* Manuscript
Click here to view linked References

Agreement attraction in comprehension: representations and processes*

Matthew W. Wagers[1], Ellen F. Lau[2] & Colin Phillips[2,3]
[1]Department of Linguistics, University of California, Santa Cruz; [2]Department of Linguistics, University of Maryland, College Park; [3]Neuroscience and Cognitive Science Program, University of Maryland, College Park

*The first two authors contributed equally to this work. Correspondence should be addressed to Matthew Wagers: *mwagers@ucsc.edu*

Abstract

Much work has demonstrated so-called attraction errors in the production of subject-verb agreement (e.g., 'The key to the cabinets are on the table', [Bock, J. K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology, 23*, 45-93]), in which a verb erroneously agrees with an intervening noun. Six self-paced reading experiments examined the online mechanisms underlying the analogous attraction effects that have been shown in comprehension; namely reduced disruption for subject-verb agreement violations when these 'attractor' nouns intervene. One class of theories suggests that these effects are rooted in faulty representation of the number of the subject, while another class of theories suggests instead that such effects arise in the process of re-accessing subject number at the verb. Two main findings provide evidence against the first class of theories. First, attraction also occurs in relative clause configurations in which the attractor noun does not intervene between subject and verb and is not in a direct structural relationship with the subject head (e.g., 'The drivers who the runner wave to each morning'). Second, we observe a 'grammatical asymmetry': attraction effects are limited to ungrammatical sentences, which would be unexpected if the representation of subject number were inherently prone to error. We argue that agreement attraction in comprehension instead reflects a cue-based retrieval mechanism that is subject to retrieval errors. The grammatical asymmetry can be accounted for under one implementation that we propose, or if the mechanism is only called upon when the predicted agreement features fail to be instantiated on the verb.

***Keywords*:** agreement, comprehension, syntax, retrieval, prediction

Introduction

The trajectory of language processing reflects sophisticated grammatical knowledge that takes advantage of complex structural cues. However, some of our best information about the processing architecture comes when its seams are exposed by characteristic 'mistakes'. The formulation of agreement in language production is one domain famously prone to such a class of mistakes, commonly known as agreement attraction. In agreement attraction errors, an agreement-bearing element, such as a verb, fails to match the agreement features of its grammatical controller and instead realizes agreement with a nearby but grammatically inaccessible distractor. In the example in (1) the plural number feature of the verb *make* mismatches with the singular number of the subject noun phrase 'the sheer weight of all these figures' and instead matches an NP within the subject-internal prepositional phrase.

(1) [$_{NP}$ The sheer **weight$_{SG}$** [$_{PP}$ of all these figures$_{PL}$]] make$_{PL}$ them harder to understand.
(based on Ronald Reagan, 13 October 1982; quoted in Francis, 1986).

Examples like these have long been noticed by grammarians and syntactic theorists (den Dikken, 2001; Francis, 1986; Jespersen, 1924; Kimball & Aissen, 1971; Quirk, Greenbaum, Leech & Svartvik, 1985; inter alia), and can be observed both in natural speech and well-edited texts like the *New York Times*. Production studies have established a well-attested and robust body of factors that affect the likelihood with which structurally-defined agreement fails and attraction is induced by a nearby element. These factors include the number features on potential attractors, their relative structural depth with respect to the grammatical controller, and linear order. (Bock & Miller, 1991; Bock & Cutting, 1992; Bock & Eberhard, 1993; Vigliocco & Nicol,1998; Hartsuiker, Antón-Méndez, & Van Zee, 2001; Haskell & MacDonald, 2003; Thornton & MacDonald, 2003).

Accounts of agreement attraction have typically appealed to the notion that multiple nouns (or noun phrases) in a complex subject have independent specifications for number (Eberhard, 1997, Eberhard, Cutting & Bock, 2005; Franck, Vigliocco & Nicol, 2002; Hartsuiker et al. 2001; Solomon & Pearlmutter, 2004; Vigliocco & Nicol, 1998). The differences among these accounts lie in what allows grammatically illicit features to influence number valuation on the verb. There are two major proposals: (1) features spread erroneously in the noun phrase due to inherent properties of hierarchically structured representations; or (2) features of simultaneously activated constituents are confused in the planning of the verb form. A key difference between the two kinds of explanation for subject-verb agreement concerns the representation of the subject itself. On the first view, incorrect agreement is realized on the verb in production because the representation of the subject itself is internally inconsistent or faulty, due to properties of the combinatorial steps by which the representation is formed and maintained. On the second view, there is nothing inherently inconsistent or unreliable in how structure is represented over time, but only in how it is that distinct constituent representations are manipulated or accessed in language processing.

