Attraction Effects for Verbal Gender and Number Are Similar but Not Identical: Self-Paced Reading Evidence from Modern Standard Arabic

Matthew A. Tucker*,a, Ali Idrissib, and Diogo Almeidaa

*Corresponding Author. Email: matt.tucker@nyu.edu

^aLanguage, Mind, and Brain Laboratory

Division of Science

New York University Abu Dhabi

P.O. Box 129188

Abu Dhabi

UAE

^bDepartment of English Literature and Linguistics

Qatar University

P.O. Box 2713

Doha

State of Qatar

Highlights:

- Agreement attraction errors in comprehension also affect verbal gender agreement morphology independent of other features/cues.
- Agreement attraction occurs in languages with rich verbal inflectional morphology.
- Errors for verbal gender are likely quantitatively and qualitatively different from errors for verbal number.
- Arabic does not show attraction effects in acceptable sentences.

Manuscript submitted for publication.

Attraction Effects for Verbal Gender and Number Are Similar but Not Identical Self-Paced Reading Evidence from Modern Standard Arabic

Matthew A. Tucker^{a,*}, Ali Idrissi^b, Diogo Almeida^a

^aLanguage, Mind, and Brain Laboratory, Division of Science, NYU Abu Dhabi,
 P.O. Box 129188, Abu Dhabi, United Arab Emirates
 ^bDepartment of English Literature and Linguistics, College of Arts and Sciences,
 Qatar University, P.O. Box 2713, Doha, State of Qatar

Abstract

Previous work on the comprehension of agreement has shown that, in many languages, incorrectly inflected verbs do not trigger responses typically seen with fully ungrammatical verbs when the preceding sentential context furnishes a possibly matching distractor noun. Four studies are presented which tested the character and timing of these errors in comprehension along the dimensions of grammatical gender and number in Modern Standard Arabic. Despite a robust verbal gender system which interacts with other inflectional features, Arabic readers show agreement attraction effects in reading comprehension for gender and number on verbs given appropriate preceding contexts with mismatching NPs. However, we also observe that these two features do not behave identically either qualitatively or quantitatively. Qualitatively, attraction for gender is not subject to markedness considerations which lead to asymmetries in attraction relative to feature combinations in the preceding NPs, yet number is subject to these asymmetries. Quantitatively, the size of gender effects are smaller and later relative to number attraction effects. These results are shown to require changes to representation and process theoretical models of agreement attraction. We also discuss how models of agreement errors require modifications in order to account for these differential results.

^{*}Corresponding Author

Email addresses: matt.tucker@nyu.edu (Matthew A. Tucker), ali.idrissi@gmail.com (Ali Idrissi), diogo@nyu.edu (Diogo Almeida)

1. Introduction

10

11

- Human language, despite being transmitted serially as a string of words, contains a myriad of relationships between words which can obtain at a distance. Subject–verb agreement as in (1) is one such relationship:
- (1) a. **The fencers** *are* divided about the best strategy for the World Championships.
 - b. **The fencers** on the French National Team that won a major award last year *are* deeply divided about the best strategy for the World Championships.
 - c. The fencers on the French National Team that won a major award last year by beating the Italian team in a very hotly contested and important match are deeply divided about the best strategy for the World Championships.

In the specific case of (1a), the choice of the fencers conditions the subsequent choice 12 of are in production or the expectation of a plural verb in comprehension. Subject-verb 13 agreement is particularly important in the study of language and its relationship with the 14 performance systems since it not only involves the very basic building blocks of a clause but also because it is a relationship that can obtain at an unbounded serial distance. This is because subjects can theoretically be separated from their verbs by an infinite amount 17 of material yet still require proper agreement — see (1b,c). Despite this potentially 18 infinite linear distance, subjects and verbs are still relatable by dependency because of 19 their structural positions in the clause. This basic fact underscores an important property of the syntax of human languages: despite of their linear externalization, sentences are 21 internally organized in a hierarchical, and not serial, fashion. 22

Therefore, from the perspective of real–time language production and comprehension, coping with potentially unbounded dependencies such as subject–verb agreement
requires attention to the encoding, maintaining, and retrieving of linguistic units from
working memory, as well as the monitoring process that oversees whether the correct
relationship between the subject and the verb has been completed. It is a remarkable
fact, then, that subject–verb agreement errors are not only sometimes observed (both in
language production Bock & Miller, 1991 and comprehension Pearlmutter et al., 1999;

- Wagers et al., 2009), but that they also seem to be at least partially systematic. Known as AGREEMENT ATTRACTION, a particularly well-studied subset of these errors are commonly seen when a subject co-occurs with a non-subject argument that appears to be the target of the erroneous number agreement, as in the example in (2) from Dillon et al. (2013):¹
- The executive who oversaw **the middle managers** apparently *were* dishonest about the company's profits.

 (Dillon et al., 2013)

The characteristic property of this phenomenon is the illusion of acceptability of *prima*facie unacceptable agreement violations — despite the fact that the plural were is ungrammatical in (2), many speakers occasionally both accept and produce such utterances. In production studies such as Bock & Miller (1991) or Franck et al. (2002),
these errors surface as incorrect verb productions, whereas in comprehension studies
such as Pearlmutter et al. (1999) or Tanner et al. (2014), these errors surface as the
absence of behavioral or electrophysiological responses typically associated with the
perception of ungrammaticality. Because they represent a systematic exception to the
idea that processing is faithful to grammar during the production and comprehension
of dependencies, these attraction violations have served as a focal point for much theorizing about the nature of both grammatical agreement and dependency processing.

49 1.1. Failure of Representation or Failure of Process?

One theoretical approach to illusory dependency licensing in the literature conceives of attraction effects as arising as a function of the dynamics of memory encoding and retrieval of agreement–relevant material in memory. This line of research draws on cue–based retrieval theories of language processing (Lewis & Vasishth, 2005) and extends their logic to agreement dependencies (Badecker & Kuminiak, 2007; Badecker & Lewis, 2007; Wagers et al., 2009). This theory, building upon the observation that dependency resolution is subject to retrieval interference, posits that sentence process-

¹Here the (correct) subject appears in italic face, the attractor/distractor NP in bold face, and the target region in both bold and italic.

ing contains instances of working memory retrievals which access long-term memory stores in a parallel, cue-based manner. It is therefore a theory in which the probability that memory chunks are retrieved is a function of the similarity of a given chunk to other items in memory as well as the number of dimensions upon which a chunk matches the cues in the goal of the retrieval event. When more than possible retrieval target matches the goal cues, erroneous retrievals of non-subjects can occur. For explicit modeling of agreement in this system, see Badecker & Kuminiak (2007); Badecker & Lewis (2007); Dillon et al. (2013); Wagers et al. (2009); and Tucker et al. (2015), but what all these models have in common is the notion that agreement attraction is a failure of *process* in the memory retrieval system underwriting language use.

A theory which is often contrasted with the cue-based retrieval model is one in 67 which structural representations themselves can be erroneously represented (see Bock & Eberhard, 1993; Eberhard et al., 2005; Franck et al., 2008; Nicol et al., 1997; Pearlmutter et al., 1999 and the discussion in Engelmann et al., 2015b; Wagers et al., 2009). 70 In these models, stochastic fallibility in the encoding or maintenance of structural rep-71 resentations leads to misrepresentation of the true subject's features in terms of relevant 72 features of the distractor, resulting in attraction through the normal processes of subjectverb agreement in the proportion of cases where this erroneous representation obtains. 74 One way to formalize this notion is to say that the representations responsible for main-75 taining features and syntactic constituents in memory allow for targets of agreement to have their features overwritten in the presence of a distractor with mismatching features. 77 Proponents of these models have advanced several distinct mechanisms for achieving this misrepresentation including degradation of structural representation (Eberhard et al., 2005; Staub, 2009), erroneous feature percolation (Nicol et al., 1997), and fallible 80 feature copying (Franck, 2011; Franck et al., 2008). Here we abstract away from con-81 siderable detail to jointly consider these models as sharing the notion that agreement 82 attraction is a failure of representation in the processing systems subserving language use.

Despite these differences, there is at least one important dimension along which these two families of theories are similar, namely the way in which they incorporate cross-linguistic differences pertaining to the process of subject—verb agreement. Lin-

guistic theory generally takes person, number, and gender features to be equipotent in 88 agreement phenomena (e.g., Chomsky, 1995; Pollock, 1989; Preminger, 2011; though see Béjar, 2003; Béjar & Rezac, 2009 for a different approach), and processing theories which rely directly upon these linguistic notions predict (at least limited) equivalency in attraction effects for each of these features without having to resort to nontrivial exten-92 sions. Because misrepresentation theories attribute attraction to normal mechanisms of 93 feature spreading, differences in attraction strength for different features are only predicted if representational considerations constrain spreading, overwriting, or copying. Cue-based models, on the other hand, posit that cues are typically treated equally by the retrieval system. Any observed difference between how different agreement cues are 97 processed would necessitate positing a more complex view of these cues or how they ٩R are weighed or retrieved within the memory system.

With this backdrop, it is therefore important to determine whether the basic assumption shared by the two most popular families of theories of agreement errors — namely 101 that all agreement features are equipotent — is in fact supported by the evidence. While 102 number agreement comprehension is relatively well-studied, considerably less work has 103 been conducted to address gender verbal agreement and the extant data do not provide 104 any consensus on the matter in either production or comprehension (which we discuss 105 below). This is an important issue for both representational and cue-based theories, 106 since both would, all else being equal, take verbal gender to be equivalent to verbal 107 number in attraction terms. In a series of four comprehension studies in Modern Stan-108 dard Arabic, we directly test these foundational assumptions by comparing the process of subject-verb gender agreement with subject-verb number agreement, in an effort to 110 document the ways in which these two processes are similar (as predicted or assumed 111 by the two existing theoretical approaches to agreement errors) and the ways in which 112 they are different (and therefore the ways in which they would challenge these theories 113 to revise their assumptions). 114

1.2. Equivalency in Error Studies

115

116

Given the preceding discussion it is important, theoretically, to understand whether attraction for verbal gender agreement is identical to attraction for verbal number agree-

ment. But how does one assess equivalency between gender and number in subject-118 verb agreement comprehension? Here we identify five distinct ways in which gender 119 and number could be equivalent in regards to comprehension attraction effects:

a. EXISTENCE: Do both features participate in attraction?

120

121

122

123

124

125

127

128

- b. SIZE: Do both features yield similar attraction effect magnitudes?
- c. Grammatical Asymmetry: Do both features participate in asymmetries based on grammaticality of the verb?
- d. MARKEDNESS ASYMMETRY: Do both features participate in asymmetries based on markedness of the agreeing elements?
- e. TIMING: Do both features exhibit attraction effects with the same timecourse?

An evaluation of the extant literature reveals that the answers to these questions 129 are muddled when it comes to verbal gender. With respect to the existence of attrac-130 tion effects and their size (3a, b), Lorimor et al. (2008) reported no gender attraction in 131 production in Russian, whereas Badecker & Kuminiak (2007) and Malko & Slioussar 132 (To Appear) report gender attraction respectively in production in Slovak and in per-133 ception in Russian, but without assessing its comparative effect with the one observed 134 for number. Interestingly, verbal gender agreement in Slavic languages, where present, 135 is restricted only to past tense verbs; verbs in other tenses do not show gender agree-136 ment at all. While gender can appear on a verb in Slavic, gender in the nominal system 137 is often conflated with case morphology, an issue which leads to ambiguity that has been known to influence agreement attraction rates and sizes (Badecker & Kuminiak, 2007; Hartsuiker et al., 2003; Häussler & Bader, 2009). Deutsch & Dank (2009, 2011) 140 directly compare gender and number attraction effect sizes in a series of Hebrew pro-141 duction studies and find similar rates of attraction which are nevertheless subject to 142 different morphophonological influences, though only inanimate NPs were tested. It therefore seems clear that verbal gender could in principle be subject to attraction, but whether it is of the same size as the effects for verbal number remains to be determined. 145 In comprehension studies of subject-verb number agreement attraction, it is often 146 (but not always, see Engelmann et al., 2015b for review) observed that these effects are

asymmetric in nature: attraction effects are easily observable in ungrammatical sen-148 tences, but less prominently so in grammatical ones (Wagers et al., 2009; though see 149 also Franck et al., 2015 and Engelmann et al., 2015b). In addition, Tanner et al. (2014) have found that ERPs to English attraction configurations do not show evidence of mor-151 phosyntactic error recognition (defined as an observable P600 effect) in grammatical 152 examples. If agreement processing relies on the same structural representations or the 153 same structure of memory cues regardless of the agreement features, then one would 154 predict that a similar asymmetry would be found for gender subject-verb agreement as well. As far as we are aware, no studies directly address the presence of grammaticality 156 asymmetries (3c) between gender and number subject-verb agreement, though exam-157 ination of the results in Malko & Slioussar (To Appear) show that a grammaticality 158 asymmetry is present with Russian verbal gender. As Wagers et al. (2009) note, com-159 prehension studies assessing the grammaticality asymmetry are particularly important in this regard as (i) the traditional preamble completion task generally used in produc-161 tion studies Bock & Miller (1991) cannot provide evidence about attraction effects in 162 ultimately grammatical productions (though see Staub, 2009 for a variation argued to 163 be able to do this), and (ii) that cue-based retrieval models seem to be better able to ac-164 commodate this effect when compared to misrepresentation models, thereby providing 165 a potential empiric diagnostic that favors one family of theories over the other (Engel-166 mann et al., 2015b; Tanner et al., 2014; Wagers et al., 2009). 167

Identity of markedness (3d) is the best-understood of the five dimensions of comparisons along which number and gender subject—verb agreement can be evaluated, and it is the topic of considerable discussion (see Badecker & Kuminiak, 2007; Deutsch & Dank, 2011; and Malko & Slioussar, To Appear). By "markedness," here we mean the asymmetry observed by Bock & Miller (1991) wherein certain feature combinations in pre-verbal material cause more attraction errors than others (as in English, where *The key to the cabinets...* induces many more attraction errors than *The keys to the cabinet...*; see also Eberhard, 1997). The Slovak and Russian studies (Badecker & Kuminiak, 2007; Malko & Slioussar, To Appear) advance the conclusion that markedness in a three-valued system with masculine, feminine, and neuter should be defined in a pairwise fashion such that there is no *gestalt* markedness hierarchy but rather individual

168

170

171

172

173

176

177

relations between masculine, feminine, and neuter. However, this is at odds with results from Russian number, as Malko & Slioussar (To Appear) note, since the results of the latter suggest that number attraction profiles are not prevalent for plural subjects with singular distractors (*i.e.*, there is no markedness asymmetry). In Hebrew, on the other hand, markedness has been shown to affect production of number attraction errors but not gender (Deutsch & Dank, 2011). To our knowledge, there are no studies assessing the phenomenon of markedness in comprehension other than Malko & Slioussar (To Appear), with all other data coming from production studies.

Finally, it is also possible to ask whether or not gender and number attraction effects occur in similar time-courses (3e) — both with respect to global grammaticality effects (Lago et al., 2015; Staub, 2009, 2010) and each other. Given the preponderance of production studies in the existing literature, this question has not been properly addressed to date.

187

188

189

190

191

209

It is also worth noting that much research in the literature on attraction has attempted 192 to examine the locus of morphophonological influences (including markedness) on at-193 traction rates to one or more of the constituents involved in the attraction process — 194 the head noun/subject, the local noun/distractor, or the verb (e.g., Dank & Deutsch, 195 2010; Deutsch & Dank, 2011; Hartsuiker et al., 2003; Malko & Slioussar, To Appear; 196 Vigliocco et al., 1995). This is a fruitful line of inquiry, but one which is orthogonal to a 197 more basic concern: are there asymmetries in attraction rates to different combinations 198 of features on heads, local nouns, and verbs? Whether these effects are attributable to 199 the influence of the head or local noun per se is impossible to evaluate in a binary system like that found with English number or Arabic gender. Thus, while it is sensible to 201 talk about the influence of markedness or morphophonological ambiguity on only the 202 head or local noun, in trying to evaluate whether there is a basic equivalence among 203 agreement features, one must first establish whether any asymmetry is present based 204 on morphophonological or conceptual features before one can examine the importance of the locus of these features. We therefore retain the term "markedness" from the literature for these asymmetries but attempt, where possible, to abstract away from tying 207 the asymmetry to particular morphosyntactic positions. 208

In the same vein, it is worth noting that grammatical gender does appear as the focus

of a large number of agreement attraction studies, such as Antón-Méndez et al. (2002); 210 Dank & Deutsch (2010); Deutsch & Dank (2009, 2011); Franck et al. (2008); Vigliocco 211 & Franck (1999, 2001); Vigliocco & Zilli (1999); and Fuchs et al. (2015), to name just a few, but these works examine adjective-noun gender co-occurrence restrictions. In 213 the nominal domain, the combined results of these studies seems to support the idea 214 that attraction in gender nominal agreement proceeds in similar ways as attraction in 215 number nominal agreement. We take this point to be well-established in the production 216 literature but note that there is little reason to treat verbal and nominal agreement as 217 being, in principle, the same sort of process (for theoretical discussion, see Baker, 2008 218 and Norris, 2014; for a discussion within the psycholinguistic literature, see e.g., Tanner 219 et al., 2014). Most crucially for our purposes here, while subject-verb agreement is 220 potentially unbounded, adjective-noun agreement is by definition extremely local and 221 not a potentially unbounded dependency.

1.3. The Present Study: Context

223

The experiments reported here attempt to address several outstanding questions 224 about the equivalency of verbal gender and number with respect to the five dimensions 225 laid out above: whether they both elicit attraction effects of the same size and time 226 course, whether they do so asymmetrically with respect to language-internal marked-227 ness considerations, and whether they do so asymmetrically with respect to grammat-228 icality. In following this reasoning, our language of study, Modern Standard Arabic 229 (MSA), provides several important desiderata for studies of verbal gender (Ryding, 2005): (1) the presence of verbal gender agreement on all verbs in the language (not 231 restricted to a given tense, as in Slavic); (2) the appearance of gender marking on nom-232 inals independent of case morphology (also unlike Slavic), allowing the examination 233 of gender independently of the influence of case; (3) a demonstrated number attrac-234 tion effect in comprehension against which to compare results from gender (Tucker et al., 2015); and (4) a close typological relationship to Hebrew, allowing comparison 236 of our results with the production studies of Deutsch & Dank (2009, 2011), and Dank 237 & Deutsch (2010). Finally, the last two experiments simultaneously attempt to repli-238 cate and expand upon findings from Tucker et al. (2015) that the type of plural marking

on attractor NPs matters for agreement attraction effect sizes. These latter two experiments also provide a number contrast to the gender effects reported in Experiments 1–2 in order to assess similarities and differences in attraction rates, asymmetries, and time courses.

244 2. Experiment 1

264

265

268

Experiment 1 was designed to assess whether attraction for grammatical gender is 245 possible in MSA comprehension using a self-paced reading methodology. In the ex-246 perimental agreement attraction literature, several grammatical structures are used with some regularity: (i) PP modifiers (Bock & Miller, 1991), (ii) object relative clauses (Bock & Miller, 1991; Wagers et al., 2009), and (iii) infinitival complements (Tucker 249 & Wagers, 2010). In this study, however, we employ subject relative clauses (see, e.g., 250 Bock & Miller, 1991; Dillon et al., 2013) modifying a sentence-initial subject. One 25 Arabic-internal and one Arabic-external consideration each guide the choice of subject 252 relative clauses. Firstly, Wagers et al. (2009) have shown that spillover effects in agreement attraction studies can inadvertently impact measurements at critical verbs when 254 the immediately previous region is manipulated experimentally. A common solution to 255 the need to manipulate the features/cues of an attractor is to insert an adverb between 256 the attractor and critical verb. In Arabic, adverbs are not commonly found preverbally in Subject — Verb — Object word orders (Tucker, 2011). Adverbs and adverbial PPs 258 are preferred clause-finally, making subject relative clauses an appropriate choice given 259 the availability of a relative-clause final parse of adverbs appearing preverbally. Sec-260 ondly, number agreement attraction in this configuration has already been studied in 261 MSA (Tucker et al., 2015), allowing for direct comparison of the reaction time profiles of grammatical number and gender processing. 263

Given the prevailing theoretical and experimental conceptions of agreement, one expects to find that attraction should be possible for gender. In the formal syntactic literature, agreement is often taken to be a uniform process which simultaneously encompasses the features of gender, number, and person (to wit, the theory of Agr projections originating from Pollock, 1989 and the more modern notion of AGREE from Chomsky,

1995, 2000, 2001; *et seq.*). Furthermore, both misrepresentation models and cue-based retrieval models require added mechanics to differentiate cues for number and gender, meaning that gender should, if isolated properly, behave similarly to number in comprehension. Furthermore, if attraction is a product of the *process* of resolving agreement dependencies, then we do not expect to find attraction profiles in grammatical sentences, in line with the claims of cue-based models. If attraction is due to *fallibility* in representation of gender, on the other hand, we expect to find no differential attraction effect owing to grammaticality.

277 2.1. Participants

Participants were 104 native speakers of Arabic from the United Arab Emirates University (UAEU) student body with no history of language disorders and self-assessed proficiency with MSA (104 females; mean age 20.4 years).² All participants provided informed consent and were compensated monetarily for their time. This and all other studies reported here were approved by the NYU Abu Dhabi Institutional Review Board and the UAEU Ethics Committee.

2.2. Materials & Predictions

284

285

288

289

290

291

293

In order to assess the possibility of gender attraction in MSA, a set of 48 sentences containing a subject modified by a subject relative clause were constructed based upon the stimuli created for the experiment reported in Tucker et al. (2015). All the sentences were of the structure NPI — Complementizer — [Verb — NP2 — Adverb] — Target Verb — Continuation, where NP1 is the grammatically accessible subject and NP2 the attractor NP for agreement realized on the target verb. In MSA, however, there are comparatively few adverbs, and so in some cases adverbial PPs which comprised a single orthographic word were used instead. A complete list of stimuli for this experiment appear in Appendix A.

²The gender composition of our participant sample is due to the nature of instruction at the UAEU — there are separate campuses for male and female students, and all participant testing was conducted on the female campus.

MSA has a grammatical case marking system which interacts with orthography in nontrivial ways (Ryding, 2005, 165-204). Given that morphological case marking is known to influence attraction effects for both gender (Badecker & Kuminiak, 2007; Lorimor et al., 2008) and number (Häussler, 2009), NP1 and NP2 were both selected to be morphologically definite (marked with the definite proclitic |al-|) so that they belonged to declension classes that did not involve case marking morphology with orthographi-cally long vowels (Ryding, 2005, 182–204). This is an important desideratum because indefinite nouns in Arabic mark morphological accusative case distinct from nomina-30: tive, allowing participants to potentially disambiguate subjects and objects using this case-marking.

The result of these two constraints is that morphological case on NP1 and NP2 is expressed by short vowel diacritics which are not typically written in Arabic. Our stimuli therefore abstract away from the effects of morphologically overt case marking by not writing these short vowels, a convention which matches everyday written text in the Arab world. In fact, we matched this convention across the entire study: short vowels and other diacritics were only written in our stimuli when they would be written in everyday Arab print media. This is usually because a single vowel would disambiguate two orthographically ambiguous words or be more common by convention (i.e., the tanwiin on adverbs/adverbial PPs). For example, the MSA word بمابقاً being read incorrectly as saabaqaa ("they.DUAL raced"). An example sentence from these 48 stimuli is shown in (4):

Pal-mutarzim-u Pallaðii saasad-a **Pal-mudiir-a**the-translator-NOM COMP.MASC.SG helped-3.MASG.SG **the-manager-**ACC
Pahjaanan **ja-takallamu** xamsata luyaat-in bi-fassaahatin.
often 3.SG.MASC-**speaks** five languages-ACC with-fluency

"The translator who helped the manager often speaks five languages fluently."

In addition to the requirements discussed above, several other constraints were also placed on the creation of stimuli sentences: Firstly, the relative clause verbs were chosen such that they either took a bare NP complement or a PP complement headed by a

preposition which is orthographically encliticized to the relative clause direct object in order to ensure that all stimuli had the same number of words up to the main clause target verb. Secondly, Arabic has two distinct morphological tenses which are marked on verbs in part by distinct agreement affixes (Ryding, 2005, 439–44). In order to abstract away from the individual contributions of distinct tense/agreement affixes, the main clause target verbs were counterbalanced for the two tenses, perfect (e.g., "خار المعالمة المعال

With the NP subject and attractor, nouns were chosen which had a masculine stem which could be made feminine solely by addition of the feminine singular nominal suffix /-a/ (orthographic s̄-) — in MSA these are easiest to find in the domain of nouns which denote human occupations. While MSA does have nouns which are feminine without the presence of this suffix, restriction to these nouns was employed in order to abstract away from possible differences in the processing of nominal gender owing to whether or not the feminine gender was an inherent property of the stem versus the contribution of an overt suffix (Sicuro Corrêa et al., 2004). Moreover, the choice of an overtly suffixing feminine allows a straightforward comparison between the processing of gender in MSA and suffixal plural morphology in other languages. The result is also a set of stimuli where grammatical gender is morphologically expressed in ways similar to English pluralization with /-s/, for example.

For each experimental sentence, four variants were constructed by systematically varying the grammatical gender of the attractor (NP2) and the main clause verb (target verb). These manipulations are coded as MATCH (does the grammatical subject match the attractor in grammatical gender value?) and GRAMMATICALITY (does the grammatical subject match the verb in grammatical gender?). Note that in this design, NOMATCH conditions are conditions with feminine attractors, since all subjects are masculine. Both relevant NPs remained in the singular throughout the experiment to assess the effect of gender alone. This resulted in four experimental conditions per stimulus; a complete set of four such sentences appears in Table 2.2.

The 48 sets of four sentences were distributed across four lists in a Latin Square

Condition	NP1 R1	Comp R2	RCV R3	NP2 R4	Adv R5	V R6	Continuation R7–R _n
MATCH/GRAM	المترجم	الذي	ساعد	المدير	أحياناً	يتكلم	خمس لغات بفصاحة.
	The translator (MASC)	who	helped	the manager (MASC)	often	speaks (MASC)	five languages fluently.
MATCH/UNGRAM	المترجم	الذي	ساعد	المدير	أحياناً	تتكلم	خمس لغات بفصاحة.
	The translator (MASC)	who	helped	the manager (MASC)	often	speaks (FEM)	five languages fluently.
NoMatch/Gram	المترجم	الذي	ساعد	المديرة	أحياناً	يتكلم	خمس لغات بفصاحة.
	The translator (MASC)	who	helped	the manager (FEM)	often	speaks (MASC)	five languages fluently.
NoMatch/Ungram	المترجم	الذي	ساعد	المديرة	أحياناً	تتكلم	خمس لغات بفصاحة.
	The translator (MASC)	who	helped	the manager (FEM)	often	speaks (FEM)	five languages fluently.

