Deducing Grammar from Parsing: Self-Paced Reading and Modeling Evidence from Arabic

Matthew A. Tucker*, Ali Idrissi†, & Diogo Almeida*

*New York University Abu Dhabi, †United Arab Emirates University

Introduction

- Much theorizing about agreement dependencies comes from Agreement Attraction errors:

(1) [NP The key [PP to the cabinets]] are on the table. (Bock & Miller, 1991)

- Agreement attraction is sensitive to morphological markedness and therefore is influenced by morphological representation (Bock & Miller, 1991, *et seq.*)
- Occurs in both production (sentence completion) and comprehension (self-paced reading, eye-tracking reading, EEG, MEG)
- **However:** *All* such studies are confined to Indo-European languages
- Thus, all results are confined to similar inflectional paradigms and similar pluralization types
- Modern Standard Arabic (Arabic) provides a nice place to expand the empirical base:
  - Large amount of inflectional morphology on verbs (gender, number, person)
  - Large amount of morphology on nouns (case, number)
  - Wide array of "irregular" plural types which involve non-concatenative, abstract morphemes representing plurality

Experiment 1 — Design

Subjects:

- 84 native Arabic speakers (83 female; mean age 21.4 years)

Stimuli:

- 48 item sets of the form:
  NP Subj – Complementizer – RC Verb – NP Attr – Adv/PP – Verb – Continuation
- Adverb inserted to avoid Attr spillover effects (Wagers, *et al.*, 2009)
- Systematically manipulated for:
  - Attractor Number: Singular, Plural (Attr) (AttrNum)
  - Verb Number: Singular, Plural (Verb) (Gram)
  - Attractor Gender: Masculine, Feminine — counterbalanced across all items.
- Diacritics only used for lexical disambiguation; short-vowel case markers not written
- All subjects singular; example item appears in (2):

(2) a. المترجم الذي ساعد الرئيس احيانا يتكلم خمس لغات بفصاحة.

b. ʔal-mutarʒim-u ʔallaðii saaʕad-a ʔal-raʔiis-a ʔaħjaanan
the-translator-NOM comp.masc.sg helped-3.sg.masc the-president-acc often
ja-takallamu xamsata luɣaat-in bi-fasˤaahatin.
3.sg.masc-speaks five languages-acc with-fluency
"The translator who helped the president(s) often speak(s) five languages fluently."

- Details about gender balancing:
  - Feminine Suffixing Plurals: formed by suffixation ($t^{\varsigma}aaliba$ – $t^{\varsigma}aalib$-*aat*, "student(s) (fem.)")
  - Masculine Ablauting Plurals: formed by ablaut/vowel-change (*fajx* – *fujuux*, "sheikh(s)")
- Gender co-varies with plural types because of grammatical properties of Arabic (Ryding, 2005): case is orthographically marked on masculine suffixing plurals; there are very few feminine ablauting plurals
- Four conditions (24 masculine, 24 feminine):

Grammatical Conditions	Ungrammatical Conditions
SS singular attractor, singular verb (*The key to the cabinet. . . is*)	SP singular attractor, plural verb (*The key to the cabinet. . . are*)
PS plural attractor, singular verb (*The key to the cabinets. . . is*)	PP plural attractor, plural verb (*The key to the cabinets. . . are*)

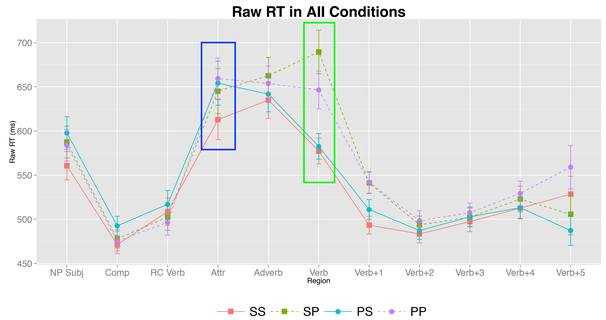
Procedure & Analysis:

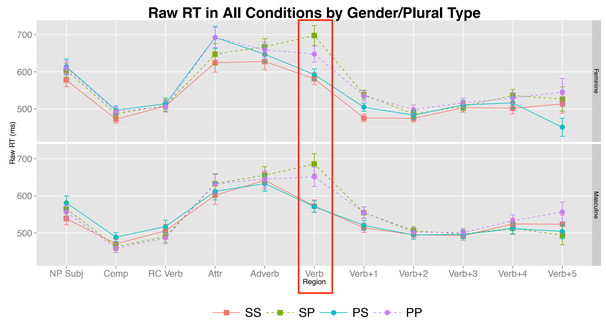
- Self-paced word-by-word moving window procedure using Linger software (Doug Rohde, MIT)
- Every item followed by a comprehension question (with feedback)
- 10% Winsorization of outliers by region and condition (not by subject)
- Mixed-effects model fitted with experimental variables, orthographic length, and 3 previous regions

Predictions:

- Main effect of Gram in verb region and spillover regions (ungrammatical > grammatical)
- Interaction of Gram × AttrNum in verb and spillover regions (Sg/Ungram > Pl/Ungram)
- Main effect of AttrNum in Attr region (Pl > Sg; Wagers, *et al.*, 2009)
- *No* impact of Attr plural type (Bock & Eberhard, 1993)