The current study examines agreement attraction in comprehension rather than production, bringing a different kind of evidence to bear on the debate. Data from seven experiments using both online and offline measures support the latter view of agreement attraction. In particular, we argue that agreement attraction in comprehension arises in the process of reaccessing information about the subject on the basis of cues at the verb, and not because the number of the subject is incorrectly represented. We present two related accounts of agreement attraction. In both, attraction reflects similarity-based interference that stems from cue-based retrieval (McElree, Foraker, & Dyer, 2003; Lewis & Vasishth, 2005).

Production

In the wake of Bock and Miller's seminal work (1991), numerous studies have established the parameters of agreement attraction in production. Most studies use an elicitation paradigm in which participants are presented with an auditory preamble constituting a subject phrase, like "The key to the cabinets," which they must repeat and finish. When the head noun is singular (i.e., *key*) and the embedded, or *local*, noun is plural (i.e., *cabinets*), there is a robust tendency for participants to use a verb form with plural morphology, even though the subject phrase is itself singular. Across studies using simple count nouns, 13% of singular-plural preambles, on average, are continued with the plural verb form (Eberhard, Cutting, & Bock, 2005). It is crucial that the head noun be singular and the local noun plural, and not vice versa. Preambles like "The keys to the cabinet" lead to very few agreement errors (on average, 3%; Eberhard, Cutting, & Bock, 2005).

For the simplest case, i.e., N_{SG}-P-N_{PL}-V,[1] it has sometimes been suggested that the linear proximity of the plural noun, with respect to the verb, is responsible for attraction (e.g. Quirk, Greenbaum, Leech, & Svartvik, 1985). Indeed agreement attraction has sometimes been referred to as proximity concord (Quirk et al., 1985; Francis, 1986). Such an account is also consistent with contemporary models of comprehension in which parsing decisions are heavily influenced by local surface statistics (e.g., Tabor, Galantucci & Richardson, 2004). However that view is undercut by the plural markedness effect, as well as by the existence of structural distance effects. Bock and Cutting (1992) and Solomon and Pearlmutter (2004) both show that local plurals embedded in a relative clause (e.g., "The editor who rejected the books") do not induce attraction as strongly as local plurals in a PP modifier. Vigliocco and Nicol (1998) demonstrate that linear adjacency is not a requirement for attraction, showing a similar proportion of subject-

verb agreement errors when participants produce inverted yes/no questions like “Is/are the helicopter for the flights safe?” relative to declarative sentences in which the verb follows the subject. Franck et al. (2002) found that when the subject phrase contained two stacked PP modifiers like “The inscription(s) on the door(s) of the toilet(s)” the medial prepositional phrase led to more attraction errors than the most deeply embedded one. These results suggest that structural distance between a potential attractor and the subject head noun, or its syntactic projection, impacts the likelihood of attraction more so than its linear distance to the verb (although see Haskell & McDonald, 2005, for arguments that small but measurable linear distance effects also exist).

Comprehension

In trying to understand the relation between agreement attraction and syntactic representation, evidence from agreement processing in comprehension provides a useful complement to evidence from production. Findings in comprehension are generally convergent with those from production: In the same scenarios in which individuals are likely to produce an agreement error, they experience less difficulty in processing an agreement error (Pearlmutter, Garnsey & Bock, 1999). Whereas the key measure in production studies of agreement attraction is the proportion of agreement errors that an individual produces, the key measure in comprehension studies is typically either acceptability judgments (Häussler & Bader, 2007; Clifton, Frazier & Deevy, 1999) or some index of processing difficulty, as reflected in a localized reading time difference (e.g. Pearlmutter, Garnsey & Bock, 1999), or a signature pattern of evoked potentials (e.g. Kaan, 2002). A unique aspect of agreement studies in comprehension is in the ability to examine the response to both grammatical and ungrammatical agreement in agreement attraction environments.

