جامعة الإمارات العربية المتحدة
United Arab Emirates University

جامعة نيويورك ابوظبي
NYU ABU DHABI

The cost of processing vowel diacritics in Arabic: Evidence from masked-priming

Diogo Almeida, Kevin Schluter, Matthew Tucker (NYUAD)
&
Ali Idrissi (UAE University)

Roadmap for today

- The problem of Visual Word Recognition
  - Specific issues related to reading Arabic
- Previous work
- Experimental design
- Results
- Discussion

Basic problem in visual word recognition

- A major problem to overcome: Perceptual confusability
  - Small inventory of basic symbols (28 in Arabic)
  - Large number of word forms (tens of thousands)
    - All words resemble each other to some extent, because symbol sharing is rampant.

- Overcoming perceptual confusability
  - Familiarity with full word form *as a unit*
    - *Gestalt* representation
  - Reduction in lexical competition by reducing form ambiguity

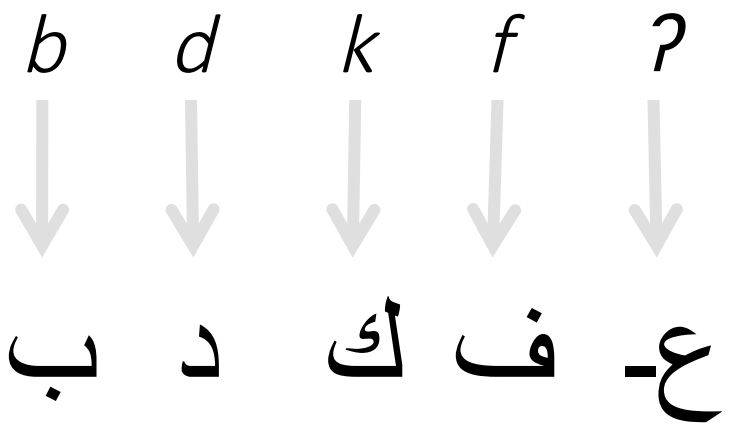
Basic problem in visual word recognition in Arabic

- Overcoming perceptual confusability
  - Familiarity with full word form *as a unit*
  - Reduction in lexical competition by reducing form ambiguity

- These two sources of help are in opposition in Arabic
  - Short vowels and long consonants are not **usually** written
    - Form ambiguity is fairly common
    - Disambiguation by context
  - Visual word forms in Arabic:
    - Most familiar tend to be the most ambiguous.
    - Least familiar tend to be the least ambiguous.

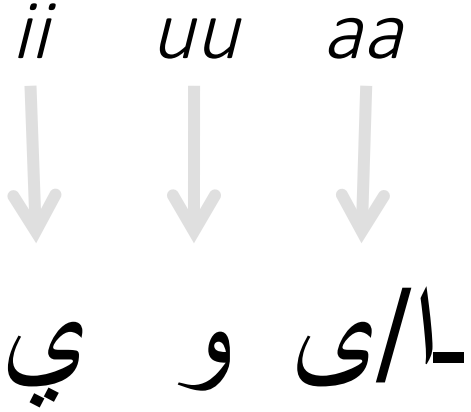
Arabic Orthography

- Only the 28 consonants are indicated with independent symbols (*abjad*);



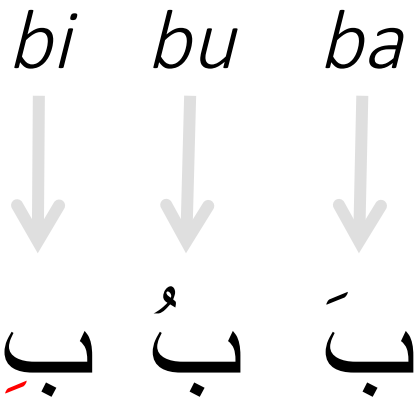
Arabic Orthography

- Only the 28 consonants are indicated with independent symbols (*abjad*);
- The long vowels ([ii], [uu] and [aa]) are rendered by the letters representing the consonants [y], [w] and [?];