Experiment 1 (w/Gender) — Results

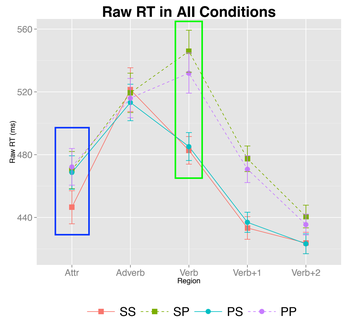


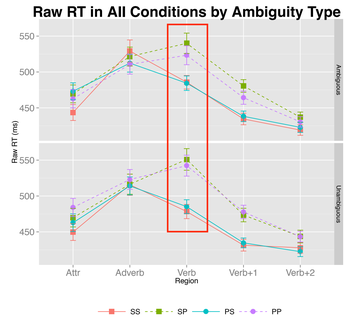


Experiment 2 — Design

- **A possible confound**: broken plurals are a heterogeneous class.
- These plurals can be characterized by whether their CV-pattern is "ambiguous" (used for both singular and plural) or "unambiguous" (used only for plural)
- Ambiguous: $C_1uC_2uuC_3$, with both singulars and plurals (*duxuul*, "entering (n.)" & $lus^{\varsigma}uus^{\varsigma}$, "thieves")
- Unambiguous: $C_1uC_2aaC_3$, only with plurals ($t^{\varsigma}ulaab$, "students")
- **Experiment 2:** replace gender manipulation with direct manipulation of ambiguity.
  - Counterbalanced for ambiguity of attractor's plural template.
  - 24 ambiguous, 24 unambiguous, in a variety of CV-patterns.
- 110 subjects from UAEU (110 females; mean age 21.0 years) in identical methodology.

Experiment 2 (w/Ambiguity) — Results





Discussion

Experiment 1:

- **Attraction:** PS read more slowly than PP at verb ($\beta = -66.31; p = 0.0002$)
- **Plural NP effect:** PX conditions read slower than SX conditions at Attr ($\beta = 31.48; p = 0.10$)
- **Gender effect:** Attraction *much* stronger in feminine conditions (fem: $\beta = -75.77; p = 0.0004$; masc: $\beta = -3.71; p = 0.87$)

Experiment 2:

- **No attraction:** No significant difference between PS and PS ($\beta = -6.80; p = 0.57$)
- **Plural NP effect:** PX conditions read more slowly than SX conditions ($\beta = 23.05; p = 0.005$)
- **No ambiguity effect:** lack of attraction not mediated by ambiguity of attractor ($\beta = -1.25; p = 0.47$)

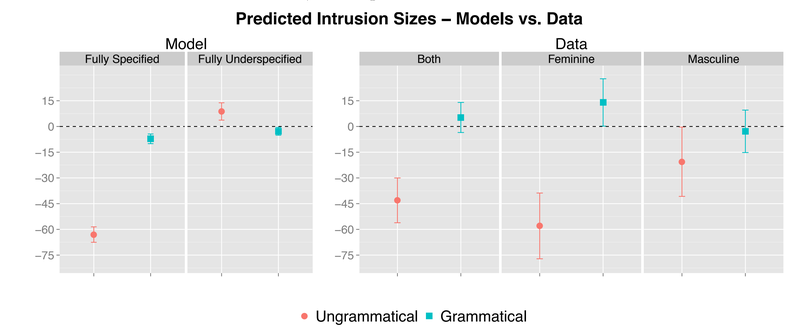
Modeling

Question: What is the representational requirement for modeling these results?

- Examine featural representation's impact on attraction effects using act-r (Lewis & Vasishth, 2005).
- Model results from 10,000 Monte Carlo runs using reasonable default model parameters.
- Dependent measure is the difference between PS and PP conditions — the *size of the intrusion/attraction effect*.

Models:

- Fully Specified: number is a binary cue with values [sg], [pl]
- Fully Underspecified: number is lexically underspecified and absent on some nouns



Underspecification is a reasonable model-theoretic approach to the reduction in attraction.

Concusions, *etc.*

Conclusions:

- Agreement attraction does occur in Arabic. . .
- . . . but suffixing feminines attract more than ablauting masculines
- . . . and this is not due to lexical ambiguity.

Agreement is about form, not meaning, matching.

Future Directions:

- Does [fem] attract like [pl]?
- Can we dissociate gender from plural type? Suffixing masculine plurals. . .
- Arabic has [dual]: how is a three-way number system represented?
- What about the influence of optional orthographic case?

Thanks & Selected References

Acknowledgments — Thanks to Eias Al-Daman, Souad Al Helou, Meera Saeed Al Kaabi, Esma Mansouri, and Samer Nehme for assistance with stimuli creation and Tommi Leung for help with participant recruitment. Thanks also to Stephen Politzer-Ahles, Joseph King, Kevin Schluter, Matt Wagers, and Shravan Vasishth for comments on this project. Thanks to Richard Lewis for sharing his ACT-R code and Dave Kush and Brian Dillon for help with modeling.

Selected References — Bock, K., & Miller, C.A. 1991. Broken agreement. *Cognitive Psychology* 23:45–93. Lewis, R.L., & Vasishth, S. 2005. An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science* 29:375–419. Ryding, K.C. 2005. *A Reference Grammar of Modern Standard Arabic*. Cambridge UP. Wagers, M.W., Lau, E.F., and Phillips, C. 2009. Agreement attraction in comprehension: Representations and processes. *JML* 61:206–237.

✉ matt.tucker@nyu.edu • http://files.nyu.edu/mat11/public/