Mechanisms

Based on the existence of hierarchical distance effects, previous authors working mainly in production have proposed that agreement attraction is a result of feature movement or 'percolation' within a syntactic representation (Nicol, Forster, Veres, 1997; Vigliocco & Nicol, 1998; Franck, Vigliocco & Nicol, 2002; Eberhard, Cutting & Bock, 2005). On this account, temporary structured representations are leaky enough that information is sometimes spuriously transmitted through the structural links between constituents. In percolation, features on a given syntactic constituent can be transferred to other, nearby constituents, but they can only be transferred one syntactic 'step' at a time; in other words, features must pass first to the immediately dominating syntactic node, then to the next, and so on. This 'stepwise' movement is reflected in the reduced likelihood of feature movement with increasing syntactic distance between nodes. In the typical agreement attraction case of a subject with a PP modifier ("*the key to the cabinets"*), the number features bound to the noun phrase within the PP percolate upward, valuing higher phrasal projections for number. In some proportion of cases, these features can percolate up to the highest projection, that of the subject noun phrase. By hypothesis, the verb or verb phrase is reliably valued by the number on the subject phrase, and so will be inappropriately valued in just that proportion of cases when the PP-object's number percolates far enough to value the subject phrase.

The mechanism of feature percolation naturally accounts for structural depth effects like those demonstrated by Bock and Cutting (1992) and Franck et al. (2002); the most deeply embedded noun in a stack of PPs (like *toilets* in "the inscription on the door of the toilets") or in a relative clause (like *books* in "the editor who rejected the books") is uncontroversially more distant from the subject head noun than the noun in a single PP modifier (like *cabinets* in "the

key to the cabinets"). Furthermore, feature percolation naturally encompasses the plural markedness effect if the feature system is privative (Trubetzkoy, 1939), i.e., singular number is represented by default in the absence of any number feature, whereas plural number is represented by the presence of plural feature (e.g., Kimball & Aissen, 1971; Eberhard, 1997). This implies that in the case of embedded singulars there should be no feature to percolate upwards, and therefore there should be no possibility for attraction to occur. In comprehension, feature percolation has been implemented through a 'head overwriting' mechanism in which the number feature on the head noun of the subject is 'overwritten' with the features of the attractor noun (Pearlmutter, Garnsey, & Bock, 1999).

A related recent model of agreement processing also assumes that feature movement is responsible for agreement attraction effects, but generalizes the availability of movement (Eberhard, Cutting, & Bock, 2005). Results from a Dutch study by Hartsuiker and colleagues showed that $Subject_{SG}$-$Object_{PL}$-Verb configurations can induce agreement attraction in production, although error rates were lower than is generally observed in PP-modifier constructions (Hartsuiker, Antón-Méndez, & van Zee, 2001; although cf. Hemforth & Konieczny, 2003). Because the subject and object are not directly structurally related, a feature percolation model would have to assume that the intervening verb could take on the number features along the way. Instead, Eberhard and colleagues propose that the path for feature movement should be defined less rigidly than is the case in strict percolation approaches, and that any number information in a structure can spread to any other part of the structure in principle. However, it is unclear to what extent this model can capture the structural distance facts discussed above without assuming principles of feature movement very similar to those assumed by the strict percolation model.

The feature movement models discussed both share the property that they allow formation of a syntactic representation (here the subject) that is inconsistent with principles of the grammar. On the other hand, other models propose that agreement attraction effects arise primarily during the construction of the representation, and are more explicitly tied to the order in which constituents are planned (Bock & Cutting, 1992; Solomon & Pearlmutter, 2004; Badecker & Lewis, 2007). Solomon and Pearlmutter (2004) argue that attraction errors occur because features of both nouns are simultaneously activated in planning, with the distractor noun interfering with agreement processes before the subject representation is completed. Structural distance between the head and the attractor should correlate with the likelihood of errors because structural distance often reflects conceptual relatedness between the two, which in turn directly affects planning. Consistent with this, they show a correlation between agreement attraction in production and semantic relatedness when structural distance is held constant: More tightly related distractor nouns induce more agreement errors (e.g., "the pizza with the yummy toppings" versus "the pizza with the tasty beverages"). In comprehension, this simultaneity of feature activation might be realized when a verb is encountered and its subject must be retrieved from among other noun phrases in memory to check the verb's number features. The key point for this type of account is that the errors that arise do not stem from a faulty representation of the subject, but rather involve errors in mapping or retrieval in which multiple elements of the structure are simultaneously active.