Table 1: A complete item set for one stimulus in Experiment 1.

design after being combined with 144 grammatical fillers of similar length for a 3:1 filler-to-item ratio. None of the fillers included the relative clause construction used in the experimental stimuli or any construction which drew attention to meaningful alternations in verbal agreement. In the final version of each list, only the experimental sentences contained ungrammaticalities, with 12.5% of the sentences in each list ungrammatical.

In this experiment, two predictions are of interest. First, ungrammatical verbal agreement is widely known to engender slower reading times, and we therefore expect a main effect of GRAMMATICALITY at the main clause/target verb region (and possibly 362 in subsequent spillover regions). Additionally, if attraction for grammatical gender in 363 MSA occurs at all, then one also expects to find an additional effect, but how attrac-364 tion should manifest is different for misrepresentation and cue-based retrieval theories. If cue-based retrieval theories are correct in asserting that attraction is not equivalent for grammatical and ungrammatical sentences, then one expects an interaction effect 367 of GRAMMATICALITY and MATCH at the target verb (or in spillover) owing to a slow-368 down of smaller magnitude in the NOMATCH/UNGRAM condition as compared to the 369 MATCH/UNGRAM condition. Alternatively, one could view this expectation as an er-370 roneous facilitation relative to the ungrammatical baseline in the MATCH/UNGRAM 371 condition. On the other hand, if misrepresentation of gender were the operative theo-372 retical mechanism, then one would expect to find only a main effect of MATCH and no 373 interaction. 374

375 *2.3. Procedure*

376

377

378

380

381

382

Subjects were seated comfortably up to eight at a time at a table in a quiet room in front of Apple iMac computers running Windows 7 natively via a Boot Camp partition on which the experimental software had been pre-loaded. Sentences were presented using the Linger software (Rhode, 2003) in a self-paced word-by-word moving window paradigm (Just et al., 1982). Each trial began with the display of a screen containing the sentence masked by dashes (including spaces and punctuation). Each time the participant pressed the space bar, a single word was revealed and the previous word re-masked with no look-back allowed. All items were presented in the Courier New Arabic font in

28pt bold type. A yes/no comprehension question followed each sentence, appearing on the screen all at once. Comprehension questions were designed in such a way that the answer could be provided independent of experimental manipulations — no questions asked about the attractor NP or the main clause verb. None of our comprehension 387 questions required lexical elaboration of the item or difficult semantic processing. A 388 majority of the comprehension questions asked about the relative clause verb or the 389 post-critical region continuation. As an example, the item The student who saw the professor(s) yesterday studied electrical engineering at the university was followed by the question Did the student study electrical engineering? Participants responded via a 392 dual Arabic/English keyboard where the 'f/ت' key was used for "yes (نعم)" and the 'j/ت' 393 key used for "no (Y)." Onscreen feedback was provided for both correct and incorrect 394 answers. Participants were instructed to read at a natural pace ensuring comprehension and were not alerted to the presence of grammatical errors in the stimuli, but they were warned that sentences read out of context might seem pragmatically odd. The 397 order of sentence presentation within each list was randomized for each participant. 398 Four practice items were presented before the start of the experiment, one of which was 399 ungrammatical and three of which were followed by a question. 400

401 2.4. Analysis

410

411

412

All data were analyzed in the R statistical software platform (R Core Team, 2015).

Error rates on comprehension questions were analyzed by fitting a logistic mixed effects model with the experimental manipulations MATCH and GRAMMATICALITY as
fixed effects and the experimental item and participant identity as random effects. We
adopted as a subject-exclusion criterion answering less than 60% of the comprehension questions correctly on the assumption that subjects below this threshold were not
properly attentive to the task.

For reaction time data, only data from sentences in which the comprehension question was answered correctly were included for analysis. Previous work attentive to the contribution of different portions of the reaction time distribution to agreement attraction configurations has shown that the canonical comprehension attraction effects are contained disproportionately in the right tail of reading times in regions where effects

exist (see Lago et al., 2015; Staub, 2009, 2010; and Tucker et al., 2015). Therefore, we deliberately chose a conservative method of by-region outlier treatment: Winsorization at 1% of the by-region mean (see Ratcliff, 1993 for discussion). No other exclusion criteria were used.

Reaction time data were analyzed by fitting a by-region linear mixed effects regression with the following predictors as fixed effects: (1) NP MATCH, (2) verb GRAM-MATICALITY, (3) the interaction of (1-2), (4) the order of presentation of the stimulus in the overall experimental order, (5) word length in characters,³ and (6) the reaction time to the immediately preceding region. We chose to model order of presentation, word length, and previous region RT because they are well-known predictors of reading time at a given region; where they are significant in the models, we do not comment on them in the discussion. Participants and items were modeled as random effects, as in the logistic models for error data. Where condition means are reported in the text, they are grand averages computed across subject averages for the conditions in question. A fixed effect is considered significant if the absolute value of its *t*-ratio was greater than two, a reasonable heuristic for its 95% confidence interval not including zero (Gelman & Hill, 2006). We report marginal effects (based upon the 90% confidence interval) when the absolute value of *t* is greater than 1.65, as well.

Finally, a word of note is in order about the fitting of mixed effects models and the interpretation of these models as there has been some debate about the amount of random effects structure to include in model building recently (see Barr et al., 2013 and Bates et al., 2015). In nearly all the models discussed in this paper, the maximal model which converged (in a reasonable amount of time) for all regions in a given experiment was one with random intercepts only for both subjects and items. In no cases did models

³Calculating word length is not a simple matter in Arabic, owing to decisions which must be made concerning the (non-)representation of short vowels. Given that this study used minimal short vowel representation and was concerned only with orthographic comprehension, length was computed over a Roman representation of the Arabic generated by applying a Buckwalter transliteration to the string appearing in our items. This yields a length count which includes only the short vowels we included and not those short vowels which are present in our strings but not represented orthographically. It is therefore probably overly conservative insofar as it represents an orthographic diacritic on par with a full-fledged character.

with more articulated random effects structures give qualitatively different results.

439 2.5. Results

440 2.5.1. Comprehension Question Accuracy

No participants met the criterion for exclusion due to low comprehension question 441 accuracy for this experiment. Overall comprehension question accuracy across all sub-442 jects was 88.5% for all items, with an accuracy of 90.2% for fillers and 83.4% for ex-443 perimental items. The accuracy for matching {subject, attractor} sentences was 86.8% 444 (95% CI = 84.8% - 88.6%) with grammatical verbs and 83.8% (95% CI = 81.5% - 85.7%)with ungrammatical verbs. Accuracy for non-matching {subject, attractor} sentences 446 was 86.8% (95% CI = 84.7%-88.6%) with grammatical verbs and 76.3% (95% CI = 447 73.8%-78.7%) with ungrammatical verbs. This difference emerged due to an effect of 448 both GRAMMATICALITY ($\hat{\beta} = -0.25; z = -2.18; p = 0.03$) and an interaction between GRAMMATICALITY and MATCH ($\hat{\beta} = -0.48; z = -3.04; p = 0.002$). Qualitatively. subjects were less likely to answer the comprehension question correctly when the verb 451 was ungrammatical, and even less likely when the verb was ungrammatical and the 452 attractor was feminine. Plausibly this is related to miscomprehension of the sentence 453 as a whole when the verb is ungrammatical. It seems unlikely that the interaction is 454 due to misidentification of the subject since the converse effect is not seen in Experiment 2 with feminine subjects (Bock & Miller, 1991), though the effect here could be 456 due to misrepresentation of the subject's features (Patson & Husband, 2015). At any 457 rate, neither the grammaticality effect nor the interaction term affected comprehension 458 in Experiment 2 or in the experiments on number (3-4), so we treat these effects as spurious here. In order to be conservative, we report the results with all incorrect com-460 prehension question items excluded from analysis, though the qualitative results of the 461 self-paced reading data do not change with the inclusion of these items, suggesting that 462 the higher error rates in the NOMATCH/UNGRAM condition did not unduly change the 463 pattern of results.

5 2.5.2. Self-Paced Reading

Only the sentences for which the comprehension question was answered correctly
were included for subsequent analysis of the self-paced reading data. This resulted
in the exclusion of 12.80% of the raw collected data (across all conditions, regions,
and participants). Mean reading times for each region and condition in Experiment 1
appear in Figure 1. Grand averages of the raw reading times for the critical verb and first
spillover region appear in Table 2 and means for all other regions appear in Appendix D.
Coefficient tables for the mixed effects models appear for the critical verb region and
subsequent spillover region in Table 3. Tables for all other regions appear in Appendix
D.

Condition	Mean	SD				
Critical Verb Region						
MATCH/GRAM	475	158				
MATCH/UNGRAM	502	195				
NoMatch/Gram	479	145				
NoMatch/Ungram	506	201				
First Spillover	Region					
MATCH/GRAM	418	107				
MATCH/UNGRAM	494	174				
NoMatch/Gram	422	115				
NoMatch/Ungram	473	166				

Table 2: Raw condition grand average reading times across participant means in milliseconds for regions of interest in Experiment 1.

At the main clause region there was a significant effect of Grammatical sentences ($\hat{\beta}$ = 30.17; s.e. = 11.02; t = 2.74). None of the other experimental manipulations were significant in this region (all |t| < 0.43).

Downstream from the target verb one region, however, a different pattern emerged. While the contribution of grammaticality persisted and increased in magnitude ($\hat{\beta}$ = 79.03; s.e. = 7.86; t = 9.96), there was also now a marginal interaction of MATCH and Grammaticality such that there was a difference between values of MATCH in ungrammatical sentences but not in grammatical sentences ($\hat{\beta}$ = -22.50; s.e. = 11.36; t =

Coefficient	β	SE	t			
Critical Verb Region						
Intercept	542.84	33.38	16.26			
Match:No	4.40	10.40	0.42			
Gram:Ungram	30.17	11.02	2.74			
Item Order	-1.14	0.07	-16.28			
Length	0.94	6.92	0.14			
Previous Region RT	0.07	0.01	6.46			
Match:No × Gram:UNGRAM	-0.79	15.04	-0.05			
First Spillove	r Region					
Intercept	446.84	17.20	25.98			
Match:No	3.43	7.86	0.44			
Gram:Ungram	79.03	7.94	9.96			
Item Order	-0.98	0.05	-18.90			
Length	14.33	2.01	7.13			
Previous Region RT	0.01	0.00	1.33			
<i>Match:</i> No × <i>Gram:</i> UNGRAM	-22.50	11.36	-1.98			

Table 3: Mixed effects regression coefficients for regions of interest in Experiment 1. Significant coefficients (|t| > 2) are in bold and marginal coefficients (|t| > 1.65) are in italics.

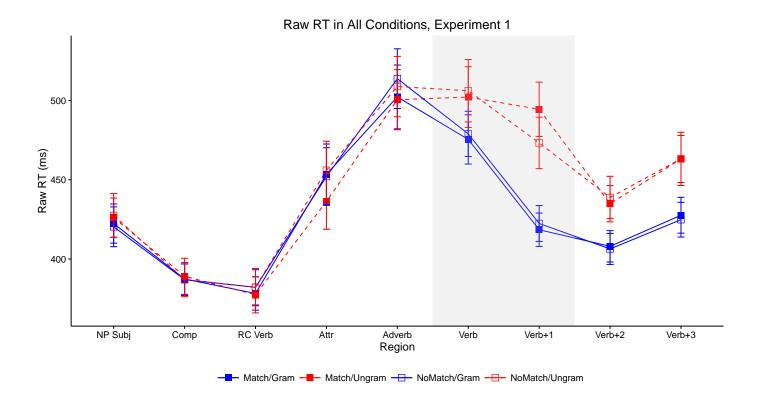


Figure 1: Mean raw reading times from Experiment 1 for all conditions and regions. Error bars represent the standard error of the condition mean across participant averages.

 484 -1.98). The computation of the p-value for this effect using the Welch-Satterthwaite approximation (Satterthwaite, 1946; Welch, 1947) for degrees of freedom was p = 0.05. Following this region, only the main effect of grammaticality persisted into the second spillover region ($\hat{\beta} = 26.51$; s.e. = 5.84; t = 4.54); the interaction effect did not (see Appendix D for complete modeling results).

489 2.6. Discussion

510

511

512

The results of Experiment 1 confirm the notion that gender, like number, can be 490 confusable in comprehension, assuming that the reaction time profiles observed here 491 correspond to illusory licensing. The longer reading times to regions including and 492 following the main clause verb suggest that readers notice verb ungrammaticalities on 493 the whole, spending longer time attempting to resolve the conflicting agreement information. However, relative to the baseline match condition, sentences in which an 495 erroneously feminine verb was preceded by a feminine relative clause object that mis-496 matched the true grammatical subject showed a marginally reduced reading time in-497 crease. Alternatively, one can view this as a relative *facilitation* of reading times in an 498 otherwise ungrammatical string. Either way, this interaction is the hallmark of agreement attraction effects in comprehension (see Dillon et al., 2013; Lago et al., 2015; 500 Pearlmutter et al., 1999; Wagers et al., 2009; Tucker et al., 2015 and references therein) 501 and plausibly interpretable as illusory licensing of ungrammatical verbs in some cases. 502 Moreover, these effects with gender are not seen in equal measure with grammatical 503 verbs. The lack of a reading time difference between masculine and feminine attractors in the grammatical conditions adds to the growing body of literature supporting the idea 505 that attraction effects in self-paced reading comprehension are limited to ungrammati-506 cal contexts (Lago et al., 2015; Tanner, 2011; Tanner et al., 2014; Wagers et al., 2009; 507 and Tucker et al., 2015). 508

It is important to emphasize that the finding of attraction for any agreement feature/cue in Arabic is striking given the relative inhospitality of Arabic to misrepresentations in agreement morphology. In our experimental stimuli, for instance, not only are attractor NPs overtly marked with a feminine suffix in the mismatch cases, both the relativizing complementizer $2alla\delta ii$ (like) and the embedded clause verb contained

overt morphology matching the correct subject. It seems untenable, therefore, to hold that comprehenders of MSA are more or less susceptible to attraction effects given the prevalence of agreement morphology in the language or a repeated reinforcement of the correct subject during the unfolding of a complex relative clause structure. All of this is true over and above any effect of relative clauses in general (see Bock & Miller, 1991 against Gillespie & Pearlmutter, 2013). This is an important cross-linguistic addition to the conclusions reached by Lago et al. (2015), for instance, that attraction effects are universal in character.

One caveat about the finding of agreement attraction for gender is that, unlike the reading time increases seen with ungrammatical verbs in general, the attraction effect is not present at the target verb in the main clause. Instead, the effect is delayed one region immediately downstream in the spillover. Since the spillover regions were not altered across conditions in a single item set, this difference must be a delayed effect of the gender mismatching agreement morphology encountered in the previous region. It is not uncommon for effects in self-paced reading to appear downstream from the point in the strings where the effect is first possible — in the seven experiments in Wagers et al. (2009), for instance, two of them show results where no effects appear at the critical verb itself (in a structure very similar to the one used here). While it is certainly interesting that this effect does not appear at the target region but nevertheless spills over (and that the opposite is true for the grammaticality effect), self-paced reading is not an appropriate methodology to address such a fine-grained timing difference owing to spillover effects.

In conclusion, it seems at least *prima facie* possible that verbal attraction for gender exists in MSA, insofar as sentences containing masculine subjects and feminine attractors show the reading time correlates of attraction. However, this is only one-half of the attraction effect profile seen for number in languages such as, *e.g.*, English. The other component to this effect is an asymmetry owing to *markedness* — attraction effects on reaction times or in productions are often found in languages when the erroneous verbal morphology is the marked version more than when it is in the unmarked version (Eberhard, 1997), but nothing in Experiment 1 has shown that this is true for MSA. As discussed in the Introduction, this is an important dimension of similarity upon which

to assess the similarity of gender and number attraction. Experiment 2, involving the manipulation of subject gender, was designed to address this question.

3. Experiment 2

Our goal in the second experiment was to assess whether the attraction effects for Arabic gender are replicable and whether they pattern along markedness lines the way other agreement features/cues have been observed to in other languages. At least three 550 papers (Badecker & Kuminiak, 2007; Badecker & Lewis, 2007; Malko & Slioussar, To 551 Appear) have all shown that gender attraction can in principle follow language-internal 552 markedness hierarchies with attraction effects sensitive to whether the verb appears in the marked or unmarked version. These findings are at odds, however, with findings 554 from Hebrew, where markedness effects do not appear to obtain in production (Dank & 555 Deutsch, 2010). Moreover, only one study (Malko & Slioussar, To Appear) has assessed 556 this phenomenon in comprehension, reporting one experiment on the three-way gender 557 system of modern Russian.

In MSA — a language with a two-valued system including masculine and femi-559 nine nouns — the marked grammatical gender is arguably feminine given that on many 560 nouns, feminine gender is overtly marked with a suffix. Furthermore, conjunctions 561 containing both masculine and feminine nouns invariably resolve to the masculine plural (Ryding, 2005). We therefore expect to find that gender attraction effect profiles 563 would appear more often in reading times when the true subject is masculine and the 564 attractor feminine, rather than the other way around, if markedness effects obtain as 565 in English number, where ungrammatical plural verbs are more acceptable with plu-566 ral attractors than ungrammatical singular verbs with singular attractors. However, it is equally possible that no markedness asymmetry obtains, as in Hebrew (Dank & Deutsch, 2010), and we would then expect no difference between masucline-feminine 569 conditions and feminine-masculine conditions. This latter result would be challeng-570 ing given the predictions of both misrepresentation and cue-based retrieval models, 571 assuming that grammatical gender features are subject to the same kinds of markedness distinctions applicable to grammatical number. 573

3.1. Participants

Participants were 128 native speakers of Arabic from the UAEU student commu-575 nity with no history of language disorder and self-assessed proficiency in MSA (128 576 females; mean age 20.4 years). All participants provided informed consent and were compensated for their participation in this study and an additional unrelated study. Par-578 ticipants were orally asked whether they had participated in Experiment 1 and excluded 579 from this experiment if they answered affirmatively. 580

3.2. Materials & Predictions 581

582

583

586

587

588

591

597

601

In order to directly assess the impact of markedness on gender attraction effects in MSA, the 48 item sets from Experiment 1 were altered to allow the main clause subject NP to also appear with the feminine suffix $-a/\bar{s}$. Where pragmatics required, the continuations were altered to allow for sensible interpretations across different genders of subject NPs. All other constraints on the creation of stimuli used in Experiment 1 were followed in this experiment as well, resulting in items which were identical to the items used in Experiment 1 save for these specific changes.

Using each of the 48 sentences as a standard, seven additional variants were con-589 structed by systematically varying the grammatical gender of both the main clause sub-590 ject and relative clause object NP as well as the main clause verb (the target verb). All feminine NPs were created by attaching the feminine suffix -a/5- to the NP used in the 592 equivalent masculine conditions. All NPs which were the target of experimental ma-593 nipulations were in the singular number and grammatically animate. We elected to use animate nouns despite the inclusion of notional gender of the referent in order to facilitate comparisons to Experiments 3 and 4 and the experiment from Tucker et al. (2015); 596 this additionally adds a new body of evidence to the production data from inanimates furnished by Dank & Deutsch (2010); and Deutsch & Dank (2009, 2011). The items 598 obtained by this choice also match English number marking in the nominal domain extremely closely: the marked alternative (here feminine, in English plural) is expressed with a single orthographic character suffix (-\vec{s} in Arabic and -s in English). The result is eight conditions per experimental sentence in a $2 \times 2 \times 2$ factorial design crossing 602 SUBJECT GENDER, GRAMMATICALITY, and MATCH.

It should also be noted that complementizers in MSA agree with the NP they modify in both grammatical number and grammatical gender (Ryding, 2005, 322), meaning that conditions with a feminine subject also contain a feminine singular definite complementizer (*Pallatii/والتي)*), in contrast to the masculine singular definite complementizer (*Pallaðii/والتي)*) found in masculine subject conditions. Additionally, whenever the subject NP was feminine, the relative clause verb also appeared in the feminine, so that the only possible agreement attraction effects occur on the main clause/target verb. A complete item set for one experimental sentence appears in Table 3.3.

The 48 sets of eight sentences were distributed across eight lists in a Latin Square design after being combined with 144 grammatical fillers of a similar length for a 3:1 filler-to-item ratio. None of the fillers used in Experiment 1 were used for this experiment, and none of the fillers included the relative clause construction used in the experimental stimuli or any construction which drew attention to meaningful alternations in verbal agreement. In the final version of each list, only the experimental sentences contained ungrammaticalities, with 12.5% of the sentences in each list ungrammatical.

In this experiment one expects a replication of the effects found in Experiment 1. One expects a main effect of GRAMMATICALITY at the target verb (and possibly into adjacent spillover) region. While the conclusions from Experiment 1 certainly lead one to expect an attraction effect in Experiment 2, what form that effect should take depends on the expectations one has about the role of markedness in gender attraction. If, following Badecker & Lewis (2007) and Badecker & Kuminiak (2007), markedness applies to gender in identical ways as it applies to number, then one expects to find a *three-way* interaction of SUBJECT GENDER, MATCH, and GRAMMATICALITY driven by the presence of an attraction effect only in MASC/NOMATCH/UNGRAM conditions (and not FEM/NOMATCH/UNGRAM conditions). On the other hand, if markedness affects different agreement cues differentially one expects to find a *two-way* interaction between MATCH and GRAMMATICALITY only, with no significant three-way effect.

3.3. Procedure

The procedure for Experiment 2 was identical to that employed for Experiment 1, save for the difference that participants were asked to participate in a second, unrelated

Condition	NP1 R1	Comp R2	RCV R3	NP2 R4	Adv R5	V R6	Continuation R7–R _n
MASC/MATCH/GRAM	المهندس	الذي	استقبل	العالم	بالصدفة	يعمل	على ابتكار جديد.
	The engineer (MASC)	who	met	the scientist (MASC)	by chance	is working (MASC)	on a new invention.
MASC/MATCH/UNGRAM	ألمهندس	الذي	استقبل	العالم	بالصدفة	تعمل	على ابتكار جديد.
	The engineer (MASC)	who	met	the scientist (MASC)	by chance	is working (FEM)	on a new invention.
MASC/NOMATCH/GRAM	المهندس	الذي	استقبل	العالمة	بالصدفة	يعمل	على ابتكار جديد.
	The engineer (MASC)	who	met	the scientist (FEM)	by chance	is working (MASC)	on a new invention.
MASC/NOMATCH/UNGRAM	المهندس	الذي	استقبل	العالمة	بالصدفة	تعمل	على ابتكار جديد.
	The engineer (MASC)	who	met	the scientist (FEM)	by chance	is working (FEM)	on a new invention.
FEM/NOMATCH/GRAM	المهندسة	التي	استقبلت	العالم	بالصدفة	تعمل	على ابتكار جديد.
	The engineer (FEM)	who	met	the scientist (MASC)	by chance	is working (FEM)	on a new invention.
FEM/NOMATCH/UNGRAM	المهندسة	التي	استقبلت	العالم	بالصدفة	يعمل	على ابتكار جديد.
	The engineer (FEM)	who	met	the scientist (MASC)	by chance	is working (MASC)	on a new invention.
FEM/MATCH/GRAM	المهندسة	التي	استقبلت	العالمة	بالصدفة	تعمل	على ابتكار جديد.
	The engineer (FEM)	who	met	the scientist (FEM)	by chance	is working (FEM)	on a new invention.
FEM/MATCH/UNGRAM	المهندسة	التي	استقبلت	العالم	بالصدفة	يعمل	على ابتكار جديد.
	The engineer (FEM)	who	met	the scientist (FEM)	by chance	is working (MASC)	on a new invention.

Table 4: A complete item set for one stimulus in Experiment 2.

experiment upon completion of the self-paced reading experiment reported here.

635 3.4. Analysis

Comprehension question accuracy data in Experiment 2 was analyzed identically to the analysis for Experiment 1. For the self-paced reading data, all of the analysis was the same as Experiment 1 save for the addition of the additional experimental manipulation of SUBJECT GENDER. Thus, the linear mixed effects model for this experiment included the fully-crossed fixed effects SUBJECT GENDER, MATCH, and GRAMMAT-

642 3.5. Results

3.5.1. Comprehension Question Accuracy

Three participants failed to meet the comprehension question accuracy criterion and were excluded from this and all further analysis. Overall comprehension question accuracy for this experiment was 86.7%, with an accuracy of 87.7% for fillers and 83.7% for experimental items. Accuracy rates for each condition is shown in Table 5. These differences, however, were not due to significant effects of the experimental manipulations (all |z|'s < 0.8).

	Subject Gender		
Condition	MASCULINE	FEMININE	
MATCH/GRAM NOMATCH/GRAM MATCH/UNGRAM NOMATCH/UNGRAM	83.0% (±3.0%) 83.1% (±3.0%) 83.5% (±2.9%) 83.3% (±3.0%)	83.4% (±3.9%) 84.9% (±2.9%) 83.4% (±2.9%) 83.8% (±2.9%)	

Table 5: Accuracy rates by condition to comprehension questions in Experiment 2. Ranges indicate 95% confidence intervals computed over condition means.

650 3.5.2. Self-Paced Reading

651

652

653

Only the sentences for which the comprehension question was answered correctly were included for subsequent analysis of the self-paced reading data. This resulted in the exclusion of 14.56% of the raw collected data (across all conditions, regions, and

participants). Mean reading times for each region and condition in Experiment 2 appear in Figure 2. Mean raw reading time grand averages appear for the critical verb and first spillover regions in Table 6. Mean RTs for all other regions appear in Appendix D. Coefficient tables for the mixed effects models appear for the critical verb region and subsequent spillover region in Table 7. Tables for all other regions appear in Appendix D.

