

**The complex structure of errors and the independent visibility of ϕ -features:
Deriving grammatical and markedness asymmetry effects in agreement attraction**

Diogo Almeida (NYU Abu Dhabi) & Matthew A. Tucker (Oakland University)

AGREEMENT ATTRACTION, the systematic pattern of production and perception errors that occur in sentences containing mismatches in ϕ -features (generally number) between their subjects and verbs (e.g., *The key to the cabinets were...*), is interesting because it suggests an unexpected split in how the parser deals with ungrammatical structures of different types ([3, 8]). Most accounts posit a *single locus* for this phenomenon, either at the representational ([4, 5]) or process level ([1]), leading to the following corollaries: (1) simple agreement failure (i.e., ungrammatical agreement in the absence of attraction) gives rise to a uniform error response by the parser, a modulation of which (in either frequency or strength) creates the attraction effect, and (2) all ϕ -features are presumed to be equipotent in computing agreement, and therefore are expected to be subject to attraction in similar ways. We present data from two large self-paced reading (SPR) experiments in Modern Standard Arabic that challenge these assumptions.

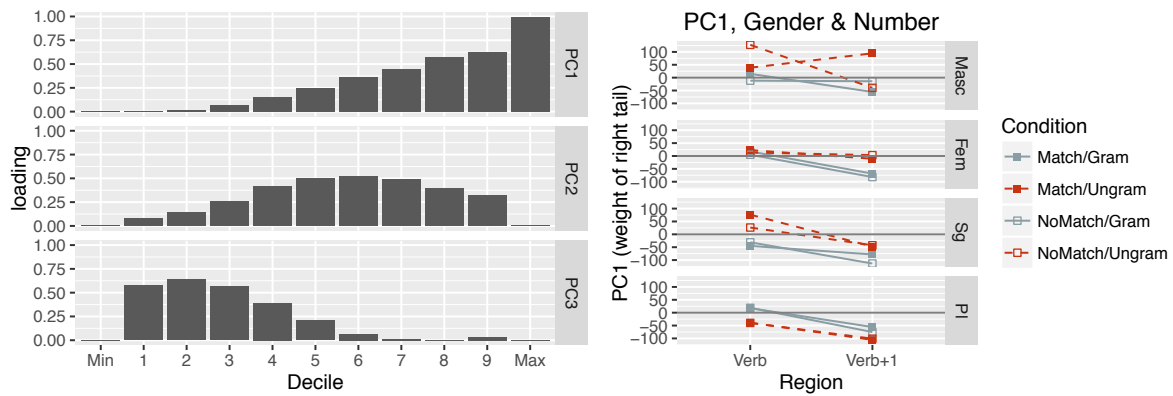
Design We manipulated agreement attraction using 48 sentences with the relative clause construction in (1) (where the *subject* is in italics, **attractor** in bold, and **critical verb** in both; all followed by a semantically plausible continuation):

(1) *ʔal-mutarzim-u* ʔallaḏii saaʕad-a **ʔal-muḏiir-a** ʔahjaanan **ja-takallamu**
the-translator-NOM COMP helped-3MS **the-manager-ACC** often **speaks.3MS**

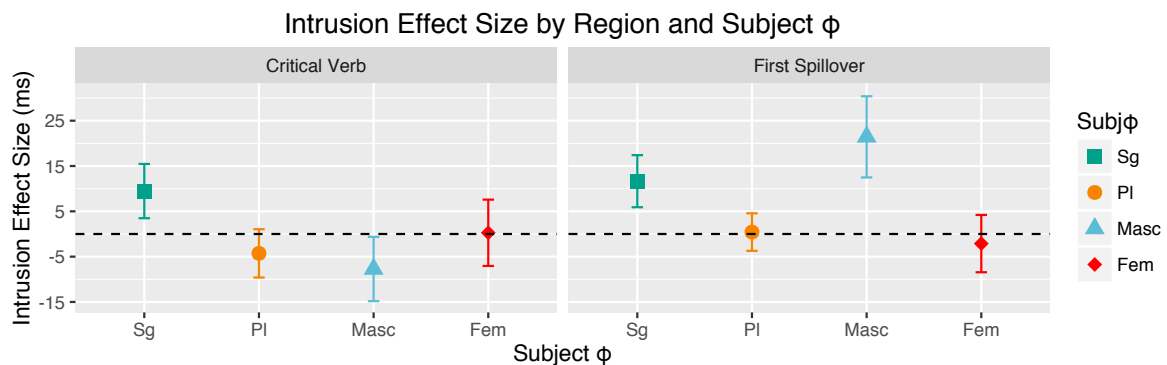
“The translator who helped the manager often speaks ...”

Both experiments manipulated: (1) the ϕ -feature of the subject (SUBJ ϕ), (2) whether the subject and attractor matched in ϕ -features (MATCH), and (3) whether the subject and verb properly agreed (GRAMMATICALITY). In the classic agreement attraction configuration, subject and verb do not agree (UNGRAMMATICAL) and the subject and the attractor do not match (NOMATCH). Exp. 1 manipulated number (SG, PL) and Exp. 2 manipulated gender (MASC, FEM). To allow for reaction-time (RT) distributional analysis, we collected large samples (N = 330) in each.