Although there are formally well understood and well-motivated syntactic rules of percolation (Jackendoff, 1977, Gazdar, Klein, Pullum, & Sag, 1985, Shieber, 1986, van Riemsdijk & Williams, 1986, Bayer & Johnson 1995), the explanations proposed in the agreement attraction literature imply a percolation mechanism that is importantly different, in

that it allows agreement relationships to be formed that fail to follow grammatical constraints. There are at least two construals of the erroneous nature of this proposed mechanism: (1) errors reflect over-application of the regular feature percolation rules (Vigliocco & Nicol, 1998; Pearlmutter, Garnsey, Bock, 1999); or (2) errors results from a rather more unconstrained system of feature spreading (Eberhard, Cutting, & Bock, 2005).

Under the first construal, the application of feature percolation is errorful because it sometimes ignores pertinent structural conditions. In grammatical systems percolation is constrained by structural notions like 'head': a syntactic projection is constrained to inherit certain features, like number or syntactic category, from its head (e.g., Pollard & Sag, 1994). In feature percolation accounts of agreement attraction, these features are inherited from either a non-head or the wrong head.

Under the second construal, the erroneous spread of features results from lower-level properties of the encoding architecture, such as spreading activation. Eberhard et al. (2005) in particular seem to have this construal in mind when they write:

> When a source of number information is bound to a temporary structural network for an utterance, it transmits its information to the structure. Within the structure, the information moves or spreads according to principles of structural organization, assembly, and dissolution (p.543) ... Because SAP [*the model's continuously-valued number feature –WLP*] may flow unobstructed throughout a structural network, number information bound anywhere within a structure has the potential to influence agreement processes. (p.544).

In spreading activation models of combinatorial structure (e.g., Selman & Hirst, 1985, van der Velde & de Kamps, 2006; see Plate, 2003, for a comprehensive review), the potential for interaction between elements bound together in a structure generally implies the presence of connectivity between those elements. If feature percolation is akin to spreading activation, then positing the 'unobstructed' spread of features would also seem to require positing denser

connectivity, since elements that the grammar does not typically relate, like the number feature of a complement and the maximal projection of its selecting head, must nonetheless influence one another. Thus, an encoding architecture that permits unconstrained feature spreading may actually be more resource intensive than one that does not.

The current study

In this study we examine the question of whether agreement attraction is due to instability in the representation of features on the subject, or rather to errors in accessing the subject representation or constructing the agreement dependency. We address this question in two different ways.

After validating the paradigm with Experiment 1, in Experiments 2 and 3 we consider an English construction that has been reported to induce attraction-like agreement patterns, but does not involve a complex subject. Kimball and Aissen (1971) report a pattern of acceptability in which sentences that contain a N_{SG}-V_{PL} sequence within a singular-headed relative clause are uncontroversially reported as unacceptable (2a vs. 2b), but when the relative clause head noun is marked plural the acceptability of illicit N_{SG}-V_{PL} sentences increases (3a vs. 3b). This configuration thus exhibits a plural markedness effect similar to the complex subject case of agreement attraction: acceptability of subject-verb agreement is not impacted by the number of a nearby noun when the subject is plural, but only when it is singular.

(2) (a) The driver$_{SG}$ who the runner$_{SG}$ waves$_{SG}$ to each morning honks back cheerfully.
(b) * The driver$_{SG}$ who the runner$_{SG}$ wave$_{PL}$ to each morning honks back cheerfully.
(3) (a) The drivers$_{PL}$ who the runner$_{SG}$ waves$_{SG}$ to each morning honk back cheerfully.
(b) ? The drivers$_{PL}$ who the runner$_{SG}$ wave$_{PL}$ to each morning honk back cheerfully.

Kimball and Aissen's study was observational (and they attributed the pattern to a Northeast US/Boston dialect), but recent experimental work using cumulative acceptability judgments

supports the original judgments (Clifton, Frazier & Deevy, 1999). These configurations have been shown to induce attraction effects in production (Bock and Miller, 1991; Franck, Lassi, Frauenfelder, & Rizzi, 2006), and we have also found these types of errors in well-edited texts (4). However, to date attraction in this construction has not been examined in comprehension, and has not been systematically compared with attraction in the more well-studied prepositional-modifier cases.