Condition	Mean	SD		
Critical Verb Region				
MASC/MATCH/GRAM	434	170		
MASC/MATCH/UNGRAM	439	181		
MASC/NOMATCH/GRAM	437	161		
MASC/NOMATCH/UNGRAM	462	189		
FEM/NOMATCH/UNGRAM	431	165		
FEM/NOMATCH/GRAM	445	163		
FEM/MATCH/UNGRAM	437	173		
FEM/MATCH/GRAM	432	148		
First Spillover Reg	ion			
MASC/MATCH/GRAM	384	110		
MASC/MATCH/UNGRAM	462	238		
MASC/NOMATCH/GRAM	390	123		
MASC/NOMATCH/UNGRAM	433	163		
FEM/NOMATCH/UNGRAM	401	135		
FEM/NOMATCH/GRAM	403	133		
FEM/MATCH/UNGRAM	418	150		
FEM/MATCH/GRAM	393	120		

Table 6: Raw condition grand avergage reading times across participant means in milliseconds for regions of interest in Experiment 2.

At the critical verb region there were no effects of the experimental manipulations (all |t| < 1.44) save for a marginal three-way interaction of SUBJECT GENDER, MATCH, and GRAMMATICALITY ($\hat{\beta} = -41.90$; s.e. = 21.47; t = -1.95) driven by longer reading times to the MASC/NOMATCH/UNGRAM condition (see Figure 2).

One word immediately downstream from the critical verb, however, a different pattern emerged. In addition to a main effect of grammaticality ($\hat{\beta} = 77.66$; s.e. = 9.24; t = 8.40), GRAMMATICALITY also participated in two two-way interactions with each of the other experimental manipulations. Thus there was an interaction of SUB-

Coefficient	$\hat{oldsymbol{eta}}$	SE	t
Critical Verb Regio	on		
Intercept	479.21	28.16	17.02
Subj:FEM	-8.18	11.08	-0.74
Match:No	-1.48	10.78	-0.14
Grammaticality:UNGRAM	3.70	11.00	0.34
Item Order	-0.88	0.05	-17.45
Length	0.43	5.78	0.08
Previous Region RT	0.08	0.01	8.35
Subj:FEM × Match:No	18.86	15.16	1.24
Subj:FEM × Gram:UNGRAM	2.64	15.98	0.17
Match:No × Gram:UNGRAM	21.74	15.24	1.43
$Subj$:Fem \times $Match$:No \times $Gram$:Ungram	-41.90	21.47	-1.95
First Spillover Reg	ion		
Intercept	337.89	17.37	19.45
Subj:FEM	7.71	9.20	0.84
Match:No	8.02	9.25	0.87
Grammaticality:UNGRAM	77.66	9.24	8.40
Item Order	-0.70	0.04	-16.21
Length	15.16	2.38	6.36
Previous Region RT	0.10	0.01	10.30
Subj:FEM × Match:No	2.21	13.00	0.17
Subj:Fem × Gram:Ungram	-49.61	13.04	-3.80
Match:No × Gram:UNGRAM	-36.81	13.07	-2.82
Subj:FEM × Match:No × Gram:UNGRAM	12.96	18.41	0.70

Table 7: Mixed effects regression coefficients for regions of interest in Experiment 2. Significant coefficients (|t| > 2) are in bold and marginal coefficients (|t| > 1.65) are in italics.

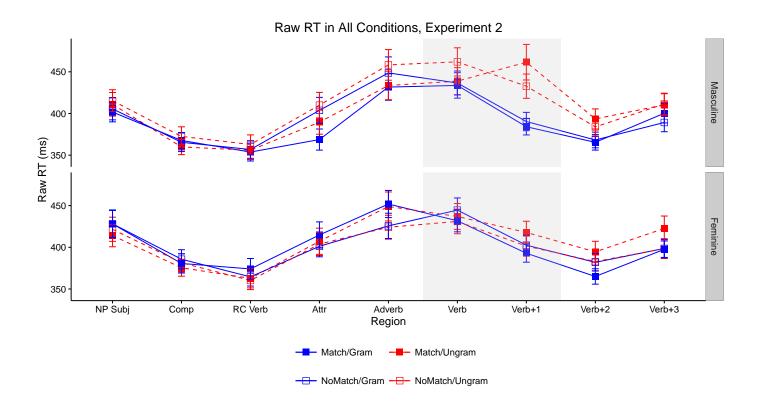


Figure 2: Mean raw reading times from Experiment 2 for all conditions and regions by SUBJECT GENDER. Error bars represent the standard error of the condition mean across participant averages.

JECT GENDER and GRAMMATICALITY ($\hat{\beta} = -49.61$; s.e. = 13.04; t = -3.80) driven by longer reading times to ungrammatical sentences with masculine subjects than to 669 sentences with feminine subjects. Additionally, there was an interaction of MATCH and Grammaticality ($\hat{\beta} = -36.81$; s.e. = 13.07; t = -2.82) driven by shorter read-67 ing times to NoMatch Ungrammatical sentences relative to Match Grammat-672 ICAL sentences. The three way interaction did not reach significance in this region 673 $(\hat{\beta} = 12.96; \text{s.e.} = 18.41; t = 0.70).$ After the first spillover region, only the main effect of grammaticality persisted 675 $(\hat{\beta} = 16.31; \text{s.e.} = 6.46; t = 2.53)$ such that grammatical sentences were read faster in 676 all conditions relative to ungrammatical sentences in the second spillover region. None 677 of the other experimental manipulations led to significant differences in reading time 678

680 3.6. Discussion

after the first spillover region (all |t| < 1.84).

The results of Experiment 2 replicates the finding that attraction for gender occurs 681 in MSA, but adds additional information about the processing of feminine morphology 682 and the differences between number and gender attraction. At the first post-verbal region in this experiment, the effect of MATCH bears all the hallmarks of an agreement at-684 traction effect but with a modulation attributable to the gender of the true subject. With 685 masculine subjects and feminine attractors (MASC/NOMATCH), reading times to un-686 grammatical verbs are not as long as they are to masculine attractors (MASC/MATCH). However, this effect is not significantly different with feminine subjects and masculine attractors — we find no evidence for a markedness effect in gender attraction in MSA 689 despite a very small numerical trend toward a smaller effect with feminine subjects. 690 Thus, our results are in line with the Hebrew production data from Dank & Deutsch 691 (2010) and distinct from the Slavic data in both production (Badecker & Kuminiak, 692 2007; Lorimor et al., 2008) and comprehension (Malko & Slioussar, To Appear). When assessing whether or not attraction is present in grammatical sentences in 694

When assessing whether or not attraction is present in grammatical sentences in
Experiment 2, some of the nuance inherent in the results comes into focus. While it is
not the case that sentences with masculine subjects show attraction in grammatical sentences, the situation is somewhat more muddled with feminine subjects. As examina-

tion of Figure 2 and the sizes of coefficients in the model show, there is hardly any sep-698 aration of the four feminine subject conditions, given that only the MATCH/UNGRAM 699 condition separates from the other three. Moreover, reading times to all ungrammatical verbs were somewhat reduced (~ 400 ms as opposed to nearly 450 ms for masculine 70 subjects) — a fact confirmed by the presence of a SUBJECT GENDER × GRAMMATI-702 CALITY interaction in our model. This suggests that participants are treating feminine 703 agreeing verbs differently than they treat masculine marked verbs. Nevertheless, this 704 depression of the grammaticality effect present in masculine verb sentences does not 705 affect the overall presence of an attraction effect insofar as there is still facilitation in 706 NOMATCH/UNGRAM relative to MATCH/UNGRAM conditions. 707

708

711

712

713

714

715

716

717

718

720

721

722

723

726

727

Further research is required to understand exactly what the nature of the grammaticality difference is between masculine and feminine subjects. One avenue worth exploring is that the third person feminine singular is homophonous with the second person singular in the imperfect tense (Ryding, 2005, 441). However, since our items were counterbalanced for tense, we have no way of assessing this hypothesis in the present data. Another alternative would take the decreased ungrammatical reading times to be a local coherency effect with the masculine attractor — an explanation which would provide some understanding of the slight increased reading times to MATCH/UNGRAM sentences in our data. Yet another possibility would be that the marked status of the head noun in feminine verb sentences changes the way attraction configurations are processed. Finally, it is also possible that the relatively lower frequency and marked nature of feminine verbal agreement leads subjects to accept masculine singular verbs as acceptable default/non-agreeing verbs — a concern which is especially intriguing given the diglossic nature of the Arab world and varying rates of proficiency with MSA from person to person. Our design in this study is not equipped to tease apart these differences.

Reaching across both these conclusions, however, is the spillover of the gender attraction effect. In Experiment 2 neither the grammaticality effect nor the attraction effect appeared until the region immediately following the critical verb. For the grammaticality effect, Experiment 2 therefore showed later onset of the longer reading times to erroneous verbs compared to Experiment 1 by one region. As noted in the discussion

for Experiment 1, not much can be made of these differences in self-paced reading, so
we simply observe the non-uniformity at present.

In an interim conclusion, gender attraction occurs in Standard Arabic and seems insensitive to markedness asymmetries found with number in English. But is this the same as number in Arabic? The only study which addresses this question is reported in Tucker et al. (2015). In that paper, the authors show that the attraction effect does occur for number in MSA, but had nothing to say about the featural asymmetry issue. Experiments 3 and 4 try to address these issues, and also clarify some unanticipated results Tucker and colleagues obtained vis-à-vis the number attraction effect in Arabic.

4. Experiment 3

738

753

754

757

In order to examine the similarities and differences between gender and number 739 attraction in MSA, one must examine whether the markedness asymmetry is present 740 in Arabic number attraction — an effect left untested in the comprehension study by 741 Tucker et al. (2015). However, testing number independent of gender in Arabic requires making a choice about which genders to include while independently manipu-743 lating number values. Since gender is orthogonal to number in MSA number agree-744 ment paradigms, the simplest option would be to simply counterbalance masculine and 745 feminine verbs across experimental items. However, the one existing study on MSA number attraction in comprehension, Tucker et al. (2015), presents findings concerning 747 the interplay of nominal gender and morphophonological effects on plural formation 748 which make this counterbalancing possibly undesirable. Since any experiment which 749 a priori restricted itself to one of two available genders in a language would need to be 750 justified, we first examine the findings from Tucker et al. (2015) in some detail with an 751 experiment designed to replicate and extend those findings. 752

We begin with a items subgroup issue in the study of Tucker et al. (2015). In that work, the authors leave unresolved a peculiar difference in agreement attraction effect sizes owing to the morphosyntactic nature of the NPs involved. Arabic allows for two different strategies of plural formation: SOUND/SUFFIXING plurals and BROKEN/ABLAUTING plurals. The former take their plurals with a regular, shape-invariant

suffix (in that study, $-aat/\Box$), whereas the latter mark plurality by a change in the the vowel and syllabic structure of the singular noun. In the traditional descriptive work on Arabic, this collection of vowels and prosodic structure is known variously as the CV-template, skeleton, or pattern. The vast majority of words in Arabic can be decomposed into a prosodic template and root consisting of 2-4 consonants, as (5) exemplifies for the root \sqrt{drs} :

```
(5) Words Containing \sqrt{drs}: (Wehr, 1976, 321)
```

```
a. darasa/در س "he studied/learned"
```

- b. darrasa/درّس "he taught/caused to learn"
- c. dars/درس "lesson/chapter"

- d. diraasa/در اسة "study/written"
- e. darraas/درّاس "student"
- f. madrasa/مدرسة "school"
- g. mudarris/مدرس "teacher/instructor"

Typologically, Arabic is unique in the high number of broken/ablauting plurals relative to other languages which utilize alteration of the CV-template — indeed, they are arguably more frequent than suffixing/sound plurals insofar as many of the high-frequency nouns in the language take broken plurals. Here, examining just English would lead to a different conclusion, such as that reached by Bock & Eberhard (1993), who demonstrate that attractors with irregular plurals in English do not condition different attraction rates in production than those with regular plurals.

As Ryding (2005) and Tucker et al. (2015) note, masculine animate nouns tend to take broken plurals and feminine animate nouns tend to take sound plurals. In Tucker et al. (2015), the authors demonstrated that the size of the number agreement attraction effect in MSA is modulated by whether the NPs in the pre-critical region are feminine and the attractor takes a plural with a regular suffix ("sound" plurals in the Arabic literature) or masculine and the attractor takes a plural by alteration of the CV-template/ablaut ("broken" plurals). Specifically, they showed that broken plural attractors cause smaller intrusion effect sizes at ungrammatical verbs than sound plural attractors do. This effect can be seen in the difference between the top and bottom panels of Figure 3 (Tucker et al., 2015, Fig.2). Whereas masculine/broken plural attractor sentences involve only

a modest attraction effect, feminine/sound plural attractor sentences involve a much larger attraction effect, with the attraction condition nearly identical to grammatical sentences. Given that all the subjects in this experiment were singular, Tucker and colleagues reason that this might be due to the salience of morphological plural marking on the attractor insofar as sound plurals contain a morphological or orthographic unit (the suffix) which is clearly associated with plurality, whereas comprehension of a broken plural *qua* plural requires decomposition of a word into its root and CV-template.

789

790

792

793

794

797

798

799

800

802

803

804

805

806

807

808

809

81

812

813

814

815

817

818

However, one issue that study does not address is whether there might be differentiations to be made inside the class of broken plurals such that the distinction in attraction effect sizes is not due to broken plurals per se, but instead is due to more general factors known to influence processing. One such property is AMBIGUITY of the morphological marking. As hinted at above, one of the distinctions between sound and broken plurals is that sound plural suffixes unambiguously mark plural number, whereas template alterations are commonplace across Arabic and serve to mark many morphological distinctions. Whether morphophonological properties of the attractor plays a role in modulating attraction rates is currently an open question at present: Whereas Vigliocco et al. (1995) and Malko & Slioussar (To Appear) find that they do not, Badecker & Kuminiak (2007) and Dank & Deutsch (2010) find that they do, and the experiments in Hartsuiker et al. (2003) suggest that they do, but at smaller rates than those of the head noun and only for certain kinds of ambiguity (i.e., grammatical case). We thus wish to rule out whether ambiguity of morphophonology on the attraction might be confounding the decreased magnitude in the broken plural attraction effect reported in Tucker et al. (2015).

Furthermore, Tucker et al. (2015) leave open whether the difference between sound and broken plural attractors is a categorical or gradient one: both the idea that broken plurals do not engender any attraction as well as the idea that they engender considerably smaller rates of attraction are compatible with their results. Here, we design an experiment which aims to clear up both these outstanding issues from Tucker et al. (2015) while simultaneously re-examining the timing of number agreement attraction. In order to do this, we exploit a fortunate property of Arabic broken plurals wherein some CV-templates underwriting broken plurals are used exclusively to mark plural

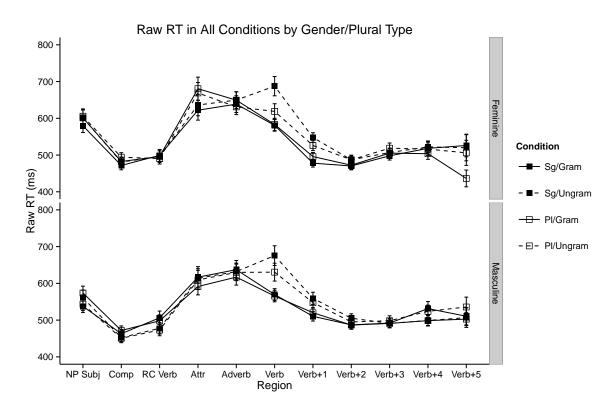


Figure 3: Mean raw reading times from Tucker et al. (2015) segregated by subject/attractor gender. All subjects are singular and Sg/Pl in the condition label refers to attractor number. Error bars represent standard error of the mean computed over subject averages.

number on nouns and some are not. For example, the CV-template associated with the 820 plural noun الصوص/lus $^{\varsigma}uus^{\varsigma}$, "thieves" — $C_1uC_2uuC_3$ — is also found in singular nouns, 82 such as the deverbal nominalization لدخول/duxuul, "entering (n.)" and is therefore morphologically ambiguous with respect to number marking. This can be contrasted with a 823 "scientists" (scientists — such as C₁uC₂aC₃aa? as in the noun علماء/Sulamaa? 824 - which is found only with plural nouns and can be considered morphologically un-825 ambiguous with respect to number. We therefore designed an experiment which tested only masculine attractors taking broken plurals and varied whether the template of those broken plurals is ambiguous or not. The result is a higher-powered replication of the 828 masculine half of the study in Tucker et al. (2015) (with twice as many items) and a 829 further investigation of the role of ambiguity in MSA number agreement attraction. 830

831 4.1. Participants

Participants were 110 native speakers of Arabic from the UAEU community with no history of language or other cognitive disorders and self-assessed proficiency in MSA (110 females; mean age 21.1 years). All participants provided informed consent and were compensated for approximately 45 minutes of time.

836 4.2. Materials & Predictions

In order to jointly assess the reliability of a lack of an attraction effect for number 837 in masculine broken plurals from Tucker et al. (2015) and the contribution (if any) of template ambiguity, 48 sentences were constructed of the form NPI — Complementizer 839 — [Verb — NP2 — Adverb] — Target Verb — Continuation, exactly as in the previous 840 two experiments and in Tucker et al. (2015). This is twice the number of items with 841 masculine pre-critical NPs compared to the subgroup in Tucker et al. (2015), where only 842 24 such items appeared. In this experiment, however, both NP1 and NP2 were specified as masculine grammatically and took their plural form in a broken pattern and not with a 844 suffix. Additionally, broken plurals were classified into two categories — AMBIGUOUS 845 and UNAMBIGUOUS plurals. Plural ambiguity was assigned based on the prosodic/CV-846 template pattern that the plural contained. Templates were considered ambiguous if the second author and a collection of other native speaker consultants could easily think of singular nouns which appeared in that same CV-pattern and unambiguous otherwise. A
complete list of the templates and classifications used in the construction of the stimuli
for this experiment appear in Table 8. In order to keep the duration of the experiment
manageable, the ambiguity of NP2 was manipulated across the 48 sentences. The result
was 24 items with NP2s that took ambiguous plurals and 24 items with NP2s that took
unambiguous plurals. All other constraints on the creation of stimuli in Experiments
1 and 2 were followed, where applicable to number instead of grammatical gender. A
complete list of sentences for this experiment appears in Appendix B.

Ambiguous	Unambiguous
$C_1aC_2aC_3a$ $C_1aC_2iiC_3$ $C_1iC_2aaC_3$ $C_1iC_2C_3aan$ $C_1uC_2C_3aan$ $C_1aC_2C_3aa$ $C_1uC_2C_2aaC_3$ $C_1uC_2uuC_3$	C ₁ uC ₂ aC ₃ aa? ?aC ₁ aaC ₂ iC ₃ a C ₁ awaaC ₂ iC ₃ ?aC ₁ C ₂ aaC ₃ ?aC ₁ C ₂ iC ₃ aa?

Table 8: Templates and ambiguity assignments for broken plural templates in Experiment 3.

The 48 sentences were then individually converted into four conditions by systematically varying the grammatical number (singular, plural) of both NP2 and the target verb. The resulting collection of four conditions for each of the 48 sentences comprised a 2 × 2 × 2 factorial design crossing MATCH (yes, no) and GRAMMATICALITY (grammatical, ungrammatical) and a between-items manipulation of AMBIGUITY. The 2 × 2 subset collapsing over ambiguity is therefore an indentical design to Experiment 1. However, in this study, all the NOMATCH conditions contained a singular NP1 and a plural NP2, and ungrammatical verbs were always plural. A complete item set for one of the experimental sentences appears in Table 4.2.

These 48 sets of four sentences were distributed across four lists in a Latin Square design and combined with 144 grammatical fillers for a 3:1 filler:item ratio where 12.5% of the items were ungrammatical. None of the fillers used in Experiments 1 or 2 were used for this experiment, and fillers varied in length from four to fifteen

Condition	NP1 R1	Comp R2	RCV R3	NP2 R4	Adv R5	V R6	Continuation R7–R _n
MATCH/GRAM	الطفل	الذي	رأى	الساحر	بانبهار	صفق	بشدة خلال العرض.
	The child (sG)	who	watched	the magician (sG)	with amazement	applauded (SG)	hysterically during the show.
MATCH/UNGRAM	الطفل	الذي	رأ <i>ي</i>	الساحر	بانبهار	صفقوا	بشدة خلال العرض.
	The child (sG)	who	watched	the magician (sG)	with amazement	applauded (PL)	hysterically during the show.
NoMatch/Gram	الطفل	الذي	رأ <i>ي</i>	السحرة	بانبهار	صفق	بشدة خلال العرض.
	The child (sg)	who	watched	the magicians (PL)	with amazement	applauded (SG)	hysterically during the show.
NoMatch/Ungram	الطفل	الذي	رأ <i>ي</i>	السحرة	بانبهار	صفقوا	بشدة خلال العرض.
	The child (sg)	who	watched	the magicians (PL)	with amazement	applauded (PL)	hysterically during the show.

Table 9: A complete item set for one stimulus in Experiment 3.

words long. None of the fillers contained the relative clause construction at the core of the experimental sentences.

Given that Tucker et al. (2015) report a diminished attraction effect in masculine items with broken plurals, one would expect to find only main effects of GRAMMAT-873 ICALITY in this experiment but no interaction of MATCH and GRAMMATICALITY, 874 though one could feasibly expect to see a numerical trend toward attraction which is 875 not very large in magnitude. If ambiguity of number marking is relevant for the effect reported by Tucker and colleagues, then we additionally would expect an effect 877 of AMBIGUITY interacting with MATCH, meaning that attraction is modulated by the 878 level of AMBIGUITY. Were that to obtain, whether or not GRAMMATICALITY is also 879 part of the interaction would be dependent upon the choice among misrepresentation 880 and cue-based retrieval models, exactly as in Experiments 1 and 2.

882 4.3. Procedure

The procedure for Experiment 3 was exactly the same as the procedure for Experiment 2.

885 4.4. Analysis

Comprehension question accuracy data for Experiment 3 was analyzed identically to the comprehension question accuracy data from Experiments 1 and 2. For the self-887 paced reading data, models were fit identically to Experiment 2 with SUBJECT GENDER 888 replaced by the value of the counter-balanced Ambiguity of the distractor NP. Addi-889 tionally, because Ambiguity was potentially not an influencer of reading times given the previous results in the production literature and a between-items manipulation in 891 our design, we also fit a 2 × 2 model with only MATCH and GRAMMATICALITY, but 892 only report properties of this model where they are informative to the discussion, as in 893 most cases the qualitative patterns of these two models were identical. 894

895 4.5. Results

896 4.5.1. Comprehension Question Accuracy

None of the participants in this experiment met the criteria for exclusion based on global comprehension question accuracy, and so all were included in the subsequent

analyses. Overall comprehension question accuracy for this experiment was 88.8% with accuracy rates of 86.8% for fillers and 89.5% for experimental items. Accuracy rates to matching attractor sentences were 88.6% (CI = 86.8-90.3%) to grammatical sentences and 87.2% (CI = 85.2%-88.9%) to ungrammatical sentences. Accuracy for non-matching attractors was 87.2% (CI = 85.3-89.0%) to grammatical sentences and 84.1% (CI = 82.0-86.0%) to ungrammatical sentences. These differences were not due to significant effects of the experimental manipulations (all |z|'s < 1.3).

906 4.5.2. Self-Paced Reading

916

Only sentences for which the comprehension question was answered accurately 907 were included in the subsequent reading time analysis. This resulted in the exclusion 908 of approximately 13.01% of the raw collected data (across all conditions, participants, 909 and items). Mean reading times across participant averages for each region are shown 910 in Figure 4. Grand average raw reading times for the critical verb and first spillover region are shown in Table 10. Means for all other regions appear in Appendix D. Co-912 efficient tables for the mixed effects models appear for the critical verb region and first 913 spillover region in Table 11. Tables for all the other regions included in the statistical 914 analysis appear in Appendix D.

Condition	Mean	SD
Critical Verb R	Region	
MATCH/GRAM	500	152
MATCH/UNGRAM	565	229
NoMatch/Gram	496	154
NoMatch/Ungram	563	235
First Spillover	Region	
MATCH/GRAM	440	125
MATCH/UNGRAM	488	151
NoMatch/Gram	446	111
NoMatch/Ungram	482	137

Table 10: Raw condition grand avergage reading times across participant means in milliseconds for regions of interest in Experiment 3.

At the critical verb region, none of the experimental manipulations were significant predictors of reading times at the verb, in either the ambiguous or unambiguous

Coefficient	\hat{eta}	SE	t		
Critical Verb Region					
Intercept	574.21	39.94	14.38		
Ambiguity: AMBIG	-10.53	20.10	-0.52		
Match:No	-14.41	14.94	-0.96		
Grammaticality:UNGRAM	26.00	22.61	1.15		
Item Order	-1.62	0.07	-22.19		
Length	17.14	8.68	1.98		
Previous Region RT	0.03	0.01	3.49		
Amb:UNAM × Match:No	15.16	21.61	0.70		
Amb:Unam × Gram:Ungram	14.99	21.62	0.69		
Match:No × Gram:UNGRAM	-0.16	21.26	-0.01		
Amb:UNAM × Match:No × Gram:UNGRAM	4.98	30.78	0.16		
First Spillover Region	on				
Intercept	467.12	17.52	26.65		
Ambiguity: AMBIG	2.83	11.79	0.24		
Match:No	3.99	9.76	0.41		
Grammaticality:UNGRAM	49.87	9.74	5.12		
Item Order	-1.14	0.05	-24.00		
Length	13.83	2.38	5.81		
Previous Region RT	0.04	0.01	5.51		
Amb:UNAM × Match:No	-1.16	14.12	-0.08		
Amb:UNAM × Gram:UNGRAM	-11.77	14.13	-0.83		
Match:No × Gram:UNGRAM	-15.85	13.89	-1.14		
$Amb: Unam \times Match: No \times Gram: Ungram$	14.95	20.11	0.74		

Table 11: Mixed effects regression coefficients for regions of interest in Experiment 3. Significant coefficients (|t| > 2) are in bold and marginal coefficients (|t| > 1.65) are in italics.

