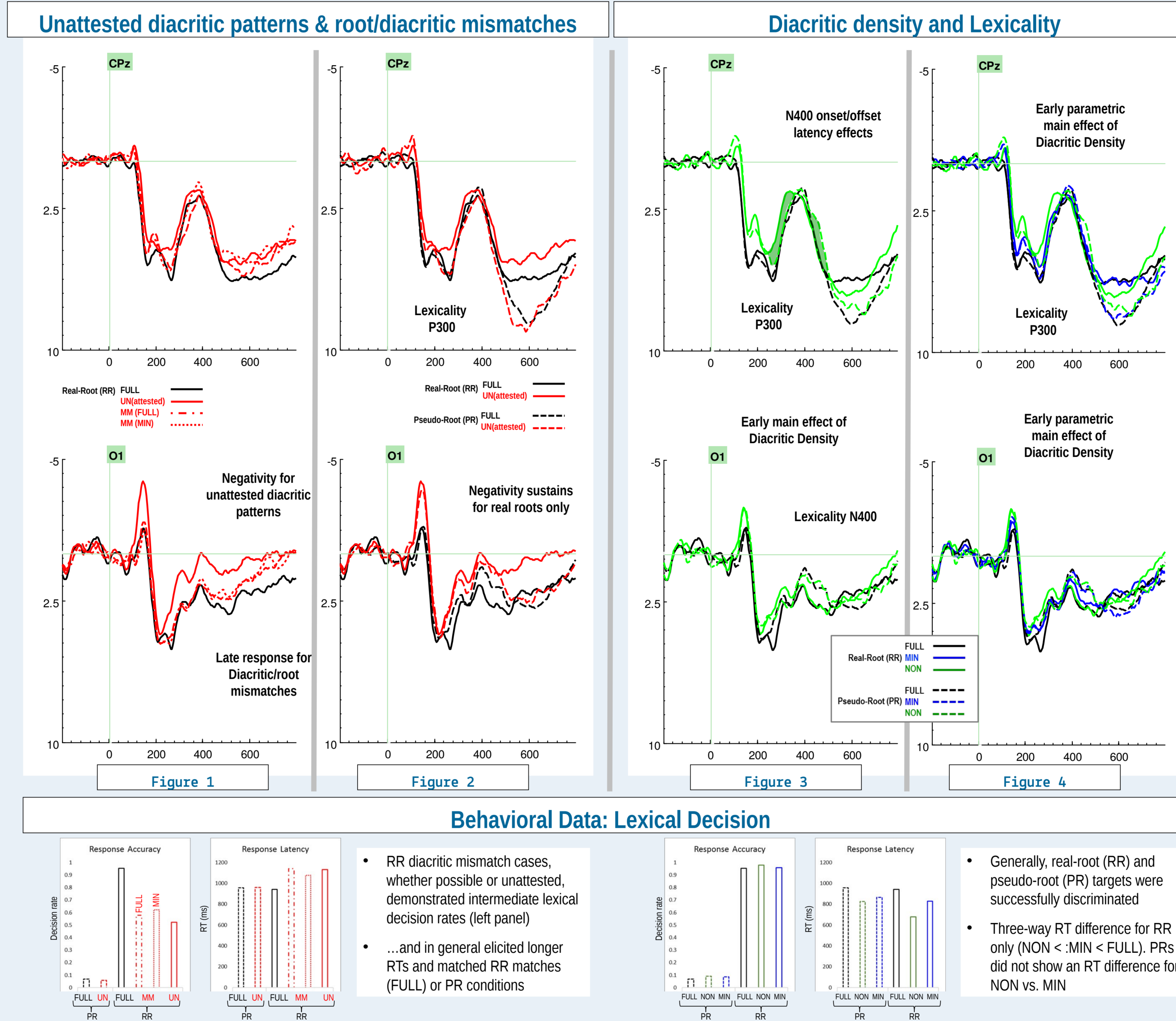
Real Root (RR) Conditions	Correct	Mismatch (possible)	Mismatch (unattested)
Fully marked (FULL)	شَغَفْ	شَغَفْ	شَغَفْ
Minimally Marked (MIN)	شغَف	شغَف	شغَف
Non-marked (NON)	شغف		

Pseudo Root (PR) Conditions

Methods

- **Participants.** 33 right-handed adult native speakers of Gulf Arabic; 2 participants excluded due to noisy EEG. Results presented here are thus based on 31 participants.

- **Materials & Procedure.** Participants judged 720 isolated words centrally presented in white font against a black background. Half constituted real words, half pseudo words (360 items each).