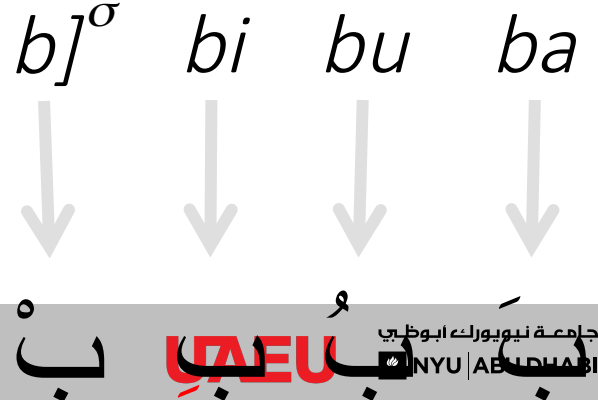
Arabic Orthography

- Only the 28 consonants are indicated with independent symbols (*abjad*);
- The long vowels ([ii], [uu] and [aa]) are rendered by the letters representing the consonants [y], [w] and [?];
- Short vowels are indicated with diacritics above or beneath letter symbols;



Arabic Orthography

- Only the 28 consonants are indicated with independent symbols (*abjad*);
- The long vowels ([ii], [uu] and [aa]) are rendered by the letters representing the consonants [y], [w] and [?];
- Short vowels are indicated with diacritics above or beneath letter symbols;
- The absence of a vowel (in coda position) is marked with a sukuun;



Arabic Orthography

- Vowels (and diacritics in general) are not used in everyday Arabic texts;
- Exception: instruction materials for young children or second-language learners;
- Religious text;
- Mainly for disambiguation purposes.

Form Ambiguity in Arabic

‘to cause to carry’

حَمَّلَ

‘to carry’

حَمَلَ

حُمِّلَ

‘to be caused to carry’

حمل

حُمِلَ

‘to be carried’

حِمْل

‘burden’

حَمْل

‘pregnancy’

حَمَل

‘a lamb’

Basic problem in visual word recognition in Arabic

- Visual word forms in Arabic:
  - most familiar tend to be the most ambiguous.
  - Least familiar tend to be the least ambiguous.

- **Research question:**
  - What is the ***relative contribution*** of form familiarity and form ambiguity to visual word recognition in Arabic?

Previous work – Equivocal findings

- **Research question:**
  - What is the ***relative contribution*** of form familiarity and form ambiguity to visual word recognition in Arabic?

- Sometimes vowel diacritics help (citations)
  - Facilitatory role due to reduction in form ambiguity
- Sometimes vowel diacritics hinder (citations)
  - Inhibitory role due to controlled processes (e.g., grapheme-to-phoneme conversion strategy)
  - Implies that **form familiarity alone** is enough to recognize words, relatively automatically.

Basic problem in visual word recognition in Arabic

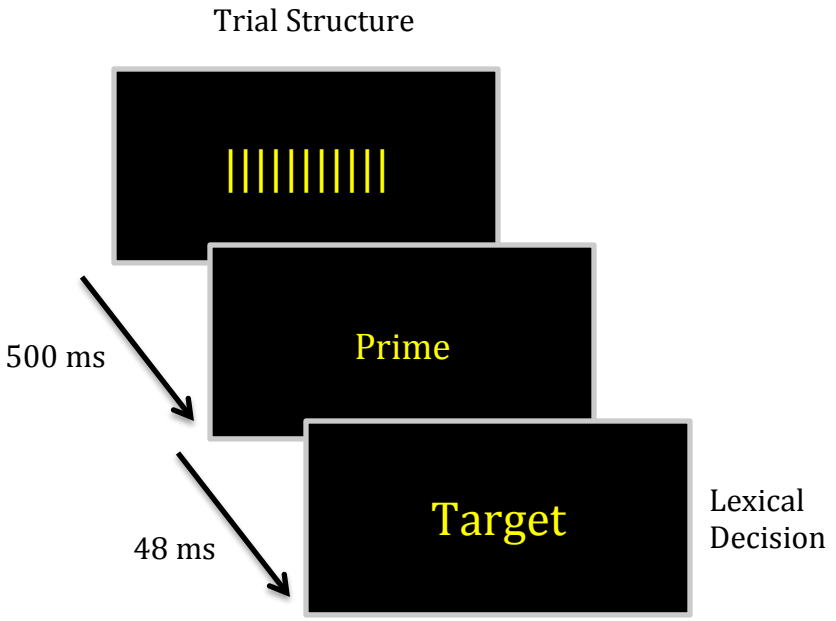
- Beyond theoretical concerns:
  - Arabic is heavily diglossic
    - spoken varieties are not written
    - written variety is seldom spoken
  - Experimental psycholinguistic work growing. Frequent questions:
    - Should we use vowel diacritics in our study or not?
    - When and where are vowel diacritics appropriate?
    - What would be the impact of deciding to use/omit them on the results?