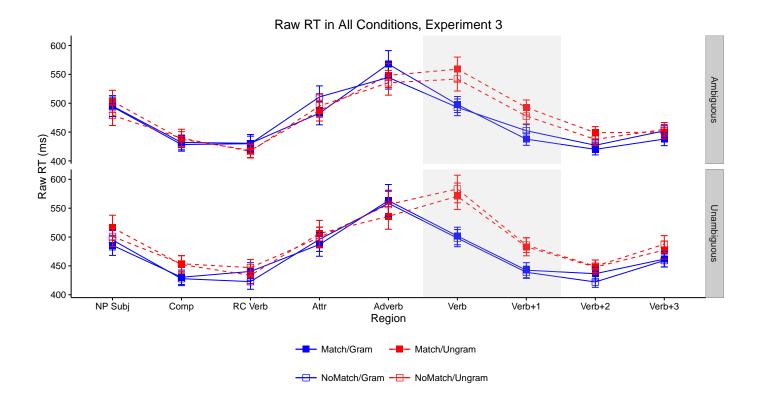


Figure 4: Mean raw reading times from Experiment 3 for all conditions and regions by attractor Ambiguity. Error bars represent the standard error of the condition mean across participant averages.

attractor sentences (all |t| < 1.16). Despite a sizable difference visually and trend in mean reading times, grammaticality was not a significant predictor of reading times in this region in either the full ($\hat{\beta} = 26.00$; s.e. = 22.61; t = 1.15) or the reduced models ($\hat{\beta} = 32.22$; s.e. = 19.85; t = 0.11), likely due to large standard errors.

After the critical verb one region, an effect of grammaticality emerged such that ungrammatical sentences were read considerably longer than grammatical sentences ($\hat{\beta}$ = 49.87; s.e. = 9.74; t = 5.12;). However, none of the other experimental manipulations were significant predictors of reading times in this region, either (all |t| < 1.15). The grammaticality effect continued in the second spillover region but diminished slightly in size ($\hat{\beta}$ = 20.02; s.e. = 7.43; t = 2.69) and none of the other manipulations had an effect in that region, either (all |t| < 1.17). Therefore, the only consistent effect throughout this experiment was the effect of grammaticality.

930 4.6. Discussion

The results of Experiment 3 serve as a replication of one-half of the experiment 931 reported in Tucker et al. (2015), insofar as it contained items with masculine NPs and 932 attractors that take broken plurals. In this experiment, we also fail to find any significant evidence of attraction effects in reading times. If one collapses over ambiguity in 934 Experiment 3 (see Figure 5), then this serves as a higher-powered replication of the mas-935 culine portion of Tucker et al. (2015) (with double the number of items; 24 in Tucker 936 et al., 2015 versus 48 here) and can lend support to the interpretation of the findings in 937 that paper wherein the authors suggest that agreement attraction is actually not present with masculine broken plural attractors, instead of simply diminished in intrusion effect 939 size. This can be seen in Experiment 3 when one considers the complete absence of an 940 attraction interaction in any of the target or spillover reasons, despite the presence of a 941 grammaticality main effect. Because grammaticality differences are being noticed by participants (leading to slowdowns in reading time), it is clear that participants are attending to the agreement morphology. However, the absence of intrusive facilitation in 944 the NOMATCH conditions means that agreement attraction is not occurring. More con-945 servatively, one might simply maintain that there is an important distinction to be made between the feminine sound plurals examined in Tucker et al. (2015) and the masculine

broken plurals re-examined here, one which must be taken into account when considering appropriate items for comparing attraction across gender and number features.

However, Experiment 3 can also provide an additional piece of information concerning what this difference might not be attributed to. In this experiment we find no evidence that the morphological ambiguity of the CV-template of the attractor drives this difference between masculines and feminines in Tucker et al. (2015). As outlined above, it was reasonable to wonder whether this could be the case, given the considerations that the ease with which participants recover morphological number information could underly broken versus sound plural differences. However, we find no evidence that this is actually occurring and therefore find converging evidence with that reported by, for instance, Vigliocco et al. (1995) and Malko & Slioussar (To Appear), that morphological ambiguity of the attractor not relating to case morphology plays little or no role in modulating attraction rates.

In order to directly compare the attraction effects of gender and number features in MSA, it was necessary to ensure that the plural-type asymmetry from Tucker et al. (2015) was replicable. With this in mind, we now turn to a domain in which agreement attraction effects for number *are* expected in MSA: sentences with sound plural attractors (the other subgroup from Tucker et al., 2015) in order to directly compare the results of Experiments 1–2 with similar effects for number.

5. Experiment 4

While Experiment 3 seems to confirm the claim that MSA number agreement attraction is not present when the pre-critical region contains masculine NPs and/or broken plural attractors, there remain several open questions about the nature of number agreement attraction in MSA given the results from Tucker et al. (2015) and the first three experiments. First, while it has been claimed above that gender attraction effects mirror agreement attraction effects in directionality but not markedness, this latter property has not been evaluated for Arabic number agreement in any fashion. The predictions are clear: given that English number attraction only gives rise to attraction RT profiles when the unmarked singular (*i.e.*, *is*) is replaced by the marked plural (*i.e.*, *are*), one

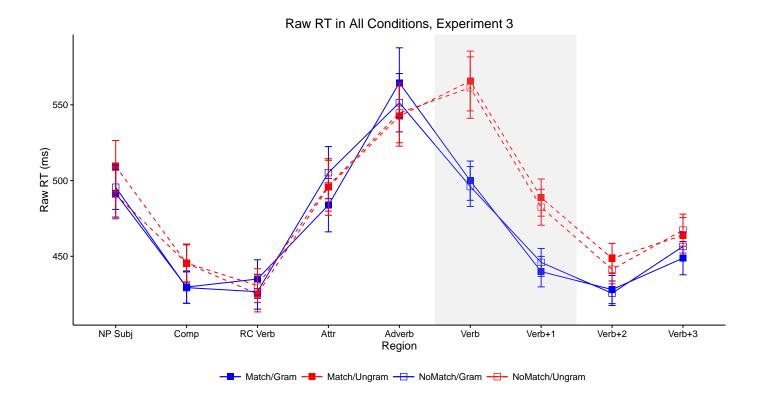


Figure 5: Mean raw reading times from Experiment 3 for all conditions and regions collapsed across attractor ambiguities. Error bars represent the standard error of the condition mean across participant averages.

could expect that attraction proceeds in the same way in MSA, different to what occurs
with MSA gender. Conversely, one could expect, in line with the predictions of both
representation and cue-based models and the results of Experiment 2, that number and
gender behave identically in not displaying markedness asymmetries in MSA. Furthermore, given that English and Arabic belong to distinct and somewhat disparate language
families where different notions of markedness are could be at play, it is important to examine whether plural-to-singular attractions give rise to attraction RT profiles in MSA,
as well.

Finally, the exact experimental design used by Tucker et al. (2015) was, as the au-985 thors themselves admit, not designed to observe the true strength of agreement attrac-986 tion effects after acknowledging a difference between masculine broken/ablauting and 987 feminine sound/suffixing animate plural attractors. This differential effect was an unexpected subgroup effect which should be examined more closely. In Experiment 3 we provided converging evidence that ablaut plurals in MSA do not show agreement 990 attraction, which means that attraction in that language should be solely a function of 991 suffixing plurals. To these ends, we designed an experiment exactly like Experiment 992 2, but which utilized only the feminine/sound plural attractor subgroup of items from 993 Tucker et al. (2015). The result is an experiment designed to replicate the presence of 994 attraction for number cues at the verb while simultaneously testing for the presence or 995 absence of a markedness asymmetry in MSA feminine number agreement attraction 996 effects. 997

998 5.1. Participants

1003

1004

1005

Participants were 112 native speakers of Arabic from the UAEU community with no history of language or other cognitive disorders as well as proficiency in MSA, both by self-report (112 females; mean age 20.6 years). All participants provided informed consent and were compensated for approximately one hour of time.

5.2. Materials & Predictions

In order to assess the effects of markedness in MSA number attraction as well as replicate the findings of Tucker et al. (2015) with respect to feminine attractors, 54

sentences were constructed of the form NP1 — Complementizer — [Verb - NP2] — $Adverb \]$ — $Target \ Verb$ — Continuation, exactly as in the previous three experiments. However, in this experiment both NP1 and NP2 were constrained to be grammatically feminine nouns bearing the feminine suffix $-a/\bar{s}$ —. Given that these nouns had singulars ending in $-a/\bar{s}$ —, their plurals were all suffixal, ending in $-aat/\bar{s}$ —. This choice was made for two reasons: (1) it allowed for higher-powered replication of the feminine subset of results from Tucker et al. (2015) (with 54 items compared to 24 in Tucker et al., 2015) and (2) if Tucker and colleagues' hypothesis that suffixing attractors provide for greater attraction rates, then these feminines provide the greatest opportunity to observe attraction with erroneous unmarked feminine singular verbs. All other constraints applied to items in Experiment 3 and in Tucker et al. (2015) were followed, where possible.

The 54 sentences were then individually converted into eight conditions by systematically varying the grammatical number (singular or plural) of the word in the *NP1*, *NP2*, and the *Verb*. The result was a collection of eight variants organized in a 2 × 2 × 2 factorial design crossing SUBJECT NUMBER (singular, plural), MATCH (yes, no), and GRAMMATICALITY (grammatical, ungrammatical). A complete item set for one of the experimental sentences appears in Table 5.2 and a complete list of experimental sentences appears in Appendix C.

These 54 sets of eight sentences were distributed across eight lists in a Latin Square design and combined with 144 fillers for a filler-to-item ratio of 2.67:1. The fillers were randomly selected from the collection of fillers used in Experiments 1–3, none of which contained the construction used in the experimental items (subject relative clauses attached to a subject) and varied in length from four to sixteen words long. All the fillers were grammatical with a total of 13.6% of the sentences ungrammatical in any given list.

If the results from the subset of feminine items in Tucker et al. (2015) replicate, then one expects to find a main effect of GRAMMATICALITY beginning at the main clause/target verb along with an interaction of MATCH and GRAMMATICALITY corresponding to the agreement attraction effect. These effects may spill over into the subsequent spillover regions but, given the effects in the previous study by Tucker and colleagues, one expects to find that the number attraction effect begins and is largest at

Condition	NP1 R1	Comp R2	RCV R3	NP2 R4	Adv R5	V R6	Continuation R7–R _n
SG/MATCH/GRAM	المدربة	التي	اهتمت	باللاعبة	جداً	اشتغلت	في الأكاديمية الوطنية للمبارزة.
	The coach (SG)	who	was interested	in.the player (SG)	very	worked (SG)	at the National Fencing Academy.
SG/MATCH/UNGRAM	المدربة The coach (SG)	التي who	اهتمت was interested	باللاعبة in.the player (SG)	بداً جداً verv	worked (SG) اشتغان worked (PL)	في الأكاديمية الوطنية للمبارزة. at the National Fencing Academy.
SG/NoMatch/Gram	المدربة	التي	اهتمت	باللاعبات	جداً	اشتغلت	في الأكاديمية الوطنية للمبارزة.
	The coach (SG)	who	was interested	in.the players (PL)	verv	worked (SG)	at the National Fencing Academy.
SG/NoMatch/Ungram	المدربة	التي	اهتمت	باللاعبات	جداً	اشتغلن	في الأكاديمية الوطنية للمبارزة.
	The coach (SG)	who	was interested	in.the players (PL)	very	worked (PL)	at the National Fencing Academy.
PL/NoMatch/Gram	المدربات	اللواتي	اهتمن	باللاعبة	جداً	اشتغلن	في الأكاديمية الوطنية للمبارزة.
	The coaches (PL)	who	were interested	in.the player (SG)	verv	worked (PL)	at the National Fencing Academy.
PL/NoMatch/Ungram	المدربات	اللواتي	اهتمن	باللاعبة	جداً	اشتغات	في الأكاديمية الوطنية للمبارزة.
	The coaches (PL)	who	were interested	in.the player (SG)	verv	worked (SG)	at the National Fencing Academy.
PL/MATCH/GRAM	المدربات The coaches (PL)	اللواتي who	اهتمن were interested	باللاعبات in.the players (PL)	بور) جداً verv	اشتغان worked (PL)	في الأكاديمية الوطنية للمبارزة. at the National Fencing Academy.
PL/MATCH/UNGRAM	المدربات	اللوات <i>ي</i>	اهتمن	باللاعبات	جداً	اشتغلت	في الأكاديمية الوطنية للمبارزة.
	The coaches (PL)	who	were interested	in.the players (PL)	very	worked (SG)	at the National Fencing Academy.

Table 12: A complete item set for one stimulus in Experiment 4. Note that NP1, NP2, RCV and V are all morphologically feminine.

the critical verb.

1038 5.3. Procedure

The procedure followed for Experiment 4 was exactly the same as the procedure for Experiment 3.

1041 5.4. Analysis

1042

1043

1044

1047

1048

1049

1050

1051

1052

1053

1054

Comprehension question accuracy for Experiment 4 was analyzed identically to the comprehension question accuracy analysis in Experiments 1–3. For the self-paced reading data, raw reading times were fit using a linear mixed-effects model exactly as in Experiment 2 save for the substitution of Subject Number for Subject Gender.

1046 5.5. Results

5.5.1. Comprehension Question Accuracy

One subject met the criteria for exclusion due to low accuracy based upon global comprehension question scores; she was therefore excluded from the subsequent analyses. Overall comprehension question accuracy for this experiment was 89.6% with accuracy rates of 89.4% for fillers and 89.7% for experimental items. Accuracy rates for all conditions appear in Table 13. These differences were not large enough to reach significance according to a logistic mixed-effects model fitted with the experimental manipulations as fixed effects (all |z|'s < 1.38).

	Subject Number			
Condition	SINGULAR	PLURAL		
MATCH/GRAM	89.3% (±2.5%)	89.5%(±2.5%)		
NoMatch/Gram	90.9% (±2.3%)	88.4%(±2.6%)		
MATCH/UNGRAM	90.0% (±2.4%)	88.5%(±2.6%)		
NoMatch/Ungram	89.6% (±2.5%)	89.2%(±2.5%)		

Table 13: Accuracy rates by condition to comprehension questions in Experiment 4. Ranges indicate 95% confidence intervals computed over condition means.

5.5.2. Self-Paced Reading

Only sentences for which the comprehension questions were answered correctly were included in the reading time analysis. This resulted in the exclusion of approximately 10.69% of the raw data acquired from the experimental sentences (across all conditions, participants, and items). Mean reading times across participant averages for all conditions by subject number appear in Figure 6. Grand average raw reading times for the critical verb and first spillover region appear in Table 14. Means for all other regions appear in Appendix D. Coefficient tables for the linear mixed effects models appear for the critical verb region and first spillover region in Table 15. Tables for all the other regions included in the statistical analysis are included in Appendix D.

Condition	Mean	SD			
Critical Verb Region					
SG/MATCH/GRAM	417	132			
SG/MATCH/UNGRAM	467	196			
SG/NoMatch/Gram	429	148			
SG/NoMatch/Ungram	434	164			
PL/NoMatch/Ungram	439	159			
PL/NoMatch/Gram	429	149			
PL/MATCH/UNGRAM	421	133			
PL/MATCH/GRAM	452	181			
First Spillover Region					
SG/MATCH/GRAM	381	96			
SG/MATCH/UNGRAM	428	137			
SG/NoMatch/Gram	383	100			
SG/NoMatch/Ungram	408	120			
PL/NoMatch/Ungram	394	119			
PL/NoMatch/Gram	394	117			
PL/MATCH/UNGRAM	394	107			
PL/MATCH/GRAM	394	109			

Table 14: Raw condition grand avergage reading times across participant means in milliseconds for regions of interest in Experiment 4.

In the critical verb region, the three-way interaction of the experimental manipulations was significant ($\hat{\beta} = 75.55$; s.e. = 20.16; t = 3.75). In the singular subject sentences there were slower reading times to MATCH/UNGRAM sentences than to NO-MATCH/UNGRAM, a difference which did not appear with plural subjects (see Fig-

Coefficient	\hat{eta}	SE	t
Critical Verb Region			
Intercept	424.10	25.10	16.90
SubjNum:PL	26.23	10.31	2.54
Match:No	10.37	10.04	1.03
Grammaticality:UNGRAM	43.26	10.30	4.20
Item Order	-0.88	0.05	-19.42
Length	7.58	4.53	1.67
Previous Region RT	0.10	0.01	11.93
SubjNum:PL × Match:No	-29.00	14.25	-2.04
SubjNum:PL × Grammaticality:UNGRAM	-63.50	14.87	-4.27
Match:No × Grammaticality:UNGRAM	-40.86	14.21	-2.88
SubjNum:PL × Match:No × Grammaticality:UNGRAM	75.55	20.16	3.75
First Spillover Region			
Intercept	382.71	12.16	31.47
SubjNum:PL	9.45	7.00	1.35
Match:No	1.56	6.97	0.22
Grammaticality:UNGRAM	43.66	7.00	6.24
Item Order	-0.68	0.03	-21.36
Length	8.05	1.37	5.89
Previous Region RT	0.08	0.01	11.10
SubjNum:PL × Match:No	1.24	9.89	0.13
SubjNum:PL × Grammaticality:UNGRAM	-36.43	9.92	-3.67
Match:No × Grammaticality:UNGRAM	-18.89	9.87	-1.91
SubjNum:PL × Match:No × Grammaticality:UNGRAM	13.85	14.01	0.99

Table 15: Mixed effects regression coefficients for regions of interest in Experiment 4. Significant coefficients (|t| > 2) are in bold and marginal coefficients (|t| > 1.65) are in italics.

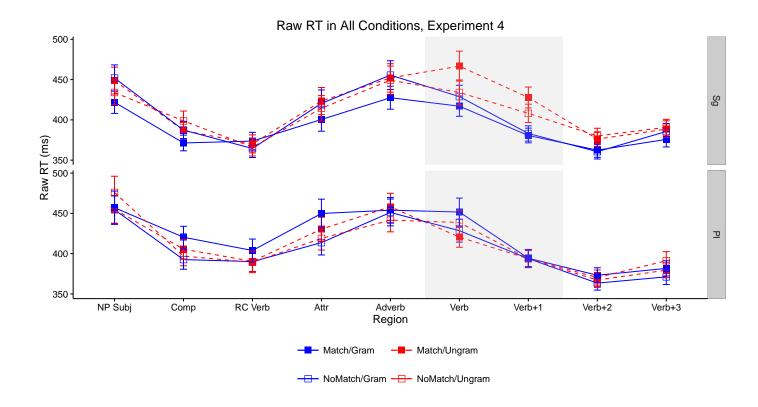


Figure 6: Mean raw reading times from Experiment 4 for all conditions and regions by subject number. Error bars represent the standard error of the condition mean across participant averages.

ure 6). There were also significant effects of all the two-way interactions: (1) Subject Number and Match ($\hat{\beta} = -29.00$; s.e. = 14.25; t = -2.04), (2) Subject Number and Grammaticality ($\hat{\beta} = -63.50$; s.e. = 14.87; t = -4.27), and (3) Match and Grammaticality ($\hat{\beta} = -40.86$; s.e. = 14.21; t = -2.88). There were also effects of Subject Number ($\hat{\beta} = 26.23$; s.e. = 10.31; t = 2.54) and Grammaticality ($\hat{\beta} = 43.26$; s.e. = 10.30; t = 4.20) alone.

Immediately after the critical verb region, the effect of GRAMMATICALITY re-1075 mained, albeit slightly diminished in size ($\hat{\beta} = 43.66$; s.e. = 7.00; t = 6.24). The 1076 interaction between SUBJECT NUMBER and GRAMMATICALITY also persisted ($\hat{\beta}$ = 1077 -36.43; s.e. = 9.92; t = -3.67), with plural subject sentences having faster reading 1078 times in ungrammatical sentences relative to sentences with singular subjects. Addi-1079 tionally, the interaction between MATCH and GRAMMATICALITY ($\hat{\beta} = -18.89$; s.e. = 1080 9.87; t = -1.91;) also persisted albeit marginally, though visual inspection of the reading times clearly reveal an attraction effect only in singular subject items. None of 1082 the other terms were significant in this region (all |t| < 1.36). None of these effects 1083 persisted any further downstream (all |z| < 1.18; see Appendix D). 1084

5.6. Discussion

1086

1087

1088

1090

1091

1092

1093

109

1096

The results of Experiment 4 largely replicate the results found by Tucker et al. (2015) for the feminine suffixing plural subgroup of items in their experiment. Specifically, we can observe here that participants are able to recognize grammaticality manipulations early — upon being presented with the ungrammatical verb. Also like in the first three experiments in this study, this effect continues for at least 500 ms — the amount of time it takes participants to complete reading the verb and move on to the next word.

As to the question of whether or not the subgroup effect owing to suffixing feminine plurals in Tucker et al. (2015) can be relied upon, the answer here seems to be an affirmative one. The interaction of MATCH and GRAMMATICALITY which appears at the verb for singular subject sentences is a direct analogue of the attraction effect in English and a replication of the previous results reported by Tucker and colleagues.

Moreover, this effect is largest at the verb region, though it continues into the immediately postverbal spillover region.

In addition to the existence and timing of the number attraction effect, we can also observe that, at a coarse level, attraction effects for number in MSA are subject to the same markedness effect seen in English (Eberhard, 1997). Inside the paradigm of feminine verbal agreement in MSA, plural is the clearly marked option. In addition to being morphologically expressed as the third person masculine prefix *ya-/\ceig-* plus a marked suffix (-na/\ceig-), the feminine plural is less frequent than the feminine singular. Finally, the default resolution of groups of plurals in MSA is to the masculine (*i.e.*, groups of mixed genders are referred to using the masculine plural). Given the absence of an attraction effect with feminine plural subjects and feminine singular attractors/verbs, one can say that MSA, like English, allows attraction to the marked form of a verb somewhat regularly for number, but allows attraction to the unmarked form of a verb either not at all or to a much smaller degree. This is a distinctly different result than that observed in Experiment 2 for gender and in languages such as English (Bock & Miller, 1991; Eberhard, 1997).

However, it should be noted that in this experiment, there is a trend toward attraction in NoMatch conditions with plural subjects; the absence of an interaction term in the statistical analysis is largely due to the complete *absence* of a grammaticality effect for plural subjects. However, even *modulo* this caveat, we believe it is possible to conclude that number attraction respects markedness for two reasons: (1) isolating the trend in facilitation between Plural NoMatch/Ungram and Match/Ungram conditions shows that the effect is much smaller, if real at all (18 ms of facilitation compared to 33 ms of facilitation for singular subjects) and (2) the numerical trend is confined to the critical verb region for plural subjects, whereas with singular subjects the effect spills over. These differences suggest that if an attraction effect is present in plural-to-singular verb attractions, it is considerably smaller than the reverse effect and/or has a different time course. Additionally, it is worth observing how our results

⁴Note that since our items were counterbalanced for tense, the prefix portion with *ya-/ç-* only appears in half of items; all the items contained the suffix portion, however.

add to a discussion present elsewhere in the attraction literature, namely whether attraction is present in grammatical sentences. As the introduction and Engelmann et al. (2015a,b) make clear, the literature is somewhat inconsistent with respect to whether attraction is present in grammatical sentences in the form of inhibition in NOMATCH conditions. Here, as in Experiment 2, we observe no such attraction. While there is a trend present in the plural sentences toward an attraction profile in the reaction times, there is nowhere near equivalent effects in both grammaticalities for both kinds of subjects.

1126

1127

1128

1129

1130

1131

1132

1133

1134

1135

1136

1137

1138

1139

1140

1141

1142

1143

1144

1145

1146

1148

1149

1150

1151

1152

1153

1154

1155

Finally, our results reveal a somewhat remarkable general absence of grammaticality effects with plural-marked verbs in MSA. As can be seen clearly in the models and Figure 6, ungrammatical sentences with plural subjects do not engender significantly slower reading times from grammatical sentences with plural subjects. If reaction-time slowdowns are equated with error recognition, then this result points to the conclusion that singular verbs in the context of plural subjects are not as easily recognized as ungrammatical by our participants. Since we did not expect this effect, our experiments were not designed to explain why this might be, but we can offer two suggestions here: First, MSA is an inherently diglossic "community" insofar as it is no speaker's native language — the language itself is quite distinct from the spoken varieties of Arabic common throughout the Arab world and all Arabs learn MSA as a second language (albeit one that can be used with a near-native degree of fluency with proper study; Ryding, 2005). When this is coupled with the observation that all the subjects in Experiment 4 were feminine, one could imagine that participants have atypical reactions to feminine plural verbs, marked by morphology which is relatively infrequently encountered given the conjunction and group reference resolution rules in the MSA (all of which resolve to masculine plural). Alternatively, it could be the case that participants are willing to accept singular verbs as a kind of default agreement — again not unreasonable given grammatical rules in the language which cause many plural inanimates to appear with singular feminine agreement morphology (Ryding, 2005). Further research is required to disentangle these possible explanations.

In summary, the results from Experiment 4 confirm the notion that number agreement attraction in MSA is present at erroneous verbs in the feminine morphological

paradigm, given the presence of suffixing distractors. Moreover, this effect is timed similarly to other number results insofar as it peaks at the critical verb and decays quickly thereafter. Moreover, attraction in grammatical sentences does not appear to occur in MSA.

1161 6. General Discussion

The results of the four experiments reported here confirm the notion that errors in agreement dependency comprehension are, at their core, universal in scope. Despite the universality of the errors, however, the studies reported here have uncovered some important differences between number and gender agreement in comprehension which have ramifications for theories of agreement attraction.

1167 6.1. Dimensions of Similarity

As the results of these four studies show, whether or not one concludes that gender and number attract identically depends on which dimension one assesses similarity
upon. Here we conclude that gender and number are largely qualitatively similar in their
attraction profiles, with the exception of markedness effects. Quantitatively, however,
it appears that these two features attract differently.

Existence. The first and perhaps most obvious way in which gender and number can 1173 be similarly involved in attraction is the basic fact that both these features give rise to 1174 attraction RT profiles in the comprehension of verbs. In Experiments 1 and 2, the RT 1175 profiles at and immediately following the critical verbs include a facilitation to No-MATCH/UNGRAM conditions relative to the large reading time spike seen in response 1177 to MATCH/UNGRAM conditions. This is the classic attraction profile in comprehension 1178 observed for number in Arabic in Experiment 4 and in many other languages (Dillon 1179 et al., 2013; Malko & Slioussar, To Appear; Pearlmutter, 2000; Pearlmutter et al., 1999; 1180 Tucker et al., 2015; Wagers et al., 2009; i.a.). This is an important conclusion despite 118 its obviousness given that no major theory of attraction effects could, in principle or 1182 without alteration, ensure that grammatical number is subject to attraction effects in 1183 verbal comprehension but grammatical gender does not. 1184

Grammatical Asymmetries. Another important dimension along which attraction for gender and number emerges as identical in our studies is the asymmetry of the attrac-tion effects with respect to the grammaticality status of the verb. In all four of the experiments reported here, attraction RT profiles, if they are present, are present only in ungrammatical sentences. Modulo Experiment 3, where no attraction is present at all, throughout all the experiments a difference in the MATCH versus NOMATCH con-ditions emerges only when the verb is grammatically unacceptable. While there is some contention about the generality of this finding (see Engelmann et al., 2015b and Franck et al., 2015), here we can add three more experiments to the list of those which do not observe attraction effects in grammatical sentences (e.g., Dillon et al., 2013; Tanner et al., 2014; Wagers et al., 2009). As noted in the introduction, one empirical point of distinction between competing theories of attraction effects has to do with the equiva-lency of attraction effects in both grammatical and ungrammatical sentences — process theories are arguably better-equipped to handle these asymmetries than representation theories, a point to which we return below.