(4) (a) We can live with the [$_{NP}$ errors$_{PL}$][$_{RC}$ that **classification software**$_{SG}$ make$_{PL}$ ...]
(Nunberg, 2003, p. 5)
(b) These include ... the [$_{NP}$ problems$_{PL}$][$_{RC}$ that **incrementality**$_{SG}$ pose$_{PL}$ ...]
(Byram, 2007, p. 58)

On a percolation account, such constructions should not induce strong attraction effects. The hierarchical relation between the agreeing subject and the attractor is inverted, in comparison to the PP cases. Firstly, traditional percolation mechanisms are specified for percolation 'upward' through a tree only, and thus do not predict that features can also travel 'downward', as they would have to in this case. Secondly, even if downward percolation were also allowed, the probability that the features of the attractor could percolate through the numerous intervening nodes to reach the subject should be fairly low, and certainly lower than in the PP-modifier case. Furthermore, the attractor does not linearly intervene between the subject and the verb, so the features of the attractor should presumably be less activated relative to those of the head noun. Thus, finding attraction effects of comparable size to the PP-modifier case would provide evidence against classic percolation models of agreement attraction.

In Experiments 4-7 we focus on a prediction not just of percolation models, but any account that assumes that errors result from incorrect representation of the subject. The prediction is that the reduced disruption due to agreement attraction in the comprehension of ungrammatical

sentences should be mirrored by an increased disruption of approximately the same size due to agreement attraction in grammatical sentences. On such accounts, agreement attraction in comprehension is due to the fact that in some proportion of trials, simple N_{SG}-P-NP_{PL} subjects are incorrectly valued as plural ([N_{SG} [P-NP_{PL}]]$_{PL}$) through percolation or some other mechanism, before the verb is ever encountered. Whenever this happens, N_{SG}-P-NP_{PL}-V_{PL} sequences should incorrectly be considered grammatical, and importantly, N_{SG}-P-NP_{PL}-V_{SG} sequences should incorrectly be considered *ungrammatical.* Therefore, on such accounts, the rate of errors in grammatical and ungrammatical sentences should be tightly linked.

The existing data relevant to this prediction of attraction symmetry in grammatical and ungrammatical sentences is somewhat mixed. Nicol, et al. (1997) showed that offline responses were slower in grammatical sentences with a PP-internal plural attractor (*The key to the cabinets is…*) than in sentences with a singular attractor, and Pearlmutter et al. (1999) provide evidence from eyetracking and self-paced reading that in grammatical sequences, PP-internal plural attractors lead to greater processing difficulty in the verb region. However, when ungrammatical and grammatical cases have been included in the same experiment (e.g. Exps. 1 & 2 in Pearlmutter et al., 1999), the timing and duration of attraction effects are significantly different between the two cases, which is unexpected on the percolation account. More importantly, comparing grammatical, singular attractor sentences and grammatical, plural attractor sentences is potentially confounded by the additional processing load that may be incurred by adding a plural attractor NP to the sentence. Plurals are morphologically and arguably conceptually more complex than singulars, and semantically composing plurals with other predicates in the sentence may also be more costly. Noun number is well known to affect lexical decision times (New, Brysbaert, Segui, Ferrand & Rastle, 2004; Lehtonen, Niska, Wande, Niemi & Laine, 2006; Lau,

Rozanova, & Phillips, 2007). When the plural cost and the attraction effect go in opposite directions, as in ungrammatical attraction cases where the attractor improves the sentence, there is no danger of one being mistaken for the other; the danger arises when the potential effects go in the same direction, as in grammatical sentences. This is particularly problematic for online experiments in which the attractor region is immediately adjacent to the verb, as the strongest effects of plural processing complexity would likely be felt here.

In five experiments (Exps. 2-6) we revisit the prediction of attraction symmetry in ungrammatical and grammatical contexts by comparing the effect of attractor number on reading times for ungrammatical and grammatical sentences in both the traditional PP-modifier attraction constructions as well as the relative clause constructions discussed above, in which the plural attractor region is well separated from the critical verb position. A finding of 'grammatical asymmetry' in which attraction effects disappear in grammatical sentences would be problematic for any model that assumes that it is the faulty representation of the subject noun phrase itself that is responsible for attraction in comprehension.

Note that although we describe the results separately for ease of presentation, with the exception of Experiment 1, a test session always included materials from both a relative-clause manipulation and a prepositional modifier manipulation. In particular, the materials from Experiments 3 and 5 were tested in the same session and the materials from Experiments 2 and 6 were tested in the same session.