Assumption 1: Simple agreement errors are uniform. The analysis of SPR data assumes changes only in the means of the RT distributions. However, recent evidence has shown that attraction effects involve distributional shifts on the right tails ([6] and references therein). Here we follow this line of inquiry by capitalizing on Principal Components Analysis (PCA) to summarize the shape differences of RT deciles. As Fig. 1 shows, the first principal component expresses the weight of the right tail. If ungrammatical agreement RT slowdowns reflected a lengthening in the right tail, agreement attraction might be expected to reduce it, and that is what we find: agreement attraction effects are characterized by right tail differences between ungrammatical MATCH sentences (heaviest right tails) versus NOMATCH sentences (second heaviest tails). Crucially, however, the latter pattern was restricted to cases in which the subject had an *unmarked* ϕ -feature value (i.e., SG or MASC). Ungrammatical sentences with subjects with *marked* ϕ -feature values (PL, or FEM) failed to exhibit longer-than-average right tails (and, potentially as a consequence, attraction effects), *even in simple ungrammatical agreement without suitable attractors* (MATCH/UNGRAM). These results suggest that simple agreement errors do *not* have a uniform impact in processing—some cases of ungrammatical agreement may drive shifts in the RT distribution and others may drive changes to its shape—which may in turn derive the apparent sensitivity of agreement attraction to morphophonological properties of the head noun (the markedness effect).



Assumption 2: ϕ -features are all equipotent in the processing of agreement. One key debate in formal linguistics which relates to discussions of homogeneity in agreement attraction is whether ϕ -features are bundled together for the purposes of agreement (e.g., [7]) or distinctly active in syntax (e.g., [2]). Alongside this debate in psycholinguistic work is the prediction in processing models such as the cue-based models of [1] that agreement cues are equally salient in working memory retrieval events. Taken together, bundling models of formal syntax and cue-based models of parsing predict that agreement features should be equally available for attraction in equal measure. To assess this, we computed INTRUSION EFFECT SIZES for both gender and number as well as Cohen's d effect size estimates by subtracting mean RTs to NOMATCH/UNGRAM sentences from MATCH/UNGRAM sentences within each experiment. Positive values equate to attraction effects in the RT distributions.



The results of the quantitative analysis and inspection of the intrusion sizes show that gender and number only attract when the subject is unmarked ([5]). Within those unmarked subjects, number attracts in the critical verb ($d = 0.06$) and first spillover ($d = 0.09$), whereas gender attracts only in the first spillover region ($d = -0.05$ at the verb versus $d = 0.12$ postverbally). These results suggest that these two ϕ -features do not attract in equal measure, in either space or time.

Conclusions (1) Consistent with prior work we find that attraction arises due to a modulation of the length of the right tail of the RT distribution, but we also find that (2) the lengthening of the right tail associated with ungrammaticality in simple (no attraction) agreement errors is dependent on the subject having an unmarked ϕ -feature value. (3) Both gender and number participate in illusory subject-verb agreement licensing, but they exhibit non-trivial differences in the timing and amount of illusory licensing. These results present a challenge for existing psycholinguistic theories of agreement, which posit a *single locus* for agreement attraction effects. In their place, we take our results to indicate that illusory agreement licensing is grammatically dependent insofar as it differentially accesses markedness and ϕ -feature values provided by the representational syntax. Error-driven cue-based retrieval proposals (e.g., [1, 8]), which account for the grammatical asymmetry in attraction by positing that cue-based retrieval errors only occur in ungrammatical sentences, may be amended to include another gating error component, namely, the markedness status of the ϕ -feature of the subject. This simple account naturally derives **both** the grammatical and the markedness asymmetries observed in the agreement attraction literature, but it does so by suggesting that agreement errors are not uniform but instead depend on the specific ϕ -features involved, and can interact with other linguistic processes (like cue-based retrieval) *independently* – a result which is

compatible with the differential visibility of ϕ -features observed in our results and the formal syntactic literature on agreement relations.

References [1] **Badecker, W. & Kuniak, F.** 2007. Morphology, agreement and working memory retrieval in sentence production: Evidence from gender and case in Slovak. *JML*, 56(1):65–85. [2] **Béjar, S.** 2003. *Phi-Syntax: A Theory of Agreement*. PhD thesis, University of Toronto. [3] **Bock, K. & Miller, C.** 1991. Broken Agreement. *Cog. Psych.* 23(1):43–93. [4] **Bock, K. & Middleton, E. L.** 2011. Reaching agreement. *NLLT*, 29:1033–69. [5] **Franck, J., et al.** 2008. The interplay of syntax and form in sentence production: A cross-linguistic study of form effects on agreement. *LCP*, 23(3):329–374. [6] **Lago, S., et al.** 2015. Agreement in Spanish comprehension. *JML*, 82:133–49. [7] **Pollock, J.-Y.** 1989. Verb movement, universal grammar, and the structure of IP. *LI*, 20:365–424. [8] **Wagers, M., et al.** 2009. Agreement attraction in comprehension: Representations and processes. *JML*, 61(2):206–237.