Markedness Asymmetries. One way that gender and number are not identical, however, has to do with the presence of the asymmetry that we have been calling markedness-based. In MSA, plural number is marked (in the sense of Trubetskoy, 1939/1958) relative to singular and feminine gender is marked relative to masculine. If gender and number are equivalent along the markedness dimension and in line with the markedness results reported for English (Bock & Miller, 1991; Eberhard, 1997), one would expect that attraction RT profiles are present and/or strongest for singular subjects with plural attractors and masculine subjects with feminine attractors. In contrast, one would expect attraction RT profiles to be absent or greatly reduced for plural subjects with singular attractors and feminine subjects with masculine attractors. For number, this is what we observe, where Experiment 4 shows a greatly reduced attraction profile for plural subjects relative to the attraction profiles seen with singular subjects — so much of a difference that the statistical models report a reliable difference. On the other hand, Experiment 2 does not show the expected effect for gender — there feminine subjects show equal attraction effect sizes to masculine subjects. Qualitatively, then, one can

report that while markedness asymmetries in attraction do exist for number in MSA, they do not exist for gender.

RT Effect Size. Another important dimension along which to assess the similarity of attraction effects is the dimension of effect size. A priori, one could imagine two distinct quantities which define the quantity to be examined: the number of attraction incidents and the amount of reading time attraction change. Since this study involved only reading time, we have no direct way to assess the former, as individual trials do not provide such information given the latin square design (ensuring no subject saw all the relevant conditions). What we do have access to, however, is what Dillon et al. (2013) term the INTRUSION EFFECT SIZE — the amount of reading time change from fully ungrammatical sentences in non-attraction configurations to ungrammatical sentences in attraction configurations. In our experiments, this number corresponds to the difference between NOMATCH/UNGRAM and MATCH/UNGRAM conditions in the four experiments and corresponds to the amount of facilitation provided at the ungrammatical verb by the matching distractor NP.

In order to assess the various effect sizes across our experiments, we computed these intrusion effect differences across raw subject RT means as well as the corresponding Cohen's d (Cohen, 1988) for each of the four experiments reported here. We additionally computed the same values in the published data from Tucker et al. (2015). Table 16 gives these values, with shaded cells corresponding to significant main effects in the statistical models.

Comparing Experiments 1–2 to Experiments 3–4 and Tucker et al. (2015), one can see that there is plausible but weak evidence for a quantitative difference between gender and number attraction. Number attraction effects generate mildly larger effect sizes in both Experiment 4 and Tucker et al. (2015). Particularly striking is the difference of nearly 50 ms between the intrusion effect sizes in Experiment 2 (gender) and Tucker et al. (2015). As discussed below, this asymmetry is important to consider given the extent to which different theories of attraction predict identical effect sizes for features involved in agreement dependencies.

	Region			
Condition	Verb Region	Verb+1 Region		
	Experiment 1			
Masc-Fem	0.02 ~ 3 Experiment 2	0.12 ~ -21		
Masc-Fem	0.13 ~ 23	0.14 ~ -29		
Fem-Masc	$0.03 \sim -6$	$0.11 \sim -16$		
	Experiment 3			
Sg-Pl (Broken)	$0.02 \sim -2$	0.05 ~ -6		
	Experiment 4			
Sg-Pl (Sound)	0.18 ~ -33	0.12 ~ -20		
Pl-Sg (Sound)	$0.12 \sim 18$	< 0.001 ~ 0		
Tucker, et al., 2015				
Sg-Pl (Broken)	$0.15 \sim -40$	$0.07 \sim -11$		
Sg-Pl (Sound)	$0.30 \sim -73$	$0.18 \sim -25$		

Table 16: Absolute values of Cohen's d Intrusion Effect Sizes for the crticial verb and immediately postverbal spillover region for Experiments 1–4 and the experiment in Tucker, $et\ al.$ (2015). Numbers after the \sim are the means (in ms) of intrusion effect sizes computed across subject averages. Condition labels refer to the head noun/subject-attractor/verb and shaded cells correspond to significant effects in the statistical models reported in the text.

Effect Timing. Finally, it is worth considering whether the studies reported here provide any evidence for similarities or differences in timing in the appearance of agreement attraction effects. This is especially topical given the recent observations by Lago et al. (2015) that attraction effects can, in principle, appear after grammaticality effects in self-paced reading data. The question therefore arises as to whether gender and number show the appearance of attraction effects after grammaticality effects, and whether these profiles are the same or different.

1244

1246

1247

1248

1249

1250

1251

1252

1253

1255

1256

Here our results are somewhat inconclusive due to the self-paced reading methodology employed in this study. However, our results are at least suggestive (see also Table 16): Experiment 1 revealed the appearance of a grammaticality effect at the critical verb region, followed by attraction effect in the first spillover region. Experiment 2 reveals a mild grammaticality effect at the verb (in masculine subject sentences, where the grammaticality effect is present), and an attraction effect in the immediately postverbal spillover region. We therefore see weak evidence that gender attraction effects might be delayed by one region relative to grammaticality effects.

For number, Experiment 3 shows that grammaticality effects can, in principle, appear at the critical verb and continue into the following spillover region. However, since no attraction was observed in that experiment, it is impossible to assess attraction in that study. Experiment 4, however, is slightly more instructive: in that experiment, grammaticality effects *and* attraction effects both appear at the critical verb (in singular subject sentences, where attraction is present) and continue into the postverbal spillover region. This is, incidentally, the same qualititative pattern observed in Tucker et al. (2015).

It seems at least plausible, therefore, that differences in timing could exist between gender and number attraction effects which correspond to their intrusion effect sizes: smaller, later effects for gender and earlier, larger effects for number. However, confirmation of this result should be obtained via replication with methodology other than self-paced reading in which processing time is measured more continuously, such as eye-tracking or electro- or magnetoencephalography.

6.2. Implications for Representing Features and Cues

Given the importance of representational commitments to both major kinds of theories of agreement attraction, it is crucial to consider whether our results could be accounted for in ways neutral to processing theories by way of changes to the ways that linguistic features are used in processing or mapped onto cues for memory retrieval. Here we consider two approaches to featural representation: (1) an approach which localizes the difference in the valency of feature representation (*i.e.*, Fuchs et al., 2015) and (2) one which localizes the difference in the location of gender information in grammar and processing (*i.e.*, Deutsch & Dank, 2011).

One approach to asymmetries between gender and number would be to assert that these features are simply represented differently in grammar or processing. For instance, one could follow the approach of Fuchs et al. (2015) and assert that agreement features which show markedness asymmetries are PRIVATIVE — they are represented only in the marked value and not present otherwise. Features which do not

show markedness contrasts are instead EQUIPOLLENT — they are represented by the presence of features regardless of markedness. Fuchs et al. (2015), use this idea to represent the differential activity of gender and number in Spanish agreement attraction, and one could extend it to Arabic by positing that gender is bivalent ([± MASC]) whereas number is privative ([PL] or \emptyset). From this assumption one could tie either misrepresentation or cue-based retrieval models to this featural specification.

The problem with this approach is that it is not sufficiently supported by the distributional properties of the MSA grammar. For one, equipollent featural representations are typically used to encode three-way contrasts, which gender is not in Arabic — there is no neuter gender in MSA. While this is not an insurmountable representational issue, it does mean that the only evidence for equipollent gender in MSA would be the very markedness patterns that must be explained. A larger issue, however, has to do with number. Grammatical number in MSA is not a two-way system, but instead a three-way system, including a morphological DUAL which is used for sets of cardinality two (Ryding, 2005). Three-way distinctions are more difficult to encode in privative feature systems since privative representations are meant to encode two-way contrasts. What is needed to properly assess this question is a comparison of our results concerning singular and plural number with similar data concerning the dual in MSA.

A different approach to these issues would be to assert that gender and number are represented in different components of the processing system. For instance, Deutsch & Dank (2011) suggest that one could capture an identical pattern to our results but for Hebrew gender and number production data by assuming that gender is an inherent property of the lexical lemma and not part of the morpho-phonological properties of the word (see also Sicuro Corrêa et al., 2004). Grammatical number, on the other hand, is not an inherent property of the lemma, since any given lemma can be either singular or plural. Since the computation of number on nominals is part of the morpho-phonological process translating a lemma into a spoken word, it can be subject to principles of morpho-phonological markedness (see Deutsch & Dank, 2011 for details on a particular implementation of this idea in the Marking and Morphing model of Eberhard et al., 2005).

This approach certainly has some conceptual and empirical intuitiveness given that

grammatical gender is not typically meaningful in the same way as grammatical number and that the approach was designed to account for a similar set of facts in a closely related language — Modern Hebrew. However, while this approach is very well-suited to gating the presence or absence of attraction based upon markedness, it is incapable of attenuating or strengthening attraction effects in similar dimensions. Our results show that gender and number attraction effects are not simply different in quality, they are different in quantity, as well. In fact, one can step back and see that *any* attempt to explain our results based upon the representational structure or geometry of the features involved will be incapable of explaining the quantitative results we have observed in this study.

6.3. Implications for Theories of Attraction

Given that a simple representational change is not sufficient for explaining the differential effects that we observe for agreement attraction with gender and number, we now return to the two major classes of theories discussed in the introduction in light of these results. While both kinds of theories require nontrivial changes to their architectures to account for differences between gender and number, we ultimately suggest that cue-based retrieval theories require less drastic modifications (*i.e.*, such as those proposed in Engelmann et al., 2015b). What is over-arching to both discussions, however, is a need for a shift in the empirical domain of investigation for agreement attraction studies in particular and illusory dependency licensing studies in general: whereas current work has derived much of its insights from studies of the qualitative profiles of number in Indo-European languages, we believe that much insight can be gained by examining typologically diverse languages/features as well as the *quantitative* patterns of attraction in several comprehension methodologies.

6.3.1. (Mis)representation Theories

Our results present two major challenges for misrepresentation theories broadly speaking: (1) the differential quantitative strength of gender and number attraction and (2) the absence of agreement attraction RT profiles in grammatical sentences. Both of these challenges stem from a similar prediction common to representational theo-

ries: since theories that attribute attraction effects to failures of representation take the agreement process itself to be undisturbed when attraction occurs, they predict parity of attraction effects across identically represented subject NPs. What causes attraction in, *e.g.*, the theories of Eberhard et al. (2005); Franck et al. (2002); Nicol et al. (1997); Vigliocco & Nicol (1998) is a process by which structural representations of the subject are malleable enough to allow features of the attractor to be copied erroneously to the verb by the normal processes of subject-verb agreement. It is a corollary of this assumption that attraction should occur in equal measure in structurally identical subject NPs (Wagers et al., 2009).

But this is not what we observe for gender attraction. Our results suggest a smaller quantitative profile of attraction for gender in MSA than for number. Given that our experiments involved structurally identical subject and attractor NPs across all four experiments, these results cannot be explained by reference to different structural configurations leaking attractor features in different strengths. Gender attraction appears diminished in strength relative to number when compared directly in a subject relative clause configuration in both cases.

Here one could appeal to the quantitative Marking and Morphing Model of Eberhard et al. (2005) to attempt to derive these effects, but this attempt runs afoul of our use of animate human-denoting NPs in all four experiments. In the Marking and Morphing Model, one of the ways that structural representations are malleable is that top-level phrases are assigned featural strengths based upon a function of the features of their contained constituents in addition to their own feature values. An NP such as *The key to the cabinets* therefore contains some residual plurality given the plural denotation and marking of *cabinets*. However, with animate human-denoting NPs where grammatical gender is arguably semantically contentful, models such as the Marking and Morphing model should predict *stronger* attraction for gender given the clear morphological marking and semantic content of feminine gender in our MSA stimuli.

More broadly, however, both quantitative and qualitative misrepresentation models struggle with the lack of attraction consistently observed in our studies in grammatical sentences. As Wagers et al. (2009) have argued, these models cannot predict anything other than parity in the rates of attraction, since the malleable or leaky representation of

subjects occurs blind to what happens at the verb. Here it is not even enough to dispute the qualitative appearance of grammatical agreement attraction (*pace* Engelmann et al., 2015b and Franck et al., 2015), as we have shown a large quantitative difference which cannot be accounted for under misrepresentation approaches. The only misrepresentation approach which could account for these sorts of effects is the degraded memory representation model of Staub (2009, 2010), though this model too needs modifications to successfully predict differential quantitative strengths of attraction for number and gender.

6.3.2. Cue-based Retrieval Theories

Cue-based retrieval theories, on the other hand, deal much more successfully with the lack of attraction in grammatical sentences. In these models (such as those deriving from Lewis & Vasishth, 2005 and Badecker & Lewis, 2007), attraction occurs when cue-mismatches between subjects and attractors lead to the erroneous retrieval of the attractor during a working memory retrieval event triggered by the verb. There are two distinct ways to concretize this idea: either the retrieval event occurs in all instances or it only occurs upon the presentation of ungrammatical verbs. In either case, however, grammatical attraction is not predicted. In the first case, the complete cue match between subject and grammatical verb causes the probability of attractor retrieval to plummet relative to ungrammatical sentences. In the latter case, one simply constructs the model not to consider attraction in grammatical sentences by fiat.

What is less easily representable in these theories is the lack of quantitative symmetry between gender and number attraction in our results. Cue-based retrieval models are dependent upon the cue structure posited in the model. Setting the issues of markedness discussed in the previous section aside, it is difficult to see how cue structures can be posited that simultaneously cause attraction and also do so at different strengths. In the model of Lewis & Vasishth (2005), for example, the strength of a cue can only be modulated as a function of the number of other cues in memory, not the intrinsic properties of the cue itself. Thus, gender could be relatively weak relative to number, but only if gender uniformly occurred as a cue in contexts where more cues were available in the system in general. Obviously, this is not a feasible assumption for MSA, where gender

and number always co-occur (Ryding, 2005).

However, cue-based retrieval models are the subject of much active research and are constantly evolving. Engelmann et al. (2015b), for instance, have proposed two new mechanisms based upon a literature review in order to account for unrelated effects in the literature, DISTRACTOR PROMINENCE (a quantitative adjustment giving more activation to attractors as functions of their position and discourse prominence) and CUE CONFUSABILITY (the ability for cues to be related to features quantitatively instead of categorically). The latter of these ideas is an intriguing notion, though the specific formulation of the idea in Engelmann et al. (2015b) would not accommodate our results. Further work is needed to see if the cue-based retrieval models can be enriched in such ways to predict quantitatively different effects based on grammatically equipotent linguistic features but it does seem clear what this work would look like: an expanded theory of cue confusability which allows cues to be differentially weighted in isolation, as well as differentially confusible with one another.

Finally, it is worth noting that neither misrepresentation nor cue-based retrieval models could account for differences in timing of gender and number attraction effects. Attraction is a *verbal* process, meaning that the representations and processes responsible for these effects should be keyed at the verb, not later. As we observed, it is possible that our evidence weakly hints at the delayed appearance of attraction for gender relative to attraction for number. However, since self-paced reading methodologies commonly involve spillover effects with no clear theoretical explanation, we leave this idea and its implications for theories of attraction for future research.

6.4. Conclusions

We have demonstrated that verbal gender agreement attraction occurs in comprehension. Moreover, these results obtain in an inflectionally rich language in relative clause configurations where attraction should be smaller in effect, all else equal. We have also demonstrated that attraction for gender and number is not qualitatively or quantitatively identical in Arabic, with gender attraction failing to show markedness distinctions of the same kind showed by number in the same configurations. Quantitatively, we also demonstrated that agreement attraction for gender has a diminished

strength relative to number attraction and may occur later in time. We also added addi-1438 tional evidence to the body of work suggesting that comprehension attraction effects do 1439 not occur in grammatical sentences, for gender or number. These results were shown 1440 to be largely more compatible with cue-based retrieval models over misrepresentation 1441 models insofar as the former are capable of accounting for grammaticality asymmetries 1442 and require fewer alterations to account for quantitative differences among agreement 1443 features. Finally, we suggested that much progress can be made in theorizing about attraction by moving from qualitative work on grammatical number to quantitative work on other features and languages. 1446

1447 Acknowledgments

The authors would like to thank the members of the Language, Mind, and Brain 1448 Lab (LaMBLab) and Neuroscience of Language Lab (NeLLab) at NYU Abu Dhabi and 1449 audiences at CUNY 2015 (USC), AMLaP 2015 (L-Università ta' Malta), and the 2015 1450 Annual Symposium on Arabic Linguistics (University of Wisconsin — Milwaukee) 1451 for valuable feedback on this project. Thanks to Matt Wagers and Stephen Politzer-1452 Ahles for insightful discussion. Thanks to Souad Al Helou for assistance with stimuli 1453 construction and experimental procedure. Thanks also to Nizar Habash, Salam Khalifa, 1454 and Anas Shahrour for assistance with stimuli construction. Finally, thank you to three anonymous reviewers for very thoughtful comments on early versions of this paper. 1456

1457 Appendix A. Complete Materials – Experiments 1–2

- A.1. The translator who helped the manager occasionally speaks five languages fluently.
 - المترجم الذي ساعد المدير أحياناً يتكلم خمس لغات بفصاحة.
- A.2. The student who saw the professor yesterday studied electrical engineering at the university.
- الطالب الذي رأى الأستاذ بالأمس درس الهندسة الكهربائية في الجامعة.
- A.3. The engineer who met the scientist by chance is working on a new invention.

 المهندس الذي استقبل العالم بالصدفة يعمل على ابتكار جديد.
- A.4. The cook who scolded the waiter forcefully works in an expensive restaurant during the summer.
- الطباخ الذي وبّخ النادل بشدة يشتغل في مطعم غال خلال الصيف.
- A.5. The analyst who advised the minister intelligently discusses the Palestinian issue in depth.
- المحلل الذي نصح الوزير بذكاءِ يتناول القضية الفلسطينية بعمق.
- A.6. The child who saw the prince before visits the royal family each week.

 الطفل الذي رأى الأمير سلفاً يزور العائلة الملكية كل أسبوع.
- A.7. The teacher who taught the child dedicatedly attended the graduation party of the students.
- المعلم الذي علّم الطفل بتفانِ حضر حفل تخرج الطلاب.
- A.8. The consultant who warned the president yesterday found a solution for the financial problem.

 1478 لمستشار الذي حذّر الرئيس بالأمس وجد الحل للمشكلة المالية.
- A.9. The driver who accompanied the ambassador regularly works seven days a week. السائق الذي رافق السفير بانتظام يعمل سبعة أيام في الأسبوع
- A.10. The jailor who tortured the prisoner constantly cleans the cells nightly.

 1482 السجّان الذي عذّب السجين باستمرار ينظف الزنازين كل ليلة.
- A.11. The broadcaster who talked to the activist yesterday trained in a famous company. المذيع الذي كلم الناشط بالأمس تدرب في شركة مشهورة.
- A.12. The employee who helped the colleague humbly gained the confidence of colleagues at work.
- الموظف الذي ساعد الزميل بتواضع استطاع كسب ثقة باقى الزملاء في العمل.
- A.13. The singer who challenged the poet arrogantly has lost for not showing up on time.
- المنشد الذي تحدى الشاعر بغرور خسر لعدم حضوره في الوقت المحدد.

- A.14. The beginner who questioned the expert daily has acquired good experience. المبتدئ الذي سأل الخبير يومياً اكتسب خبرة جيدة. 1493
- A.15. The man who hosted the friend with pleasure slept in the basement of the house. 1494 الرجل الذي استضاف الصديق بسرور نام في الطابق السفلي من المنزل. 1499
- A.16. The patient who consulted the doctor yesterday returned home satisfied. 1496 المريض الذي استشار الطبيب بالأمس عاد الى المنزل مطمئنا.
- A.17. The teacher who met the writer happily likes reading about literature. 1498 المعلم الذي استقبل الأديب بسعادة يحب القراءة عن الأدب.
- A.18. The manager who phoned the partner in the morning plans to expand the branches 1500 of the company. 1501 المدير الذي هاتف الشريك بالصباح يخطط لتوسيع فروع الشركة. 1502
- A.19. The chef who invited the guest in the evening masters preparing various delicious 1503 dishes. 1504
- الطاهي الذي دعا النزيل بالمساء يتقن إعداد أصناف لذيذة من الطعام. 1505
- A.20. The child who watched the magician with amazement applauded hysterically dur-1506 ing the show. الطفل الذي رأى الساحر بانبهار صفق بشدة خلال العرض.

1508

- A.21. The young man who helped the wounded man in the morning spends every week 1509 reading. 1510
- الشاب الذي ساعد الجريح صباحاً يقضى كل الأسبوع في المطالعة. 1511
- A.22. The doctor who healed the captive quickly is joining the national guard (army). 1512 الطبيب الذي عالج الأسير بسرعة ينخرط في فرقة الدفاع المدني. 1513
- A.23. The ruler who imprisoned the criminal previously practiced justice and equality 1514 among the people. 1515
- الحاكم الذي سجن المجرم سابقاً مارس العدل والمساواة على الشعب. 1516
- A.24. The coach who trained the partner efficiently possesses more than 10-years of 1517 experience. 1518
- المدرب الذي درّب الرفيق بإتقان يمتلك خبرة تفوق العشر سنين. 1519
- A.25. The Imam who advised the neighbor wisely lives in the local neighborhood. الإمام الذي نصح الجار بحكمة يعيش في المنطقة المجاورة. 1520 1521
- A.26. The policeman who questioned the murderer seriously asked the pedestrians the 1522 reasons for the crime. 1523
- الشرطى الذي استجوب القاتل بجديّة سأل المشاة عن سبب الجريمة.
- A.27. The lawyer who called the inheritor in the morning discussed the issue of the 1525 inheritance distribution 1526
- المحامي الذي استدعى الوريث بالصباح ناقش موضوع تقسيم الإرث. 1527

- A.28. The teacher who taught the student well worked as a host in television. الأستاذ الذي علّم الطالب جيداً عمل كمذيع في التلفزيون. 1529
- A.29. The immigrant who spoke with the visitor for a long time feels nostalgic for the 1530 country always. المهاجر الذي حدَّث الزائر طويلاً بشعر بالجنين للوطن دائماً.

1532

1551

- A.30. The journalist who interviewed the leader persistently publishes the article in the 1533 national newspaper. 1534
- الصحافي الذي استجوب الحاكم بالحاح نشر المقال في الحريدة الوطنية. 1535
- A.31. The pilot who greeted the airline attendant warmly asked many questions during the journey. 1537
- الطيار الذي حيّا المضيف بحرارة سأل أسئلة كثيرة خلال الرحلة. 1538
- A.32. The man who talked to the monk intelligently works in the post office. 1539 الرجل الذي كلم الراهب بفطنة يشتغل في مكتب البريد. 1540
- A.33. The policeman who arrested the driver quickly helps the pedestrians in crossing 1541 the street.
- الشرطى الذي اعتقل السائق بسرعة يساعد المارة على عبور الشارع. 1543
- A.34. The lawyer who startled the witness cunningly stopped the accusation of his client 1544 in the court.
- المحامى الذي أربك الشاهد بدهاء أبعد التهم عن موكله في المحكمة. 1546
- A.35. The worker who helped the soldier yesterday drives a large truck for the company. العامل الذي ساعد الجندي بالأمس يسوق شاحنة كبيرة لفائدة الشركة. 1548
- A.36. The artist who served the king devotedly gave a portrait to the ambassador of the 1549 United States 1550 الفنان الذي خدم الملك بتفان تبرع بلوحة لسفير الولايات المتحدة.
- A.37. The journalist who hosted the star brilliantly raised very embarrassing questions. 1552 الإعلامي الذي استضاف النجم بتألق طرح أسئلة محرجة جداً.
- A.38. The nurse who is treating the patient carefully studies at the university hospital. الممرض الذي عالج المريض بعناية يدرس في مستشفى الجامعة. 1554 1555
- A.39. The seller who thanked the customer enthusiastically was happy with the large 1556 profit. البائع الذي شكر الزبون بحماس فرح بالفائدة الكبيرة. 1558
- A.40. The coach who was very interested in the player worked at the National Fencing 1559
- Academy. 1560 المدرب الذي اهتم باللاعب جداً اشتغل في الأكاديمية الوطنية للمبارزة. 1561

- A.41. The soldier who met the policewoman yesterday loved the atmosphere at the air base.
- الجندي الذي قابل الشرطي بالأمس أحبّ العمل في القاعدة الجوية.
- 1565 A.42. The singer who met the dancer previously sings with the city orchestra.

 1566 Ibade, المطرب الذي استقبل الراقص سابقاً يغنى مع أوركسترا المدينة.
- 1567 A.43. The producer who enthusiastically chose the actress produces a film every month.
 1568 المخرج الذي اختار الممثل بشغف ينتج فيلما كل شهر.
- A.44. The maid who helped the caregiver earnestly cleans the rooms seven days a week. الذي ساعد المربي بجدّ ينظف الغرف سبعة أيام في الأسبوع.
- 1571 A.45. The announcer who invited the historian nicely presented a program on TV.
 1572 المذيع الذي دعا المؤرّخ بلطف قدم برنامجاً في التلفزيون.
- A.46. The assistant who served the pharmacist loyally collected all of the old reports.

 1573 المساعد الذي خدم الصيدلانيّ بإخلاص جمع كل التقارير القديمة.
- 1575 A.47. The journalist who interviewed the winner by chance writes in many newspapers.
 1576 الصحفى الذي قابل الفائز صدفةً يكتب في صحف كثيرة.
- 1577 A.48. The bedouin who visited the farmer at night lives in the middle of the desert.