- **EEG Data** recorded at 25 Ag / AgCl scalp electrode sites; 250 Hz sampling rate; Reference: Left-mastoid, re-referenced to linked mastoids offline; Ground electrode: AFZ; Offline filter: 0.1 – 30 Hz band-pass; 1000 ms epochs isolated for ERP analysis (-200 to 800 ms post target onset).

- **ERP Analyses.** Mean amplitudes were examined for effects of Mismatches & Lexicality (Figures 1 & 2), and for effects and interactions of Diacritic Density (FULL/MIN/NON) and Lexicality (Figures 3 & 4). N400 latency was also examined via peak and 50% fractional peak latency measures.

References. Bar-Kochva, I. & Z. Breznitz 2012. Does the reading of different orthographies produce different brain activity patterns? An ERP Study. PLoS One, 7:5, 1-9. **Bourisly et al. 2013.** Neural correlates of diacritics in Arabic: An fMRI study. *Journal of Neurolinguistics* 26, 195-206. **Mountaj, N. et al. 2015.** Vowelling and semantic priming effects in Arabic. International Journal of Psychophysiology 95, 46-55. **Weiss et al. 2015.** Many ways to read your vowels—Neural processing of diacritics and vowel letters in Hebrew. *Neuroimage* 121, 10-19.

Results

Unattested diacritic patterns & root/diacritic mismatches

Figure 1

Figure 2

Diacritic density and Lexicality

Figure 3

Figure 4

Behavioral Data: Lexical Decision

- RR diacritic mismatch cases, whether possible or unattested, demonstrated intermediate lexical decision rates (left panel)

- …and in general elicited longer RTs and matched RR matches (FULL) or PR conditions

- Generally, real-root (RR) and pseudo-root (PR) targets were successfully discriminated

- Three-way RT difference for RR only (NON < :MIN < FULL). PRs did not show an RT difference for NON vs. MIN

SUMMARY

- Unattested diacritic patterns gave rise to an early posterior negativity beginning just before N1 peak, equally for RR and PRs (no Lexicality interaction; see O1 Fig.2)

- For RR's only, this negativity sustained throughout the rest of the measurement epoch (800 ms). PR's tracked this pattern only up to ~500 ms (see O1 Fig.2)

- Possible but root-mismatching diacritic patterns(MM (FULL) & MM (MIN)) did not show any early response relative to controls; But,… a late effect similar to the sustained negativity for unattested diacritics emerged (also after ~500 ms; see O1 Fig.1)

- Diacritic density manifested as an early (parametric) effect: relative negativities for NON > MIN > FULL beginning around the N1 peak and sustaining for ~200 ms (Fig.4).

- Presence/absence (FULL vs. NON) of diacritics exhibited differential N400 latency effects for RRs and PRs (see CPz, Fig.3):
  - N400 **onset** *earlier* for NON than FULL when targets were RRs
  - N400 **offset** *later* for NON than FULL when targets were PRs

- Lexicality N400 + LPC/P300 (PR > RRs)

Discussion

- The striking **early responses to unattested diacritic patterns** (Figures 1 & 2) **demonstrate rapid sensitivity to more than mere presence/absence of diacritics**, contra previous interpretations of early ERP effects tied to these elements in Arabic and Hebrew (Mountaj et al. 2015; Bar-Kochav & Breznitz 2012). Further, the combination of patterns for RRs and PRs, including responses to possible diacritic patterns which mismatch with particular RRs, together suggest that: (i) word pattern information represented by diacritics is detected and processed in detail immediately, but (ii) this word pattern information is only integrated with root information after the processes underlying the N400 are well underway (or completed).

- Our **parametric effect of Diacritic Density on early ERP responses** (Figure 4) **is the opposite of early effects found in prior studies** (where targets with diacritics were more negative than those without). This is also consistent with early ERP effects being sensitive to more than presence/absence of diacritics. Why we should find the opposite pattern here will require further investigation, but two features of our experiment seem relevant: (i) the proportions of our NON/MIN/FULL cases in the experiment track the parametric distribution of the early responses, and (ii) we included diacritic mismatches/violations. The contributions of these factors will have to be teased apart in further studies.

- When diacritics were absent, **RR N400 responses showed an earlier onset, and PR N400 responses showed a latter offset** (Figure 3). Such influences of orthographic depth on N400 onset/offset latency have not been previously reported in ERP studies of lexical decision. We suggest these timing effects index the size of the candidate set activated by target items. A larger such set (for NON cases, compared to FULL) yields an advantage for RR's, but delays completion of the processes underlying the N400 for PRs. Future work should examine factors such as root type frequency or neighbourhood size as continuous predictors of N400 onset/offset to replicate/extend this finding and determine its significance for lexical processing.