Experimental Design

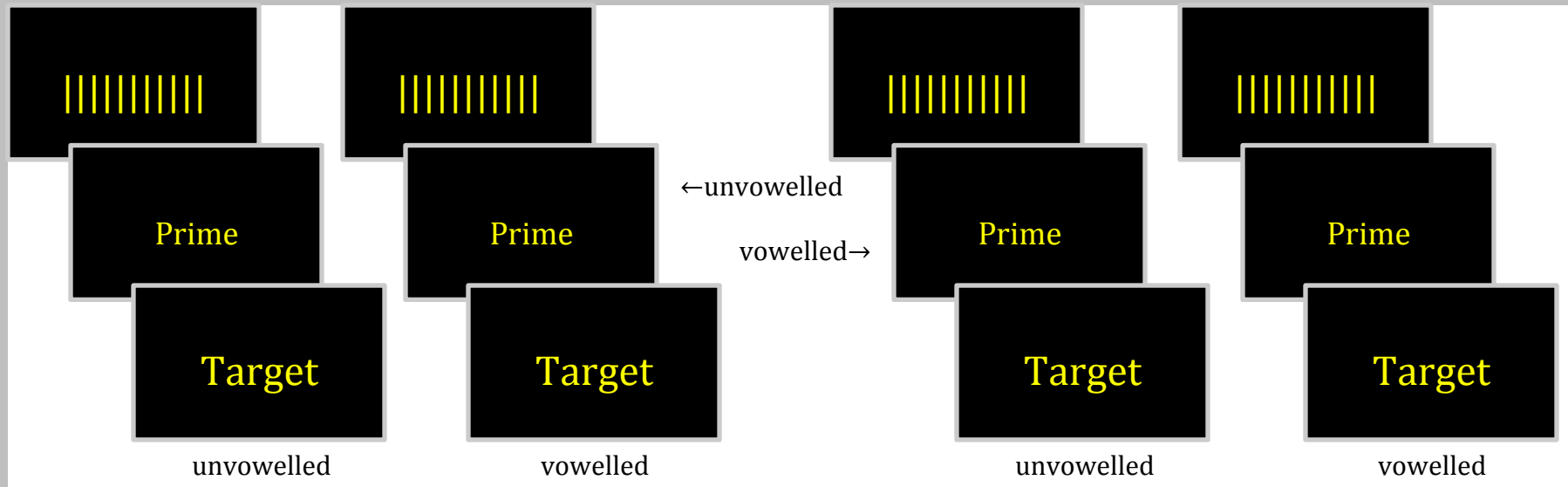
- Masked priming
  - Potential to tap into the long-term representation of visual word forms
- 2 x 2 x 2 factorial design
  - Lexicality (word vs pseudoword)
  - Prime voweling (unvoweled vs voweled)
  - Target voweling (unvoweled vs voweled)
- Lexical decision task on target
  - Measure: Relative magnitude of repetition priming effect

Experimental Design

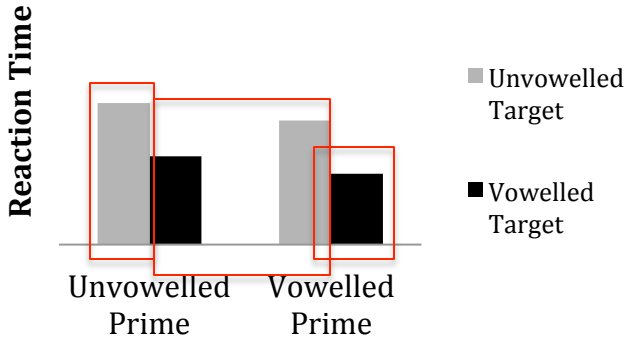
- 55 participants
  - All female (female campus of UAEU), native speakers of Arabic
  - 49 analyzed
  - 6 rejected (error rate > 15%)
- Time out = 2 s (no outlier rejection)