 1578 Ities البدوي الذي زار المُزارع ليلاً يسكن في وسط الصحراء.

1579 Appendix B. Complete Materials — Experiment 3

- B.1. The child who watched the magician with amazement applauded hysterically during the show.
- (السحرة) الطفل الذي رأى الساحر بانبهار صفق بشدة خلال العرض.
- B.2. The investigator who grabbed the robber at night installs listening devices everywhere.
- (الخونة) المحقق الذي أمسك الخائن بالليل يضع أجهزة تنصت في كل مكان.
- B.3. The young man who helped the wounded man in the morning every week reading. (الجرحى) الشاب الذي ساعد الجريح صباحاً يقضي كل الأسبوع في المطالعة.
- B.4. The doctor who healed the captive quickly is joining the national guard (army). الطبيب الذي عالج الأسير بسرعةٍ ينخرط في فرقة الدفاع المدني.
- B.5. The ruler who freed the slave in the past practiced justice and equality among the people.
- (العبيد) الحاكم الذي أعتق العبد قديمًا مارس العدل والمساواة على الشعب.
- B.6. The driver who transported the pilgrim at noon drove the car very quickly.

 المجيج السائق الذي أخذ الحاج بالظهيرة قاد السيارة بسرعة فائقة.

B.7. The carpenter who employed the man regularly made wonderful furniture for the exhibition.

1597

1603

1608

1611

1621

(الرجال) النجار الذي شغّل الرجل بانتظام أعدّ قطع أثاث رائعة للمعرض.

B.8. The coach who trained the partner efficiently possesses more than 10-years of experience.

(الرفاق) المدرب الذي درّب الرفيق بإتقان يمتلك خبرة تفوق العشر سنين.

B.9. The tv-host who presented the bridegroom skillfully met with the minister at the New Year's party.

(العرسان) المذيع الذي قدّم العريس ببراعة قابل الوزير في حفل رأس السنة.

B.10. The Imam who advised the neighbor wisely lives in the local neighborhood. الجيران) الإمام الذي نصح الجار بحكمة يعيش في المنطقة المجاورة.

B.11. The policeman who questioned the murderer harshly asked the pedestrians about the reasons for the crime.

(القتلة) الشرطي الذي استجوب القاتل بجديّة سأل المشاة عن سبب الجريمة.

B.12. The lawyer who called the inheritor in the morning discussed the issue of the inheritance distribution.

(الورثة) المحامي الذي استدعى الوريث بالصباح ناقش موضوع تقسيم الإرث.

B.13. The teacher who taught the student well worked as a host in television.

(الطلاب) الأستاذ الذي علّم الطالب جيداً عمل كمذبع في التلفزيون.

- B.14. The businessman who trained the workers frequently learned English in the UK. (العمال) التاجر الذي درّب العامل كثيراً تعلم اللغة الانجليزية في بريطانيا .
- B.15. The (football) player who admonished the referee angrily won the prize of best player.

(الحكام) اللاعب الذي عارض الحَكم بغضب فاز بجائزة أفضل لاعب.

B.16. The immigrant who spoke with the visitor for a long time feels nostalgic for the country always.

(الزوار) المهاجر الذي حدّث الزائر طويلاً يشعر بالجنين للوطن دائماً.

B.17. The journalist who interviewed the leader persistently publishes the article in the national newspaper.

(الحكام) الصحافي الذي استجوب الحاكم بالحاح نشر المقال في الجريدة الوطنية.

B.18. The pilot who greeted the knight warmly asked many questions to ask during the journey.

(الفرسان) الطيار الذي حيًا الفارس بحرارة سأل أسئلة كثيرة خلال الرحلة.

B.19. The man who talked to the monk intelligently works in the post office. (الرهبان) الرجل الذي كلم الراهب بفطنة يشتغل في مكتب البريد.

B.20. The policeman who arrested the thief quickly helps the pedestrians in crossing the street.

(اللصوص) الشرطي الذي اعتقل اللص بسرعة يساعد المارة على عبور الشارع.

B.21. The lawyer who startled the witness cunningly stopped the accusation of his client in the court.

(الشهود) المحامي الذي أربك الشاهد بدهاء أبعد التهم عن موكله في المحكمة.

- 1636 B.22. The worker who helped the soldier yesterday drives a large truck for the company.

 (الجنود) العامل الذي ساعد الجندي بالأمس يَسوق شاحنة كبيرة لفائدة الشركة.
- B.23. The artist who served the king devotedly gave a portrait to the ambassador of the United States.

(الملوك) الفنان الذي خدم الملك بتفان تبرع بلوحة لسفير الولايات المتحدة.

- B.24. The journalist who hosted the star brilliantly raised very embarrassing questions. (النجوم) الإعلامي الذي استضاف النجم بتألق طرح أسئلة محرجة جداً.
- B.25. The translator who worked for the manager occasionally speaks five languages fluently.

(المدراء) المترجم الذي ساعد المدير أحياناً يتكلم خمس لغات بفصاحة.

B.26. The student who saw the professor yesterday studied electrical engineering at the university.

1648

(الأساتذة) الطالب الذي رأى الأستاذ بالأمس درس الهندسة الكهربائية في الجامعة.

- B.27. The engineer who met the scientist by chance is working on a new invention.

 1650 العلماء) المهندس الذي استقبل العالم بالصدفة يعمل على ابتكار جديد.
- B.28. The cook who scolded the waiter forcefully works in an expensive restaurant during the summer.

(النوادل) الطباخ الذي وبّخ النادل بشدة يشتغل في مطعم غال خلال الصيف.

B.29. The analyst who advised the minister intelligently discusses the Palestinian issue in depth.

(الوزراء) المحلل الذي نصح الوزير بذكاء يتناول القضية الفلسطينية بعمق.

- 1657 B.30. The child who saw the prince before visits the royal family each week.
 1658 لأمراء) الطفل الذي رأى الأمير سلفاً يزور العائلة الملكية كل أسبوع.
- B.31. The teacher who taught the child dedicatedly attended the graduation party of the students.

(الأطفال) المعلم الذي علّم الطفل بتفانٍ حضر حفل تخرج الطلاب.

B.32. The criminal who attacked the boy viciously breaks through the checkpoint every night.

(الأولاد) المجرم الذي هاجم الولد بشراسة يخترق نقطة التفتيش كل ليلة.

```
B.33. The consultant who warned the president yesterday found a solution for the financial problem.
```

(الرؤساء) المستشار الذي حذّر الرئيس بالأمس وجد الحل للمشكلة المالية.

- B.34. The driver who accompanied the ambassador regularly works seven days a week. (السفراء) السائق الذي رافق السفير بانتظام يعمل سبعة أيام في الأسبوع.
- B.35. The jailor who tortured the prisoner constantly cleans the cells nightly. السجّان الذي عذّب السجين باستمرار ينظف الزنازين كل ليلة.
- B.36. The broadcaster who talked to the activist yesterday trained in a famous company. (النشطاء) المذيع الذي كلم الناشط بالأمس تدرب في شركة مشهورة.
- B.37. The employee who helped the colleague humbly gained the confidence of colleagues at work.

(الزملاء) الموظف الذي ساعد الزميل بتواضع استطاع كسب ثقة باقى الزملاء في العمل.

- 1677 B.38. The singer who challenged the poet arrogantly has lost for not showing up on time.
- (الشعراء) المنشد الذي تحدى الشاعر بغرور خسر لعدم حضوره في الوقت المحدد.
- 1680 B.39. The man who consulted the forgiver yesterday wants retribution of sin/guilt.

 1681 الرجل الذي استشار الشفيع البارحة يريد التكفير عن الذنب.
- B.40. The old man who has mischievously insulted the scholar strives to create problems.

1684

(الفقهاء) العجوز الذي أهان الفقيه بخبث يسعى الى افتعال المشاكل.

- 1685 B.41. The beginner who questioned the expert daily has acquired good experience.

 1686 الخُبراء) المبتدئ الذي سئل الخبير يومياً اكتسب خبرة جيدة.
- B.42. The man who hosted the friend with pleasure slept in the basement of the house. (الأصدقاء) الرجل الذي استضاف الصديق بسرور نام في الطابق السغلي من المنزل.
- 1699 B.43. The patient who consulted the doctor yesterday returned home satisfied.
 1690 (الأطباء) المريض الذي استشار الطبيب بالأمس عاد الى المنزل مطمئن.
- 1691 B.44. The teacher who met the writer happily likes reading about literature.
 1692 الأدباء) المعلم الذي استقبل الأديب بسعادة يحب القراءة عن الأدب.
- B.45. The representative who talked to the Khalif yesterday works hard to get a promotion.

(الخلفاء) النائب الذي كلم الخليفة بالأمس يجتهد في العمل للحصول على ترقية.

- 1696 B.46. The president who hurriedly called the ally tries to reign the situation.
 1697 الرئيس الذي استدعى الحليف باستعجال يحاول التحكم بزمام الأمور.
- B.47. The manager who phoned the partner in the morning plans to expand the branches of the company.
- (الشركاء) المدير الذي هاتف الشريك بالصباح يخطط لتوسيع فروع الشركة.

B.48. The chef who invited the guest in the evening masters preparing various delicious (النزلاء) الطاهي الذي دعا النزيل بالمساء يتقن إعداد أصناف لذيذة من الطعام. 1702

Appendix C. Complete Materials — Experiment 4

1703

1713

1719

- C.1. The nurse who is treating the patient carefully studies at the university hospital. 1704 الممرضة التي عالجت المريضة بعنايةٍ تدرس في مستشفى الجامعة. 1709
- C.2. The queen who looked after the princess recently appears in public every week. 1706 الملكة التي اهتمت بالأميرة حديثاً تظهر في العلن كل أسبوع. 1707
- C.3. The seller who thanked the customer enthusiastically was happy with the large 1708 profit. 1709 البائعة التي شكرت الزبونة بحماس فرحت بالفائدة الكبيرة. 1710
- C.4. The novelist who mentioned the maid of honor accurately sells many books to 1711 the public. 1712 الكاتبة التي ذكرت الوصيفة بدقة تبيع كتبا كثيرة للجمهور.
- C.5. The coach who was very interested in the player worked at the National Fencing 1714
- 1715 Academy. المدرية التي اهتمت باللاعبة جداً اشتغلت في الأكاديمية الوطنية للمبارزة. 1716
- C.6. The midwife who cared for the girl repeatedly volunteers at the university hospi-1717 1718 القابلة التي اعتنت بالفتاة تكراراً تتطوع في مستشفى الجامعة.
- C.7. The soldier (fem.) who met the policewoman yesterday loved the atmosphere at 1720 the air base. 1721
- الجندية التي قابلت الشرطية بالأمس أحبّت العمل في القاعدة الجوية . 1722
- C.8. The singer who met the dancer previously sings with the city orchestra 1723 المطربة التي استقبلت الراقصة سابقاً تغنى مع أوركسترا المدينة. 1724
- C.9. The nanny who cared for the schoolgirl affectionately traveled to a new country. 1725 المربية التي ربّت الطالبة بحنان سافرت إلى بلد جديد. 1726
- C.10. The producer who enthusiastically chose the actress produces a film every month. المخرجة التي اختارت الممثلة بشُغف تنتج فيلما كل شهر. 1728
- C.11. The photographer who photographed the witch artistically published the photos 1729 in a new book. 1730
- المصورة التي صورت الساحرة بتفنن نشرت الصور في كتاب جديد. 1731
- C.12. The maid who helped the nanny earnestly cleans the rooms seven days a week. 1732 الخادمة التي ساعدت المربية بجدِّ تنظف الغرف سبعة أيام في الأسبوع. 1733

- C.13. The announcer who invited the historian nicely presented a program on TV. المذيعة التي دعت المؤرّخة بلطف قدمت برنامجاً في التلفاز. 1735
- C.14. The model who met the accountant repeatedly owns a lot of expensive clothes. 1736 العارضة التى التقت بالمحاسبة تكراراً تملك كثيرا من الملابس الغالية.
- C.15. The assistant who served the pharmacist loyally collected all of the old reports. 1738 المساعِدة التي خدمت الصيدلانيّة بإخلاصِ جمعت كل التقارير القديمة. 1739
- C.16. The journalist who interviewed the winner by chance writes in many newspapers. الصحفية التي قابلت الفائزة صدفة تكتب في صحف كثيرة. 1741
- C.17. The bedouin who visited the farmer at night goes to the middle of the desert. 1742 البدوية التي زارت المزارعة ليلاً تذهب إلى وسط الصحراء. 1743
- C.18. The doctor who treated the girl recently discovered a cure for the terrible disease. 1744 الطبيبة التي عالجت الطفلة مؤخراً اكتشفت شفاء للمرض الرهيب. 1745
- C.19. The artist who corresponded with the publisher eagerly desired a new contract. الفنانة التي راسلت الناشرة بشغف رغبت في عَقد جديد. 1747
- C.20. The student who admired the poet greatly read many poems last year. 1748 التلميذة التي أُعجبت بالشاعرة بشدة قرأت قصائد كثيرة العام الماضي. 1749
- C.21. The director who contacted the author during the day supervises many large 1750 projects. المديرة التي اتصلت بالمؤلفة نهاراً تشرف على كثير من المشاريع الكبيرة.

1752

1758

- C.22. The dean who summoned the professor angrily observed a problem in the uni-1753 versity departments 1754
- العميدة التي استدعت الأستاذة بغضب لاحظت خللاً في أقسام الجامعة.
- C.23. The musician who accompanied the singer professionally played with the na-1756 tional music group الموسيقيّة التي رافقت المغنية بمهنيّة عزفت مع الفرقة الوطنية للموسيقي.
- C.24. The ambassador who hosted the delegate yearly spoke at the United Nations. 1759
- السفيرة التي استضافت المندوبة سنوياً تحدثت في الأمم المتحدة. C.25. The grandmother who met the neighbor suddenly talked about the neighborhood 1761
- issues. 1762 الجدّة التي صادفت الجارة فجأةً تحاورت عن أمور الحي. 1763
- C.26. The student who met the manager yesterday got high grades in the remaining 1764 subjects. 1765
- الطالبة التي قابلت المديرة البارحة نالت درجات عالية في المواد المتبقية. 1766

C.27. The accountant who talked to the employee harshly suffered from social problems. 1768 المحاسبة التي حادثت الموظفة بصرامة عانت من مشاكل اجتماعية. 1769 C.28. The study abroad student who thanked the official a lot studied at one of the best international universities. 1771 المبتعثة التي شكرت المسؤولة بكثرة درست باحدى أرقى الجامعات الدولية. 1772 C.29. The graduate who talked to the lecturer happily works for extra hours at the li-1773 brary. الخريجة التي كلمت المحاضرة بسعادة تعمل ساعات اضافية في المكتبة. 1775 C.30. The painter who excitedly interviewed the director painted wonderful paintings 1776 الرسامة التّي حاورت المخرجة بإثّارة رسمت لوحات فنية رائعة. C.31. The chef who lived next to the trader for a long time practices a cooking career 1778 skillfully. الطاهية التي جاورت التاجرة مطولاً تمارس مهنة الطبخ بمهارة. 1780 C.32. The visitor who talked to the guide in the morning gave a lecture about how to 1781 manage time. 1782 الزائرة التي حدثت المرشدة صباحاً القت محاضرة عن كنفية تنظيم الوقت. 1783 C.33. The teacher who visited the doctor yesterday masters speaking in Arabic and 1784 English. 1785 المعلمة التي زارت الطبيبة بالأمس تُجيد التحدث باللغة العربية و الإنجليزية. C.34. The lawyer who accused the guilty person angrily is trying to find the way to the 1787 truth. 1788 المحامية التي اتهمت المذنبة بغضب تحاول الوصول إلى الحقيقة. 1789 C.35. The engineer who met the colleague daily aspires to get a job at a prominent 1790 company. 179 المهندسة التي قابلت الزميلة يومياً تسعى للحصول على وظيفة في شركة مرموقة. 1792 C.36. The farmer who blamed the young lady yesterday loves working at the farm near 1793 the park. 1794 المزارعة التي لامت الشابة بالأمس تحب العمل في المزرعة المجاورة للحديقة. C.37. The beginner who helped the boss in the morning was hired for the military com-1796 1797 المبتدئة التي ساعدت الرئيسة صباحاً توظفت في الشركة العسكرية. 1798 C.38. The actress who met the interviewer in the past resigned from the acting career

1802 C.39. The guard who talked to the pupil in the morning goes home late every day. 1803 الحارسة التي كلمت التلميذة صباحاً تذهب إلى المنزل في ساعة متاخرة كل يوم.

recently.

1800

1801

الممثلة التي قابلت المذيعة بالماضي استقالت من مهنة التمثيل مؤخراً.

- 1804 C.40. The worker who gently scolded the maid cares about helping the needy.
 1805 lbala العاملة التي وبخت الخادمة برفق تهتم بمساعدة المحتاجين.
- ¹⁸⁰⁶ C.41. The analyst who patiently waited for the reporter is trying to educate people about the importance of a clean environment.
- المحللة التي انتظرت المراسلة بصبر تسعى لتوعية الشعب على أهمية نظافة البيئة.
- ¹⁸⁰⁹ C.42. The reporter who spoke to the plaintiff adeptly interviews the president at international conferences.
- المراسلة التي سألت المدعية بنباهة تقابل رئيس الدولة في المؤتمرات العالمية.
- C.43. The magician who talked to the lady quickly worked at the theater near the village.
- الساحرة التي كلمت السيدة بسرعة عملت في المسرح المجاور للقرية.
- ¹⁸¹⁵ C.44. The employee who accompanied the visitor in the morning stays at work until late.
- الموظفة التي رافقت الزائرة بالصباح تبقي حتى ساعة متاخرة في العمل.
- ¹⁸¹⁸ C.45. The novelist who pleasantly shook hands with the designer writes international and local novels about literature.
 - الروائية التي صافحت المصممة بسرور تكتب روايات عالمية و محلية في الادب.
- C.46. The researcher who calmly called the detective provides money for charity society
 - الباحثة التي هاتفت المحققة بهدوء توفر المال للجمعيات الخيرية.
- 1824 C.47. The coordinator who helped the guard devotedly was in the school courtyard.
 1825 المنسقة التي عاونت الحارسة بتفان تواجدت في ساحة المدرسة.
- C.48. The candidate who pleasantly thanked the participant took part in the electoral campaign.
- المرشحة التي شكرت المشتركة بسرور شاركت في الحملة الانتخابية.
- C.49. The judge who decisively questioned the thief ruled fairly among people.

 القاضية التي سألت السارقة بحزم حكمت بالعدل بين الناس.
- C.50. The immigrant who answered the inspector anxiously faced difficulties at the check point.
- المهاجرة التي أجابت المفتشة بقلق واجهت صعوبات عند نقطة التفتيش.
- C.51. The young girl who helped the grandmother at night works at bakery for sweets. الفتاة التي ساعدت الجدة بالليل تعمل في مخبز الحلويات.
- C.52. The tourist who met the driver on the road loves traveling to different countries. السائحة التي صادفت السائقة بالطريق تحب السفر إلى بلدان مختلفة.
- 1838 C.53. The princess who intelligently answered the journalist owns many huge palaces. الأميرة التي أجابت الصحافية بذكاء تمتلك عدة قصور كبيرة.
- C.54. The client who consulted the lawyer nervously practiced painting for a few years. الموكلة التي استشارت المحامية بتوتر مارست مهنة الرسم لعدة سنوات.

Appendix D. Supplemental Data & Model Results

1843

1844

1845

This appendix contains tables of grand average raw reading times and mixed-effects model results for regions not immediately germane to the main claims of the paper for experiments 1–4. In all cases, these values are reported identically to the analysis procedures outlined in the *Analysis* section in the main text for each experiment.

Condition	Mean	SD		
Subject Reg	Subject Region			
MATCH/GRAM	422	126		
MATCH/UNGRAM	426	126		
NoMatch/Gram	420	128		
NoMatch/Ungram	427	141		
Complementizer	Region			
MATCH/GRAM	387	105		
MATCH/UNGRAM	389	116		
NoMatch/Gram	387	101		
NoMatch/Ungram	387	110		
Relative Clause Ve.	rb Regior	η		
MATCH/GRAM	378	107		
MATCH/UNGRAM	377	117		
NoMatch/Gram	382	113		
NoMatch/Ungram	382	120		
Attractor Reg	gion			
MATCH/GRAM	453	195		
MATCH/UNGRAM	436	179		
NoMatch/Gram	452	185		
NoMatch/Ungram	456	187		
Adverb Reg	ion			
MATCH/GRAM	502	205		
MATCH/UNGRAM	501	194		
NoMatch/Gram	514	192		
NoMatch/Ungram	509	194		
Second Spillover	Region			
MATCH/GRAM	408	101		
MATCH/UNGRAM	435	116		
NoMatch/Gram	406	100		
NoMatch/Ungram	439	136		

Table D.1: Raw condition grand avergage reading times across participant means in milliseconds for additional regions in Experiment 1.

Coefficient	$\hat{oldsymbol{eta}}$	SE	t
Relative Clause	Verb Regi	on	
Intercept	331.98	14.52	22.86
Match:No	5.82	6.01	0.97
Gram:Ungram	2.66	6.07	0.44
Item Order	-0.58	0.04	-14.50
Length	6.55	2.30	2.84
Previous Region RT	0.18	0.01	19.25
Match:No × Gram:UNGRAM	-2.35	8.69	-0.27
Attractor	Region		
Intercept	328.06	52.26	6.28
Match:No	-29.02	13.76	-2.11
Gram:Ungram	-17.73	11.49	-1.54
Item Order	-0.98	0.08	-12.95
Length	25.18	7.91	3.18
Previous Region RT	0.17	0.02	10.78
Match:No × Gram:UNGRAM	23.66	16.45	1.44
Adverb R	Region		
Intercept	599.34	53.64	11.17
Match:No	8.35	11.99	0.70
Gram:Ungram	-1.84	12.11	-0.15
Item Order	-1.24	0.08	-15.49
Length	0.99	7.94	0.12
Previous Region RT	0.04	0.01	3.42
Match:No × Gram:UNGRAM	2.18	17.34	0.13
Second Spillo	ver Region	!	
Intercept	433.27	15.51	27.93
Match:No	-2.44	5.75	-0.42
Gram:Ungram	26.51	5.84	4.54
Item Order	-0.74	0.04	-19.27
Length	5.41	1.98	2.74
Previous Region RT	0.05	0.01	7.12
Match:No × Gram:UNGRAM	5.28	8.32	0.63

Table D.2: Mixed effects regression coefficients for other regions in Experiment 1. Significant coefficients (|t| > 2) are in bold and marginal coefficients (|t| > 1.65) are in italics.

Condition	Mean	SD
Subject Region	:	
MASC/MATCH/GRAM	402	131
MASC/MATCH/UNGRAM	411	158
MASC/NOMATCH/GRAM	406	148
MASC/NOMATCH/UNGRAM	415	147
FEM/NOMATCH/UNGRAM	422	162
FEM/NOMATCH/GRAM	428	187
FEM/MATCH/UNGRAM	414	151
FEM/MATCH/GRAM	428	179
Complementizer Re		1,,
MASC/MATCH/GRAM	368	107
MASC/MATCH/UNGRAM	360	103
MASC/NOMATCH/GRAM	365	125
MASC/NOMATCH/UNGRAM	372	129
FEM/NOMATCH/UNGRAM	381	122
FEM/NOMATCH/GRAM	386	127
FEM/MATCH/UNGRAM	375	113
FEM/MATCH/GRAM	381	128
Relative Clause Verb		120
MASC/MATCH/GRAM	354	119
MASC/MATCH/UNGRAM	356	125
MASC/NOMATCH/GRAM	357	118
MASC/NOMATCH/UNGRAM	363	129
FEM/NOMATCH/UNGRAM	360	122
FEM/NOMATCH/GRAM	364	123
FEM/MATCH/UNGRAM	363	128
FEM/MATCH/GRAM	374	139
Attractor Region		
MASC/MATCH/GRAM	369	141
MASC/MATCH/UNGRAM	390	161
MASC/NOMATCH/GRAM	404	174
MASC/NOMATCH/UNGRAM	410	171
FEM/NOMATCH/UNGRAM	404	153
FEM/NOMATCH/GRAM	401	141
FEM/MATCH/UNGRAM	407	176
FEM/MATCH/GRAM	415	175
Adverb Region		
MASC/MATCH/GRAM	432	176
MASC/MATCH/UNGRAM	434	191
MASC/NOMATCH/GRAM	449	215
MASC/NOMATCH/GRAM	458	205
		160
MASC/NOMATCH/UNGRAM	424	100
Masc/NoMatch/Ungram Fem/NoMatch/Ungram	424 426	
Masc/NoMatch/Ungram Fem/NoMatch/Ungram Fem/NoMatch/Gram		169 194
Masc/NoMatch/Ungram Fem/NoMatch/Ungram Fem/NoMatch/Gram Fem/Match/Ungram	426	169
Masc/NoMatch/Ungram Fem/NoMatch/Ungram Fem/NoMatch/Gram Fem/Match/Ungram	426 449 452	169 194
MASC/NOMATCH/UNGRAM FEM/NOMATCH/UNGRAM FEM/NOMATCH/GRAM FEM/MATCH/UNGRAM FEM/MATCH/GRAM Second Spillover Re	426 449 452	169 194 181
MASC/NOMATCH/UNGRAM FEM/NOMATCH/UNGRAM FEM/NOMATCH/GRAM FEM/MATCH/UNGRAM FEM/MATCH/GRAM Second Spillover Re MASC/MATCH/GRAM	426 449 452 egion	169 194 181 104 135
MASC/NOMATCH/UNGRAM FEM/NOMATCH/UNGRAM FEM/NOMATCH/GRAM FEM/MATCH/UNGRAM FEM/MATCH/GRAM Second Spillover Re MASC/MATCH/GRAM	426 449 452 egion 365	169 194 181 104 135
MASC/NOMATCH/UNGRAM FEM/NOMATCH/UNGRAM FEM/NOMATCH/GRAM FEM/MATCH/UNGRAM FEM/MATCH/GRAM Second Spillover Re MASC/MATCH/GRAM MASC/MATCH/UNGRAM MASC/MATCH/UNGRAM MASC/NOMATCH/GRAM	426 449 452 egion 365 393	169 194 181 104 135 105
MASC/NOMATCH/UNGRAM FEM/NOMATCH/UNGRAM FEM/NOMATCH/UNGRAM FEM/MATCH/UNGRAM FEM/MATCH/GRAM Second Spillover Re MASC/MATCH/GRAM MASC/MATCH/UNGRAM	426 449 452 egion 365 393 368	169 194 181
MASC/NOMATCH/UNGRAM FEM/NOMATCH/UNGRAM FEM/NOMATCH/UNGRAM FEM/MATCH/UNGRAM FEM/MATCH/GRAM Second Spillover Re MASC/MATCH/GRAM MASC/MATCH/UNGRAM MASC/NOMATCH/UNGRAM MASC/NOMATCH/UNGRAM	426 449 452 egion 365 393 368 384	169 194 181 104 135 105 122
MASC/NOMATCH/UNGRAM FEM/NOMATCH/UNGRAM FEM/NOMATCH/GRAM FEM/MATCH/UNGRAM FEM/MATCH/GRAM Second Spillover Re MASC/MATCH/GRAM MASC/MATCH/UNGRAM MASC/NOMATCH/UNGRAM MASC/NOMATCH/UNGRAM FEM/NOMATCH/UNGRAM	426 449 452 <i>egion</i> 365 393 368 384 383	169 194 181 104 135 105 122 122

Table D.3: Raw condition grand avergage reading times across participant means in milliseconds for additional regions in Experiment 2.

Coefficient	β	SE	t
Relative Clause Verb	Region		
Intercept	328.45	14.67	22.39
Subj:Fem	11.39	8.05	1.41
Match:No	1.81	7.86	0.23
Grammaticality:UNGRAM	1.02	7.85	0.13
Item Order	-0.52	0.04	-14.26
Length	4.94	2.28	2.16
Previous Region RT	0.15	0.01	17.66
Subj:FEM × Match:No	-13.32	11.05	-1.21
Subj:FEM × Gram:UNGRAM	-10.29	11.08	-0.93
Match:No × Gram:UNGRAM	-1.08	11.10	-0.10
Subj:FEM × Match:No × Gram:UNGRAM	10.19	15.64	0.65
Attractor Regio	n		
Intercept	305.26	35.77	8.53
Subj:Fem	28.15	11.44	2.46
Match:No	24.11	11.50	2.10
Grammaticality:UNGRAM	16.56	10.20	1.62
Item Order	-0.78	0.05	-16.64
Length	9.73	5.32	1.83
Previous Region RT	0.23	0.01	18.04
Subj:FEM × Match:No	-26.58	17.80	-1.49
Subj:FEM × Gram:UNGRAM	-17.19	14.39	-1.19
Match:No × Gram:UNGRAM	-18.19	14.42	-1.26
Subj:FEM × Match:No × Gram:UNGRAM	23.91	20.32	1.18
Adverb Region	!		
Intercept	390.77	27.92	14.00
Subj:FEM	10.26	11.90	0.86
Match:No	8.04	11.96	0.67
Grammaticality:UNGRAM	-4.38	11.94	-0.37
Item Order	-1.04	0.06	-18.78
Length	14.55	3.51	4.15
Previous Region RT	0.13	0.01	11.41
Subj:FEM × Match:No	-33.04	16.80	-1.97
Subj:FEM × Gram:UNGRAM	6.14	16.85	0.36
Match:No × Gram:UNGRAM	13.64	16.88	0.81
Subj:FEM × Match:No × Gram:UNGRAM	-13.56	23.79	-0.57
Second Spillover Re	egion		
Intercept	347.19	13.05	26.60
Subj:FEM	-6.94	6.38	-1.09
Match:No	0.75	6.41	0.12
Grammaticality:UNGRAM	16.31	6.46	2.53
Item Order	-0.57	0.03	-18.84
Length	3.97	1.60	2.48
Previous Region RT	0.15	0.01	16.56
Subj:FEM × Match:No	16.60	9.01	1.84
$Subj.1$ EM $\wedge Mulch.100$			
	12.74	9.05	1.41
Subj:FEM × Gram:UNGRAM Match:No × Gram:UNGRAM	12.74 -6.94	9.05 9.07	1.41 -0.77

Table D.4: Mixed effects regression coefficients for other regions in Experiment 2. Significant coefficients (|t| > 2) are in bold and marginal coefficients (|t| > 1.65) are in italics.

Condition	Mean	SD	
Subject Region			
MATCH/GRAM	490	187	
MATCH/UNGRAM	510	213	
NoMatch/Gram	494	178	
NoMatch/Ungram	490	184	
Complementizer	Region		
MATCH/GRAM	431	132	
MATCH/UNGRAM	445	148	
NoMatch/Gram	428	123	
NoMatch/Ungram	447	155	
Relative Clause Ve	rb Regior	i	
MATCH/GRAM	435	159	
MATCH/UNGRAM	426	139	
NoMatch/Gram	426	138	
NoMatch/Ungram	432	135	
Attractor Reg	gion		
MATCH/GRAM	485	211	
MATCH/UNGRAM	497	215	
NoMatch/Gram	503	210	
NoMatch/Ungram	497	196	
Adverb Reg	ion		
MATCH/GRAM	565	270	
MATCH/UNGRAM	542	227	
NoMatch/Gram	552	224	
NoMatch/Ungram	546	231	
Second Spillover	Region		
MATCH/GRAM	428	108	
MATCH/UNGRAM	448	119	
NoMatch/Gram	425	95	
NoMatch/Ungram	443	113	

Table D.5: Raw condition grand avergage reading times across participant means in milliseconds for additional regions in Experiment 3.

Coefficient	\hat{eta}	SE	t
Relative Clause Verb Re	egion		
Intercept	435.70	18.42	23.66
Ambiguity: AMBIG	0.94	11.73	0.08
Match:No	-5.74	11.48	-0.50
Grammaticality:UNGRAM	-14.26	11.44	-1.25
Item Order	-1.01	0.05	-18.31
Length	11.50	2.91	3.95
Previous Region RT	0.10	0.01	13.31
Amb:UNAM × Match:No	-4.77	16.60	-0.29
Amb:Unam × Gram:Ungram	7.02	16.61	0.42
Match:No × Gram:UNGRAM	2.26	16.33	0.14
Amb:UNAM × Match:No × Gram:UNGRAM	17.43	23.65	0.74
Attractor Region			
Intercept	579.38	66.19	8.75
Ambiguity: AMBIG	-5.92	18.53	-0.32
Match:No	22.78	16.87	1.35
Grammaticality:UNGRAM	4.07	15.91	0.26
Item Order	-1.40	0.08	-18.16
Length	-8.42	11.04	-0.76
Previous Region RT	0.20	0.01	14.45
Amb:UNAM × Match:No	9.56	23.99	0.40
Amb:Unam × Gram:Ungram	19.55	23.09	0.85
Match:No × Gram:UNGRAM	-8.68	22.71	-0.38
Amb:UNAM × Match:No × Gram:UNGRAM	-24.95	32.88	-0.76
Adverb Region			
Intercept	564.51	45.75	12.34
Ambiguity:Ambig	-11.72	20.19	-0.58
Match:No	-37.33	18.17	-2.05
Grammaticality:UNGRAM	-19.24	18.11	-1.06
Item Order	-1.72	0.09	-19.42
Length	16.93	6.15	2.75
Previous Region RT	0.13	0.01	9.76
Amb:UNAM × Match:No	34.98	26.29	1.33
Amb:Unam × Gram:Ungram	-14.30	26.30	-0.54
Match:No × Gram:UNGRAM	24.38	25.85	0.94
Amb:UNAM × Match:NO × Gram:UNGRAM Second Spillover Reg	-6.25	37.43	-0.17
	467.28	14.29	32.71
Intercent		17.47	J4./1
Intercept Ambiguity: AMBIG		9.94	0.90
Ambiguity: AMBIG	8.94	9.94 7.45	0.90
Ambiguity:AMBIG Match:NO	8.94 -2.25	7.45	-0.30
Ambiguity:Ambig Match:NO Grammaticality: Ungram	8.94 -2.25 20.02	7.45 7.43	-0.30 2.69
Ambiguity:AMBIG Match:NO Grammaticality:UNGRAM Item Order	8.94 -2.25 20.02 -1.02	7.45 7.43 0.04	-0.30 2.69 -27.72
Ambiguity:AMBIG Match:NO Grammaticality:UNGRAM Item Order Length	8.94 -2.25 20.02 -1.02 2.28	7.45 7.43 0.04 1.80	-0.30 2.69 - 27.72 1.27
Ambiguity:AMBIG Match:NO Grammaticality:UNGRAM Item Order Length Previous Region RT	8.94 -2.25 20.02 -1.02 2.28 0.10	7.45 7.43 0.04 1.80 0.01	-0.30 2.69 -27.72 1.27 11.72
Ambiguity:Ambig Match:NO Grammaticality:UNGRAM Item Order Length Previous Region RT Amb:UNAM × Match:NO	8.94 -2.25 20.02 - 1.02 2.28 0.10 -7.49	7.45 7.43 0.04 1.80 0.01 10.77	-0.30 2.69 - 27.72 1.27 11.72 -0.70
Ambiguity:AMBIG Match:NO Grammaticality:UNGRAM Item Order Length Previous Region RT	8.94 -2.25 20.02 -1.02 2.28 0.10	7.45 7.43 0.04 1.80 0.01	-0.30 2.69 -27.72 1.27 11.72

Table D.6: Mixed effects regression coefficients for other regions in Experiment 3. Significant coefficients (|t| > 2) are in bold and marginal coefficients (|t| > 1.65) are in italics.

Subject Region	Condition	Mean	SD
SG/MATCH/UNGRAM	Subject Regio	on	
SG/NOMATCH/GRAM 452 172 SG/NOMATCH/UNGRAM 434 159 PL/NOMATCH/UNGRAM 476 212 PL/NOMATCH/UNGRAM 455 180 PL/MATCH/UNGRAM 455 181 PL/MATCH/GRAM 457 217 Complementizer Region SG/MATCH/UNGRAM 387 106 SG/MATCH/UNGRAM 387 113 SG/NOMATCH/UNGRAM 399 129 PL/NOMATCH/UNGRAM 393 126 PL/NOMATCH/UNGRAM 393 126 PL/MATCH/UNGRAM 406 129 PL/NOMATCH/UNGRAM 400 129 PL/MATCH/GRAM 400 142 Relative Clause Verb Region SG/NOMATCH/UNGRAM 370 118 SG/NOMATCH/UNGRAM 364 117 SG/NOMATCH/UNGRAM 366 116 SG/NOMATCH/UNGRAM 369 129 PL/MATCH/GRAM 390 129 PL/MATCH/UNGRAM 390 129 PL/MATCH/UNGRAM 391 135 PL/NOMATCH/UNGRAM 391 135 PL/NOMATCH/UNGRAM 404 149 Attractor Region SG/MATCH/UNGRAM 404 149 Attractor Region SG/NOMATCH/UNGRAM 420 176 SG/NOMATCH/UNGRAM 420 176 SG/NOMATCH/UNGRAM 420 176 SG/NOMATCH/UNGRAM 410 152 SG/NOMATCH/UNGRAM 410 152 SG/NOMATCH/UNGRAM 410 152 SG/NOMATCH/UNGRAM 410 152 PL/NOMATCH/GRAM 410 153 SG/NOMATCH/UNGRAM 410 154 SG/NOMATCH/UNGRAM 410 157 PL/NOMATCH/GRAM 450 187 Adverb Region SG/MATCH/UNGRAM 450 187 Adverb Region SG/MATCH/UNGRAM 451 175 PL/NOMATCH/UNGRAM 452 360 360 360 360 360 360 360 360 360 360 360 360 36	SG/MATCH/GRAM	422	146
SG/NOMATCH/UNGRAM 434 159 PL/NOMATCH/UNGRAM 476 212 PL/NOMATCH/GRAM 455 180 PL/MATCH/UNGRAM 455 181 PL/MATCH/UNGRAM 457 217	SG/MATCH/UNGRAM	449	175
SG/NOMATCH/UNGRAM 434 159 PL/NOMATCH/UNGRAM 476 212 PL/NOMATCH/GRAM 455 180 PL/MATCH/UNGRAM 455 181 PL/MATCH/UNGRAM 457 217	SG/NoMatch/Gram	452	172
PL/NOMATCH/GRAM		434	159
PL/MATCH/UNGRAM	PL/NoMatch/Ungram	476	
PL/MATCH/GRAM	PL/NoMatch/Gram	455	180
SG/MATCH/GRAM 371 102	PL/MATCH/UNGRAM	455	181
SG/MATCH/GRAM 371 102 SG/MATCH/UNGRAM 387 106 SG/NOMATCH/UNGRAM 387 113 SG/NOMATCH/UNGRAM 399 129 PL/NOMATCH/UNGRAM 397 127 PL/MOMATCH/UNGRAM 393 126 PL/MATCH/UNGRAM 406 129 PL/MATCH/UNGRAM 406 129 PL/MATCH/UNGRAM 400 142 Relative Clause Verb Region SG/MATCH/GRAM 374 115 SG/MATCH/UNGRAM 370 118 SG/NOMATCH/GRAM 366 116 PL/NOMATCH/UNGRAM 389 135 PL/NOMATCH/UNGRAM 390 129 PL/MATCH/UNGRAM 391 135 PL/MATCH/UNGRAM 404 149 Attractor Region SG/MATCH/UNGRAM 400 153 SG/MATCH/UNGRAM 423 178 SG/NOMATCH/GRAM 420 176 SG/NOMATCH/UNGRAM 414 167 PL/NOMATCH/UNGRAM 414 167 PL/MATCH/UNGRAM 450 187		457	217
SG/MATCH/UNGRAM 387 106 SG/NOMATCH/GRAM 387 113 SG/NOMATCH/UNGRAM 399 129 PL/NOMATCH/UNGRAM 397 127 PL/NOMATCH/UNGRAM 393 126 PL/MATCH/UNGRAM 406 129 PL/MATCH/UNGRAM 406 129 PL/MATCH/UNGRAM 400 142 Relative Clause Verb Region SG/MATCH/GRAM 370 118 SG/MATCH/UNGRAM 366 116 PL/NOMATCH/UNGRAM 366 116 PL/NOMATCH/UNGRAM 390 129 PL/NOMATCH/UNGRAM 391 135 PL/NATCH/GRAM 404 149 Attractor Region SG/MATCH/GRAM 400 153 SG/MATCH/UNGRAM 423 178 SG/NOMATCH/UNGRAM 414 167 PL/NOMATCH/UNGRAM 414 167 PL/NOMATCH/UNGRAM 414 161 PL/MATCH/GRAM 450 187 Adverb Region SG/MATCH/UNGRAM	Complementizer I	Region	
SG/NOMATCH/GRAM 387 113 SG/NOMATCH/UNGRAM 399 129 PL/NOMATCH/UNGRAM 397 127 PL/NOMATCH/UNGRAM 393 126 PL/MATCH/UNGRAM 406 129 PL/MATCH/UNGRAM 400 142 Relative Clause Verb Region 142 SG/MATCH/UNGRAM 374 115 SG/MATCH/UNGRAM 366 116 PL/NOMATCH/UNGRAM 366 116 PL/NOMATCH/UNGRAM 390 129 PL/NOMATCH/UNGRAM 391 135 PL/NATCH/UNGRAM 391 135 PL/MATCH/UNGRAM 404 149 Altractor Region SG/MATCH/GRAM 400 153 SG/MATCH/UNGRAM 423 178 SG/NOMATCH/GRAM 423 178 SG/NOMATCH/UNGRAM 420 176 SG/NOMATCH/UNGRAM 414 167 PL/NOMATCH/UNGRAM 414 167 PL/MATCH/UNGRAM 450 187 SG/MATCH/GRAM 450 187	SG/MATCH/GRAM	371	102
SG/NOMATCH/UNGRAM 399 129 PL/NOMATCH/UNGRAM 397 127 PL/NOMATCH/UNGRAM 393 126 PL/MATCH/UNGRAM 406 129 PL/MATCH/GRAM 420 142 Relative Clause Verb Region 142 SG/MATCH/GRAM 374 115 SG/MATCH/UNGRAM 364 117 SG/NOMATCH/UNGRAM 366 116 PL/NOMATCH/UNGRAM 389 135 PL/NOMATCH/UNGRAM 390 129 PL/MATCH/UNGRAM 391 135 PL/NATCH/GRAM 404 149 MATCH/GRAM 404 149 MATCH/GRAM 404 149 MATCH/UNGRAM 423 178 SG/NOMATCH/UNGRAM 420 176 SG/NOMATCH/UNGRAM 414 167 PL/NOMATCH/UNGRAM 419 152 PL/MATCH/GRAM 430 155 PL/MATCH/UNGRAM 450 187 SG/MATCH/UNGRAM 452 186 <td>SG/MATCH/UNGRAM</td> <td>387</td> <td>106</td>	SG/MATCH/UNGRAM	387	106
PL/NOMATCH/UNGRAM 397 127 PL/NOMATCH/UNGRAM 393 126 PL/MATCH/UNGRAM 406 129 PL/MATCH/UNGRAM 420 142 Relative Clause Verb Region SG/MATCH/GRAM 370 118 SG/MATCH/UNGRAM 364 117 SG/NOMATCH/UNGRAM 366 116 PL/NOMATCH/UNGRAM 390 129 PL/NOMATCH/UNGRAM 391 135 PL/MATCH/GRAM 401 149 Attractor Region SG/MATCH/UNGRAM 402 178 SG/NOMATCH/UNGRAM 403 178 SG/NOMATCH/UNGRAM 404 149 Attractor Region SG/MATCH/UNGRAM 420 176 SG/NOMATCH/UNGRAM 410 175 PL/NOMATCH/UNGRAM 410 155 PL/NOMATCH/UNGRAM 450 187 Adverb Region SG/MATCH/UNGRAM 451 175 PL/NOMATCH/UNGRAM 451 175 PL/NOMATCH/UNGRAM 451 175 PL/NOMATCH/UNGRAM 458 177 PL/NOMATCH/UNGRAM 458 177 PL/NOMATCH/UNGRAM 458 177 PL/NOMATCH/UNGRAM 458 177 PL/MATCH/UNGRAM 458 177 PL/MATCH/UNGRAM 458 177 PL/MATCH/UNGRAM 458 177 PL/MATCH/UNGRAM 361 96 SG/NOMATCH/UNGRAM 376 95 SG/NOMATCH/UNGRAM 376 95 SG/NOMATCH/UNGRAM 370 99 PL/NOMATCH/UNGRAM 363 92	SG/NoMatch/Gram	387	113
PL/NOMATCH/GRAM 393 126 PL/MATCH/UNGRAM 406 129 PL/MATCH/UNGRAM 420 142 Relative Clause Verb Region SG/MATCH/GRAM 370 118 SG/NOMATCH/UNGRAM 364 117 SG/NOMATCH/UNGRAM 366 116 PL/NOMATCH/UNGRAM 390 129 PL/MATCH/UNGRAM 391 135 PL/NOMATCH/GRAM 391 135 PL/MATCH/UNGRAM 391 135 PL/MATCH/UNGRAM 404 149 Attractor Region SG/MATCH/UNGRAM 423 178 SG/NOMATCH/UNGRAM 423 178 SG/NOMATCH/UNGRAM 414 167 PL/NOMATCH/UNGRAM 414 167 PL/NOMATCH/UNGRAM 414 161 PL/MATCH/UNGRAM 450 187 Adverb Region SG/MATCH/UNGRAM 450 187 Adverb Region SG/NOMATCH/UNGRAM 450 187 Adverb Region SG/NOMATCH/UNGRAM 450 187 Adverb Region SG/NOMATCH/UNGRAM 442 153 PL/NOMATCH/UNGRAM 445 186 SG/NOMATCH/UNGRAM 446 188 SG/NOMATCH/UNGRAM 447 185 PL/NOMATCH/UNGRAM 448 150 SG/NOMATCH/UNGRAM 449 185 PL/NOMATCH/UNGRAM 449 185 PL/NOMATCH/UNGRAM 451 175 PL/MATCH/UNGRAM 454 165 Second Spillover Region SG/MATCH/UNGRAM 362 94 SG/MATCH/UNGRAM 362 94 SG/NOMATCH/UNGRAM 360 95 SG/NOMATCH/UNGRAM 370 99 PL/NOMATCH/UNGRAM 370 99 PL/NOMATCH/UNGRAM 370 99 PL/NOMATCH/UNGRAM 370 99 PL/NOMATCH/UNGRAM 363 92 375	SG/NoMatch/Ungram	399	129
PL/MATCH/UNGRAM 406 129 PL/MATCH/GRAM 420 142 Relative Clause Verb Region SG/MATCH/GRAM 374 115 SG/MATCH/UNGRAM 370 118 SG/NOMATCH/GRAM 366 116 PL/NOMATCH/UNGRAM 389 135 PL/NOMATCH/UNGRAM 390 129 PL/NOMATCH/UNGRAM 391 135 PL/MATCH/UNGRAM 404 149 Attractor Region SG/MATCH/UNGRAM 423 178 SG/NOMATCH/UNGRAM 423 178 SG/NOMATCH/UNGRAM 420 176 SG/NOMATCH/UNGRAM 414 167 PL/NOMATCH/UNGRAM 414 167 PL/NOMATCH/UNGRAM 414 167 PL/NOMATCH/UNGRAM 430 155 PL/MATCH/GRAM 450 187 Adverb Region SG/MATCH/UNGRAM 450 187 Adverb Region SG/MATCH/UNGRAM 450 187 Adverb Region SG/NOMATCH/UNGRAM 451 175 PL/NOMATCH/UNGRAM 442 153 PL/NOMATCH/UNGRAM 445 186 SG/NOMATCH/UNGRAM 445 186 SG/NOMATCH/UNGRAM 445 186 SG/NOMATCH/UNGRAM 445 186 SG/NOMATCH/UNGRAM 445 175 PL/MATCH/UNGRAM 451 175 PL/MATCH/UNGRAM 454 165 Second Spillover Region SG/MATCH/UNGRAM 362 94 SG/NOMATCH/UNGRAM 376 95 SG/NOMATCH/UNGRAM 376 95 SG/NOMATCH/UNGRAM 376 95 SG/NOMATCH/UNGRAM 376 95 SG/NOMATCH/UNGRAM 370 99 PL/NOMATCH/UNGRAM 370 99 PL/NOMATCH/UNGRAM 370 99 PL/NOMATCH/UNGRAM 363 92	PL/NoMatch/Ungram	397	127
PL/MATCH/GRAM	PL/NOMATCH/GRAM	393	126
Relative Clause Verb Region SG/MATCH/GRAM 374 115 SG/MATCH/UNGRAM 370 118 SG/NOMATCH/UNGRAM 364 117 SG/NOMATCH/UNGRAM 366 116 PL/NOMATCH/UNGRAM 389 135 PL/NOMATCH/GRAM 390 129 PL/MATCH/UNGRAM 391 135 PL/MATCH/GRAM 404 149 Attractor Region SG/MATCH/GRAM 404 149 SG/MATCH/GRAM 423 178 SG/NOMATCH/UNGRAM 420 176 SG/NOMATCH/UNGRAM 419 152 PL/NOMATCH/UNGRAM 419 152 PL/NOMATCH/GRAM 430 155 PL/MATCH/UNGRAM 450 187 Adverb Region SG/MATCH/UNGRAM 450 188 SG/MATCH/UNGRAM 450 188 SG/MATCH/UNGRAM 456 188 SG/MOMATCH/UNGRAM 451 175 PL/NOMATCH/UNGRAM 451 175 PL/MATCH/UNGRAM 458 177 PL/NOMATCH/UNG	PL/MATCH/UNGRAM	406	129
SG/MATCH/GRAM 374 115 SG/MATCH/UNGRAM 370 118 SG/NOMATCH/UNGRAM 364 117 SG/NOMATCH/UNGRAM 366 116 PL/NOMATCH/UNGRAM 389 135 PL/NOMATCH/GRAM 390 129 PL/MATCH/UNGRAM 391 135 PL/MATCH/GRAM 404 149 Attractor Region SG/MATCH/GRAM 400 153 SG/MATCH/UNGRAM 420 176 SG/NOMATCH/UNGRAM 420 176 SG/NOMATCH/UNGRAM 419 152 PL/NOMATCH/UNGRAM 419 152 PL/NOMATCH/UNGRAM 450 187 Adverb Region SG/MATCH/GRAM 450 187 SG/MATCH/UNGRAM 450 188 SG/NOMATCH/UNGRAM 442 153 PL/NOMATCH/UNGRAM 449 185 PL/NOMATCH/UNGRAM 451 175 PL/MATCH/GRAM 451 175 PL/MATCH/GRAM 451 175 <td>PL/MATCH/GRAM</td> <td>420</td> <td>142</td>	PL/MATCH/GRAM	420	142
SG/MATCH/UNGRAM 370 118 SG/NOMATCH/UNGRAM 364 117 SG/NOMATCH/UNGRAM 366 116 PL/NOMATCH/UNGRAM 389 135 PL/NOMATCH/UNGRAM 390 129 PL/MATCH/UNGRAM 391 135 PL/MATCH/UNGRAM 404 149 Attractor Region Autractor Region 153 SG/MATCH/GRAM 420 176 SG/NOMATCH/UNGRAM 419 152 SG/NOMATCH/UNGRAM 419 152 PL/NOMATCH/UNGRAM 430 155 PL/MATCH/UNGRAM 450 187 Adverb Region Adverb Region SG/MATCH/UNGRAM 452 186 SG/NOMATCH/UNGRAM 456 188 SG/NOMATCH/UNGRAM 441 161 PL/NOMATCH/UNGRAM 451 175 PL/NOMATCH/UNGRAM 451 175 PL/NOMATCH/UNGRAM 458 177 PL/MATCH/UNGRAM 458 177 PL/NOMATCH/UNGRAM	Relative Clause Veri	b Region	
SG/NOMATCH/GRAM 364 117 SG/NOMATCH/UNGRAM 366 116 PL/NOMATCH/UNGRAM 389 135 PL/NOMATCH/GRAM 390 129 PL/MATCH/UNGRAM 391 135 PL/MATCH/GRAM 404 149 Attractor Region 133 149 SG/MATCH/GRAM 400 153 SG/MATCH/UNGRAM 420 176 SG/NOMATCH/UNGRAM 420 176 SG/NOMATCH/UNGRAM 419 152 PL/NOMATCH/UNGRAM 419 152 PL/NOMATCH/GRAM 450 187 Adverb Region 187 428 150 SG/MATCH/GRAM 450 187 Adverb Region 188 SG/NOMATCH/UNGRAM 452 186 SG/NOMATCH/UNGRAM 442 153 PL/NOMATCH/UNGRAM 442 153 PL/NOMATCH/UNGRAM 451 175 PL/MATCH/UNGRAM 451 175 PL/MATCH/GRAM 454 165 Second Spillover Region<	SG/MATCH/GRAM	374	115
SG/NOMATCH/UNGRAM 366 116 PL/NOMATCH/UNGRAM 389 135 PL/NOMATCH/UNGRAM 390 129 PL/NATCH/UNGRAM 391 135 PL/MATCH/GRAM 404 149 Attractor Region SG/MATCH/GRAM 400 153 SG/MATCH/UNGRAM 423 178 SG/NOMATCH/UNGRAM 414 167 PL/NOMATCH/UNGRAM 414 161 PL/NOMATCH/UNGRAM 419 152 PL/NOMATCH/UNGRAM 450 187 Adverb Region Adverb Region SG/MATCH/GRAM 452 186 SG/NOMATCH/UNGRAM 452 186 SG/NOMATCH/UNGRAM 449 185 PL/NOMATCH/UNGRAM 451 175 PL/NOMATCH/UNGRAM 451 175 PL/MATCH/GRAM 454 165 SG/MATCH/GRAM 362 94 SG/MATCH/UNGRAM 376 95 SG/MATCH/UNGRAM 361 96	SG/MATCH/UNGRAM	370	118
PL/NoMatch/Ungram 389 135 PL/NoMatch/Ungram 390 129 PL/Match/Ungram 391 135 PL/Match/Ungram 404 419 Altractor Region SG/Match/Gram 400 153 SG/Match/Ungram 423 178 SG/NoMatch/Gram 423 178 SG/NoMatch/Ungram 414 167 PL/NoMatch/Ungram 414 167 PL/NoMatch/Ungram 414 161 PL/Match/Ungram 430 155 PL/Match/Ungram 450 187 Adverb Region SG/Match/Ungram 452 186 SG/Match/Ungram 452 186 SG/NoMatch/Ungram 452 186 SG/NoMatch/Ungram 442 153 SG/NoMatch/Ungram 442 153 PL/NoMatch/Ungram 442 153 PL/NoMatch/Ungram 442 153 PL/NoMatch/Ungram 451 175 PL/Match/Ungram 454 165 Second Spillover Region SG/Match/Ungram 362 94 SG/Match/Ungram 376 95 SG/NoMatch/Ungram 376 95 SG/NoMatch/Ungram 376 95 SG/NoMatch/Ungram 370 99 PL/NoMatch/Ungram 370 99 PL/NoMatch/Gram 363 92	SG/NoMatch/Gram	364	117
PL/NoMatch/Gram 390 129 PL/Match/Ungram 391 135 PL/Match/Gram 404 419 Attractor Region SG/Match/Gram 423 178 SG/NoMatch/Ungram 420 176 SG/NoMatch/Ungram 414 167 PL/NoMatch/Ungram 419 152 PL/NoMatch/Ungram 410 157 PL/Match/Ungram 430 155 PL/Match/Ungram 450 187 Adverb Region SG/Match/Ungram 452 186 SG/Match/Ungram 452 186 SG/Match/Ungram 452 186 SG/NoMatch/Ungram 456 188 SG/NoMatch/Ungram 442 153 PL/NoMatch/Ungram 442 153 PL/NoMatch/Ungram 442 153 PL/NoMatch/Ungram 451 175 PL/Match/Ungram 458 177 PL/Match/Ungram 362 94 SG/Match/Ungram 376 95 SG/NoMatch/Ungram 376 95 SG/NoMatch/Ungram 376 95 SG/NoMatch/Ungram 370 99 PL/NoMatch/Gram 363 92	SG/NoMatch/Ungram	366	116
PL/MATCH/UNGRAM 391 135 PL/MATCH/GRAM 404 404 Altractor Region SG/MATCH/UNGRAM 420 176 SG/MATCH/UNGRAM 420 176 SG/MATCH/UNGRAM 420 176 SG/NOMATCH/UNGRAM 419 152 PL/NOMATCH/UNGRAM 414 161 PL/MATCH/UNGRAM 430 155 PL/MATCH/GRAM 450 187 Adverb Region SG/MATCH/UNGRAM 450 187 Adverb Region SG/MATCH/UNGRAM 456 188 SG/NOMATCH/UNGRAM 456 188 SG/NOMATCH/UNGRAM 442 153 PL/NOMATCH/UNGRAM 442 153 PL/NOMATCH/UNGRAM 441 175 PL/MATCH/UNGRAM 451 175 PL/MATCH/UNGRAM 451 175 PL/MATCH/UNGRAM 451 175 PL/MATCH/UNGRAM 454 165 Second Spillover Region SG/MATCH/UNGRAM 361 96 SG/NOMATCH/UNGRAM 376 95 SG/NOMATCH/UNGRAM 370 99 PL/NOMATCH/UNGRAM 363 92	PL/NoMatch/Ungram	389	135
PL/MATCH/GRAM	PL/NoMatch/Gram	390	129
SG/MATCH/GRAM 400 153 SG/MATCH/UNGRAM 423 178 SG/NOMATCH/GRAM 420 176 SG/NOMATCH/UNGRAM 414 167 PL/NOMATCH/UNGRAM 419 152 PL/NOMATCH/GRAM 430 155 PL/MATCH/GRAM 450 187 Adverb Region SG/MATCH/GRAM 452 186 SG/MATCH/UNGRAM 452 186 SG/NOMATCH/UNGRAM 452 186 SG/NOMATCH/GRAM 456 188 SG/NOMATCH/UNGRAM 445 175 PL/NOMATCH/UNGRAM 441 175 PL/NOMATCH/UNGRAM 451 175 PL/NOMATCH/UNGRAM 454 165 Second Spillover Region SG/MATCH/GRAM 361 362 364 SG/NOMATCH/UNGRAM 376 95 SG/NOMATCH/UNGRAM 370 99 PL/NOMATCH/UNGRAM 370 99 PL/NOMATCH/GRAM 363 92	PL/MATCH/UNGRAM	391	135
SG/MATCH/GRAM 400 153	PL/MATCH/GRAM	404	149
SG/MATCH/UNGRAM 423 178 SG/NOMATCH/GRAM 420 176 SG/NOMATCH/UNGRAM 414 167 PL/NOMATCH/UNGRAM 414 161 PL/MATCH/UNGRAM 430 155 PL/MATCH/GRAM 450 187 Adverb Region SG/MATCH/GRAM 452 186 SG/MATCH/UNGRAM 452 186 SG/NOMATCH/UNGRAM 456 188 SG/NOMATCH/UNGRAM 445 185 SG/NOMATCH/UNGRAM 445 153 PL/NOMATCH/UNGRAM 451 175 PL/MATCH/UNGRAM 454 165 Second Spillover Region SG/MATCH/UNGRAM 454 165 Second Spillover Region SG/MATCH/UNGRAM 376 95 SG/NOMATCH/UNGRAM 376 96 SG/NOMATCH/UNGRAM 370 99 PL/NOMATCH/UNGRAM 370 99 PL/NOMATCH/GRAM 363 92	Attractor Regi	ion	
SG/NoMatch/Gram 420 176 SG/NoMatch/Ungram 414 167 PL/NoMatch/Ungram 419 152 PL/NoMatch/Ungram 430 155 PL/Match/Ungram 430 155 PL/Match/Gram 450 187 Adverb Region SG/Match/Ungram 452 186 SG/Match/Ungram 452 186 SG/NoMatch/Gram 456 188 SG/NoMatch/Ungram 449 185 PL/NoMatch/Ungram 441 175 PL/NoMatch/Ungram 451 175 PL/Match/Ungram 458 177 PL/Match/Ungram 458 177 PL/Match/Ungram 454 165 Second Spillover Region SG/Match/Ungram 362 94 SG/Match/Ungram 376 95 SG/NoMatch/Ungram 376 95 SG/NoMatch/Ungram 361 96 SG/NoMatch/Ungram 370 99 PL/NoMatch/Gram 370 99 PL/NoMatch/Gram 363 92	SG/MATCH/GRAM	400	153
SG/NoMatch/Ungram	SG/MATCH/UNGRAM		178
PL/NoMatch/Ungram 419 152			
PL/NoMatch/Gram			
PL/MATCH/UNGRAM			
PL/MATCH/GRAM			
Adverb Region SG/MATCH/GRAM 428 150 SG/MATCH/UNGRAM 452 186 SG/NOMATCH/GRAM 456 188 SG/NOMATCH/UNGRAM 449 185 PL/NOMATCH/UNGRAM 442 153 PL/MOMATCH/UNGRAM 451 175 PL/MATCH/UNGRAM 454 165 Second Spillover Region SG/MATCH/UNGRAM 362 94 SG/MATCH/UNGRAM 376 95 SG/NOMATCH/UNGRAM 361 96 SG/NOMATCH/UNGRAM 380 101 PL/NOMATCH/UNGRAM 370 99 PL/NOMATCH/GRAM 363 92			
SG/MATCH/GRAM 428 150			187
SG/MATCH/UNGRAM 452 186 SG/NOMATCH/GRAM 449 185 PL/NOMATCH/UNGRAM 442 153 PL/NOMATCH/UNGRAM 441 175 PL/MATCH/UNGRAM 458 177 PL/MATCH/GRAM 454 165 Second Spillover Region SG/MATCH/UNGRAM 362 94 SG/MATCH/UNGRAM 361 96 SG/NOMATCH/UNGRAM 361 96 SG/NOMATCH/UNGRAM 370 99 PL/NOMATCH/GRAM 363 92	Adverb Regio	on	
SG/NOMATCH/GRAM 456 188 SG/NOMATCH/UNGRAM 449 185 PL/NOMATCH/UNGRAM 442 153 PL/NOMATCH/GRAM 451 175 PL/MATCH/UNGRAM 458 177 PL/MATCH/GRAM 454 165 Second Spillover Region 56 SG/MATCH/GRAM 362 94 SG/MATCH/UNGRAM 376 95 SG/NOMATCH/UNGRAM 361 96 SG/NOMATCH/UNGRAM 380 101 PL/NOMATCH/UNGRAM 370 99 PL/NOMATCH/GRAM 363 92			
SG/NOMATCH/UNGRAM			
PL/NOMATCH/UNGRAM			
PL/NoMatch/Gram 451 175 PL/Match/Ungram 458 177 PL/Match/Gram 454 165 Second Spillover Region SG/Match/Gram 362 94 SG/Match/Ungram 376 95 SG/NoMatch/Ungram 380 101 PL/NoMatch/Ungram 370 99 PL/NoMatch/Gram 363 92			
PL/MATCH/UNGRAM			
PL/MATCH/GRAM			
Second Spillover Region SG/MATCH/GRAM 362 94 SG/MATCH/UNGRAM 376 95 SG/NOMATCH/GRAM 361 96 SG/NOMATCH/UNGRAM 380 101 PL/NOMATCH/UNGRAM 370 99 PL/NOMATCH/GRAM 363 92			
SG/MATCH/GRAM 362 94 SG/MATCH/UNGRAM 376 95 SG/NOMATCH/GRAM 361 96 SG/NOMATCH/UNGRAM 380 101 PL/NOMATCH/UNGRAM 370 99 PL/NOMATCH/GRAM 363 92			165
SG/MATCH/UNGRAM 376 95 SG/NOMATCH/GRAM 361 96 SG/NOMATCH/UNGRAM 380 101 PL/NOMATCH/UNGRAM 370 99 PL/NOMATCH/GRAM 363 92		Region	
SG/NOMATCH/GRAM 361 96 SG/NOMATCH/UNGRAM 380 101 PL/NOMATCH/UNGRAM 370 99 PL/NOMATCH/GRAM 363 92			
SG/NOMATCH/UNGRAM 380 101 PL/NOMATCH/UNGRAM 370 99 PL/NOMATCH/GRAM 363 92			
PL/NOMATCH/UNGRAM 370 99 PL/NOMATCH/GRAM 363 92			
PL/NoMatch/Gram 363 92			
	PL/MATCH/UNGRAM	367	96
PL/MATCH/GRAM 373 98	PL/MATCH/GRAM	373	98