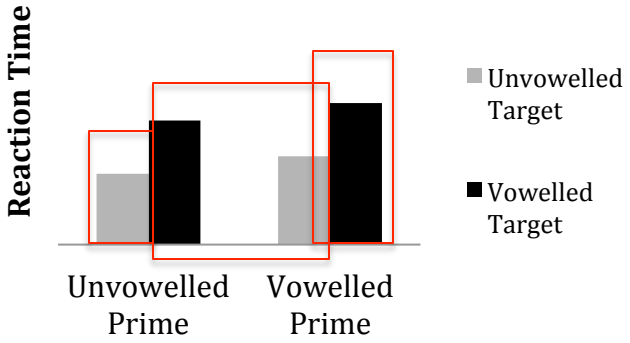
Predictions



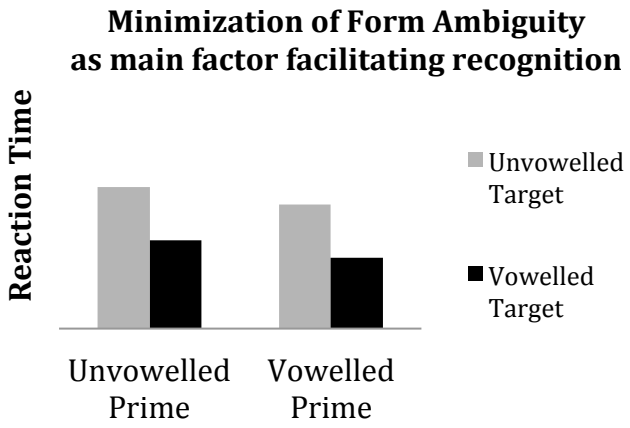
Minimization of Form Ambiguity as main factor facilitating recognition

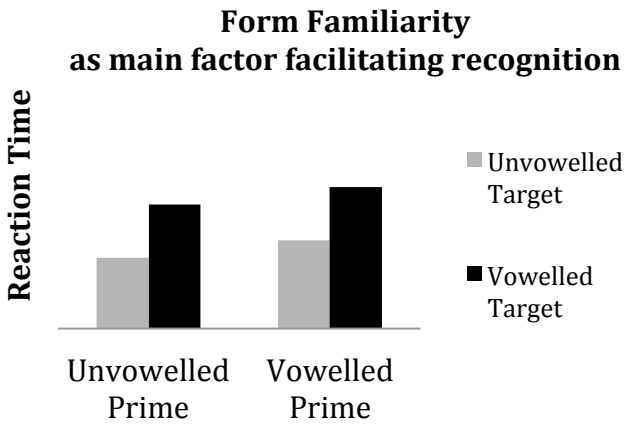


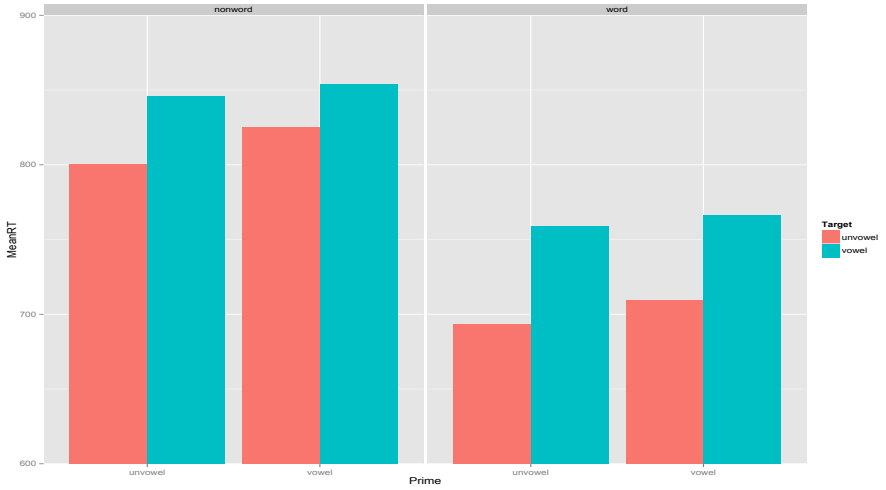
Form Familiarity as main factor facilitating recognition



Results


Minimization of Form Ambiguity
as main factor facilitating recognition
Reaction Time
Unvowelled Target
Vowelled Target
Unvowelled Prime
Vowelled Prime


Form Familiarity
as main factor facilitating recognition
Reaction Time
Unvowelled Target
Vowelled Target
Unvowelled Prime
Vowelled Prime


nonword
word
900
800
700
600
MeanRT
Target
unvowel
vowel
unvowel
vowel
unvowel
vowel
Prime

Discussion

- *Form familiarity* determines how fast Arabic readers can recognize visual word forms.
  - Perhaps voweling overtly engages the orthography-phonology interface, which takes longer?

- *Form ambiguity* does not slow Arabic readers even when reading isolated visual word forms.
  - Familiar forms have a default interpretation?
    - Example here?

Conclusion

- Full voweling, even if it can reduce form ambiguity, slows Arabic readers down in reading.
- *Form familiarity* seems to be the determining factor in how fast readers can recognize a visual word form.

- Experiments in Arabic: use full voweling only when absolutely necessary if natural automatic linguistic processes are being investigated.

جامعة الإمارات العربية المتحدة
United Arab Emirates University

جامعة نيويورك ابوظبي
NYU | ABU DHABI

Thank you!

Any questions?