Table D.7: Raw condition grand avergage reading times across participant means in milliseconds for additional regions in Experiment 4.

Coefficient	\hat{eta}	SE	t
Relative Clause Verb Region	ı		
Intercept	379.16	17.33	21.88
SubjNum:PL	24.33	8.20	2.97
Match:No	-9.12	8.17	-1.12
Grammaticality:UNGRAM	-6.22	8.19	-0.76
Item Order	-0.61	0.04	-16.76
Length	4.81	2.58	1.86
Previous Region RT	0.08	0.01	12.70
SubjNum:PL × Match:No	2.33	11.59	0.20
SubjNum:PL × Grammaticality:UNGRAM	-0.58	11.60	-0.05
Match:No × Grammaticality:UNGRAM	5.80	11.56	0.50
SubjNum:PL × Match:No × Grammaticality:UNGRAM	-2.94	16.40	-0.18
Attractor Region			
Intercept	333.03	33.49	9.95
SubjNum:PL	32.43	11.31	2.87
Match:No	13.42	11.28	1.19
Grammaticality:UNGRAM	20.01	10.54	1.90
Item Order	-0.85	0.05	-17.83
Length	9.47	4.05	2.34
Previous Region RT	0.21	0.01	17.49
SubjNum:PL × Match:No	-35.49	16.98	-2.09
SubjNum:PL × $Grammaticality$:Ungram	-28.17	14.94	-1.89
Match:No × Grammaticality:UNGRAM	-26.99	14.88	-1.81
SubjNum:PL × Match:No × Grammaticality:UNGRAM	40.35	21.12	1.91
Adverb Region			
Intercept	489.73	28.15	17.40
SubjNum:PL	20.51	11.07	1.85
Match:No	29.06	11.02	2.64
Grammaticality:UNGRAM	18.31	11.05	1.66
Item Order	-0.81	0.05	-16.14
Length	-5.08	3.99	-1.27
Previous Region RT	0.11	0.01	10.69
SubjNum:PL × Match:No	-24.19	15.64	-1.55
SubjNum:PL × Grammaticality:UNGRAM	-4.18	15.66	-0.27
Match:No × Grammaticality:UNGRAM	-24.07	15.59	-1.54
SubjNum:PL × Match:No × Grammaticality:UNGRAM	-0.18	22.13	-0.01
Second Spillover Region			
Second Spillover Region	303.85	11.14	27 27
Intercept	303.85 6 46	11.14 5.52	27.27 1 17
Intercept SubjNum:PL	6.46	5.52	1.17
Intercept SubjNum:PL Match:No	6.46 -0.32	5.52 5.50	1.17 -0.06
Intercept SubjNum:PL Match:NO Grammaticality:UNGRAM	6.46 -0.32 4.41	5.52 5.50 5.53	1.17 -0.06 0.80
Intercept SubjNum:PL Match:NO Grammaticality:UNGRAM Item Order	6.46 -0.32 4.41 -0.44	5.52 5.50 5.53 0.03	1.17 -0.06 0.80 -17.48
Intercept SubjNum:PL Match:NO Grammaticality:UNGRAM Item Order Length	6.46 -0.32 4.41 -0.44 5.11	5.52 5.50 5.53 0.03 1.42	1.17 -0.06 0.80 -17.48 3.61
Intercept SubjNum:PL Match:NO Grammaticality:UNGRAM Item Order Length Previous Region RT	6.46 -0.32 4.41 -0.44 5.11 0.20	5.52 5.50 5.53 0.03 1.42 0.01	1.17 -0.06 0.80 -17.48 3.61 23.49
Intercept SubjNum:PL Match:NO Grammaticality:UNGRAM Item Order Length Previous Region RT SubjNum:PL × Match:NO	6.46 -0.32 4.41 - 0.44 5.11 0.20 -6.26	5.52 5.50 5.53 0.03 1.42 0.01 7.80	1.17 -0.06 0.80 -17.48 3.61 23.49 -0.80
Intercept SubjNum:PL Match:NO Grammaticality:UNGRAM Item Order Length Previous Region RT	6.46 -0.32 4.41 -0.44 5.11 0.20	5.52 5.50 5.53 0.03 1.42 0.01	1.17 -0.06 0.80 -17.48 3.61 23.49

Table D.8: Mixed effects regression coefficients for other regions in Experiment 4. Significant coefficients (|t| > 2) are in bold and marginal coefficients (|t| > 1.65) are in italics.

847 References

- Antón-Méndez, I., Nicol, J. L., & Garrett, M. F. (2002). The relation between gender and number agreement processing. *Syntax*, *5*, 1–25.
- Badecker, W., & Kuminiak, F. (2007). Morphology, agreement and working memory
 retrieval in sentence production: Evidence from gender and case in Slovak. *Journal* of Memory and Language, 56, 65–85.
- Badecker, W., & Lewis, R. (2007). A new theory and computational model of working memory in sentence production: Agreement errors as failues of cue-based retrieval. Paper Presented at the 20th annual CUNY Sentence Processing Conference.

 San Diego, La Jolla.
- Baker, M. C. (2008). *The Syntax of Agreement and Concord*. Cambridge, UK: Cambridge University Press.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255–278.
- Bates, D. M., Kliegl, R., Vasishth, S., & Baayen, R. H. (2015). Parsimonious mixed models. URL: http://arxiv.org/abs/1506.04967 ms., arXiv.
- Béjar, S. (2003). *Phi-Syntax: A Theory of Agreement*. Ph.D. thesis University of Toronto Toronto.
- Béjar, S., & Rezac, M. (2009). Cyclic Agree. Linguistic Inquiry, 40, 35-73.
- Bock, K., & Eberhard, K. M. (1993). Meaning, sound and syntax in English number agreement. *Language and Cognitive Processes*, *8*, 57–99.
- Bock, K., & Miller, C. (1991). Broken agreement. Cognitive Psychology, 23, 45–93.
- ¹⁸⁷⁰ Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, N. (2000). Minimalist inquiries. In D. M. Martin, Roger, & J. Uriagereka (Eds.), *Step by Step: Essays on Minimalist Sytnax in Honor of Howard Lasnik* chapter 3. (pp. 89–156). Cambridge, MA: MIT Press.
- Chomsky, N. (2001). Derivation by phase. In M. Kenstowicz (Ed.), Ken Hale: A Life
 in Language (pp. 1–52). Cambridge, Massachusetts: MIT Press.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Earlbaum Associates.
- Dank, M., & Deutsch, A. (2010). The role of morpho-phonological factors in subject—predicate gender agreement in Hebrew. *Language and Cognitive Processes*, 25, 1380–1410.
- Deutsch, A., & Dank, M. (2009). Conflicting cues and competition between notional and grammatical factors in producing number and gender agreement: Evidence from Hebrew. *Journal of Memory and Language*, *60*, 112–143.
- Deutsch, A., & Dank, M. (2011). Symmetric and asymmetric patterns of attraction errors in producing subject–predicate agreement in Hebrew: An issue of morphological structure. *Language and Cognitive Processes*, 26, 24–46.
- Dillon, B., Mishler, A., Sloggett, S., & Phillips, C. (2013). Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language*, *69*, 85–103.

- Eberhard, K., Cutting, J., & Bock, K. (2005). Making syntax of sense: Number agreement in sentence production. *Psychological Review*, *112*, 531–559.
- Eberhard, K. M. (1997). The marked effect of number on subject–verb agreement. *Journal of Memory and Language*, *36*, 147–164.
- Engelmann, F., Jäger, L., & Vasishth, S. (2015a). Cue confusion and distractor prominence can explain inconsistent interference effects. Poster Presented at the 28th Annual CUNY Conference on Human Sentence Processing, Los Angeles, CA.
- Engelmann, F., Jäger, L. A., & Vasishth, S. (2015b). The determinants of retrieval interference in dependency resolution: Review and computational modeling. Ms. (submitted), Universität Potsdam.
- Franck, J. (2011). Reaching agreement as a core syntactic process. *Natural Language*4. *Linguistic Theory*, 29, 1071–1086.
- Franck, J., Colonna, S., & Rizzi, L. (2015). Task-dependency and structure-dependency in number interference effects in sentence comprehension. *Frontiers in Psychology*, 6, 349.
- Franck, J., Vigliocco, G., Antón-Méndez, I., Collina, S., & Frauenfelder, U. (2008).
 The interplay of syntax and form in sentence production: A cross-linguistic study of form effects on agreement. *Language and Cognitive Processes*, 23, 329–374.
- Franck, J., Vigliocco, G., & Nicol, J. (2002). Subject-verb agreement errors in French and English: The role of syntactic hierarchy. *Language and Cognitive Processes*, 17, 371–404.
- Fuchs, Z., Polinsky, M., & Scontras, G. (2015). The differential representation of number and gender in Spanish. *The Linguistic Review*, *32*, 703–737.
- Gelman, A., & Hill, J. (2006). *Data Analysis Using Regression and Multi-level/Hierarchical Models*. Cambridge, UK: Cambridge University Press.
- Gillespie, M., & Pearlmutter, N. J. (2013). Against structural constraints in subject–verb agreement production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39, 515.
- Hartsuiker, R. J., Schriefers, H., Bock, K., & Kikstra, G. (2003). Morphophonological influences on the construction of subject-verb agreement. *Memory & Cognition*, *31*, 1316.
- Häussler, J. (2009). *The Emergence of Agreement Attraction Errors During Sentence*Comprehension. Ph.D. thesis Universität Konstanz Konstanz, Germany.
- Häussler, J., & Bader, M. (2009). Agreement checking and number attraction in sentence comprehension: Insights from German relative clauses. *Travaux du Cercle Linguistique de Prague*, 7.
- Just, M., Carpenter, P., & Woolley, J. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, 111, 228–238.
- Lago, S., Shalom, D. E., Sigman, M., Lau, E. F., & Phillips, C. (2015). Agreement in spanish comprehension. *Journal of Memory and Language*, 82, 133–49.
- Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, *29*, 375–419.
- Lorimor, H., Bock, K., Zalkind, E., Sheyman, A., & Beard, R. (2008). Agreement and attraction in Russian. *Language and Cognitive Processes*, *23*, 769–799.
- Malko, A., & Slioussar, N. (To Appear). Gender features in processing: Evidence from Russian. Frontiers in Psychology: Encoding and Navigating Linguistic Representa-

- tions in Memory, 6.
- Nicol, J. L., Forster, K. I., & Veres, C. (1997). Subject–verb agreement processes in comprehension. *Journal of Memory and Language*, *36*, 569–87.
- Norris, M. (2014). *A Theory of Norminal Concord*. Ph.D. thesis University of California, Santa Cruz.
- Patson, N. D., & Husband, E. M. (2015). Misinterpretations in agreement and agreement attraction. *The Quarterly Journal of Experimental Psychology*, (pp. 1–22).
- Pearlmutter, N. J. (2000). Linear versus hierarchical agreement feature processing in comprehension. *Journal of Psycholinguistic Research*, *29*, 89–98.
- Pearlmutter, N. J., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language*, *41*, 427–456.
- Pollock, J.-Y. (1989). Verb movement, universal grammar, and the structure of IP. Linguistic Inquiry, 20, 365–424.
- Preminger, O. (2011). Agreement as a Fallible Operation. Ph.D. thesis MIT Cambridge,
 MA.
- ¹⁹⁵¹ R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing Vienna, Austria.
- Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*, 510–532.
- Rhode, D. (2003). *Linger: A Flexible Platform for Language Processing Experiments*.

 MIT.
- Ryding, K. C. (2005). *A Reference Grammar of Modern Standard Arabic*. Cambridge, UK: Cambridge UP.
- Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics Bulletin*, *2*, 110–114.
- Sicuro Corrêa, L. M., Almeida, D. A. d. A., & Porto, R. S. (2004). On the representation
 of Portuguese gender-inflected words in the mental lexicon. *Brain and Language*,
 90, 63–73.
- Staub, A. (2009). On the interpretation of the number attraction effect: Response time evidence. *Journal of Memory and Language*, *60*, 308–27.
- Staub, A. (2010). Response time distributional evidence for distinct varieties of number attraction. *Cognition*, *114*, 447–54.
- Tanner, D. (2011). Agreement Mechanisms in Native and Nonnative Language Processing: Electrophysiological Correlates of Complexity and Interference. Ph.D. thesis University of Washington Seattle, WA.
- Tanner, D., Nicol, J. L., & Brehm, L. (2014). The time-course of feature interference in agreement comprehension: Multiple mechanisms and asymmetrical attraction. *Journal of Memory and Language*, 76, 195–215.
- Trubetskoy, N. (1939/1958). *Grundzüge der Phonologie*. Göttingen: Vandenhoeck & Ruprecht.
- Tucker, M. A. (2011). The morphosyntax of the Arabic verb. In N. LaCara, A. Thompson, & M. A. Tucker (Eds.), *Morphology at Santa Cruz: Papers in Honor of Jorge Hankamer* (pp. 177–211). Santa Cruz, CA: LRC Publications.
- Tucker, M. A., Idrissi, A., & Almeida, D. (2015). Representing number in the realtime processing of agreement: Self-paced reading evidence from Arabic. *Frontiers* in *Psychology*, 6.

- Tucker, M. A., & Wagers, M. W. (2010). Agreement attraction in infinitival complements. In G. McGuire (Ed.), *UCSC LRC Laboratory Report 2010* (pp. 23–43). Santa Cruz, CA: LRC Publications.
- Vigliocco, G., Butterworth, B., & Semenza, C. (1995). Constructing subject-verb agreement in speech: The role of semantic and morphological factors. *Journal of Memory* and Language, 34, 186–215.
- Vigliocco, G., & Franck, J. (1999). When sex and syntax go hand in hand: Gender agreement in language production. *Journal of Memory and Language*, 40, 455–478.
- Vigliocco, G., & Franck, J. (2001). When sex affects syntax: Contextual influences in sentence production. *Journal of Memory and Language*, *45*, 368–390.
- Vigliocco, G., & Nicol, J. (1998). Separating hierarchical relations and word order in language production: Is proximity concord syntactic or linear? *Cognition*, 68, B13–B29.
- Vigliocco, G., & Zilli, T. (1999). Syntactic accuracy in sentence production: The case of
 gender disagreement in Italian language-impaired and unimpaired speakers. *Journal* of Psycholinguistic Research, 28, 623–648.
- Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, 61, 2000
 206–237.
- Wehr, H. J. C. e. (1976). *A Dictionary of Modern Written Arabic*. Spoken Language Services Ithaca, NY.
- Welch, B. L. (1947). The generalization of "student's" problem when several different population variances are involved. *Biometrika*, *34*, 28–35.