Experiment 1

Previous behavioral studies of agreement in comprehension have rarely tested the strength of the ungrammaticality effect in the baseline case, that is, when subject and verb mismatch in number and no attractor NP is present. The goal of Experiment 1 was to determine the size and time course of the ungrammaticality effect in self-paced reading for this baseline case. This experiment also provided an opportunity to test whether the additional complexity due to processing a plural would in itself have an effect on self-paced reading times. Finally, the contrast between singular and plural subject-verb mismatch in the no-attractor case is also of some interest, given that sentences with singular and plural subjects are differentially sensitive to attraction effects.

Participants

Participants were 32 native speakers of English from the University of Maryland community with no history of language disorders. All participants in this and following experiments provided informed consent. Participants were either compensated $10/hr or received credit in an introductory Linguistics class. No participant took part in more than one experimental session.

Materials

Materials for all experiments are presented in Appendix A. Experimental materials consisted of 24 sentence sets arranged in a 2 × 2 design with subject number (singular/plural) and grammaticality (grammatical/ungrammatical) as factors. 16 sentence sets were derived from the materials of Pearlmutter et al. (1999), with attractor nouns removed, and 8 further sentence sets were created for this study. An example set is presented in Table 1. The first five words of each experimental item always followed the same form: *determiner-adjective-noun-adverb-verb*. The verb was always either auxiliary or copular 'be'. Since singular and plural subjects differ both in

the number of characters and morphological complexity, an adverb was included to make it easier to distinguish any effects due to the noun from effects caused by the verb itself. The word following the critical verb was either an adjective or past participle and never carried agreement. The 24 sets of 4 conditions were distributed across 4 lists in a Latin Square design, and combined with 84 grammatical filler sentences of similar length (48 of these fillers consisted of two separate manipulations unrelated to agreement processing which are not presented here). This resulted in 11% of the items being ungrammatical.

Procedure

Sentences were presented on a desktop PC using the Linger software (Doug Rohde, MIT) in a self-paced word-by-word moving window paradigm (Just, Carpenter, & Woolley, 1982). Each trial began with a screen presenting a sentence in which the words were masked by dashes while spaces and punctuation remained intact. Each time the participant pressed the space bar, a word was revealed and the previous word was re-masked. A yes/no comprehension question appeared all at once on the screen after each sentence. The 'f' key was used for 'yes' and the 'j' key was used for 'no'. Onscreen feedback was provided for incorrect answers. Participants were instructed to read at a natural pace and answer the questions as accurately as possible. In all of the self-paced reading experiments presented here, participants were never informed that sentences would contain grammaticality errors. Order of presentation was randomized for each participant. 7 practice items were presented before the beginning of the experiment.

*** INSERT TABLE 1 HERE ***

Analysis

Only items for which the comprehension question was answered correctly were included in the analysis. Reading times that exceeded a threshold of 2.5 standard deviations, by region and

condition, were excluded (Ratcliff, 1993). Across all self-paced reading experiments reported in this paper, on average by region this standard deviation trim resulted in the exclusion of 2.6% of correctly-answered trials (min: 2.3%, max: 3.1%). The regions used for analysis consisted of single words. Data for each of the first 10 regions of the sentence were entered into a 2 × 2 repeated-measures ANOVA with subject number and grammaticality as factors. Using R (R Development Core Team, Vienna), ANOVAs were computed on participant mean reading times across items (F1) and on item means across participants (F2). Min F′ statistics (Clark, 1973; Raaijmakers, Schrijnemakers, & Gremmen, 1999) were also computed, although because our items were counterbalanced across lists, this test is probably too conservative (see Raaijmakers, et al., 1999 for discussion). ANOVA statistics for regions of interest are presented in Table 2: these were the subject region (R3), the adverb region (R4), the critical verb region (R5), and the four regions following the critical verb (R6-10). We performed a complementary analysis by fitting linear mixed-effect models to our data, simultaneously controlling for subject and item as random factors. Models were fit using restricted log-likelihood maximization and 95% confidence intervals (CIs) were then derived from these models by Markov Chain Monte Carlo simulation (Baayen, Davidson, & Bates, 2008). These are presented in the text.

Results

Comprehension Question Accuracy

The mean comprehension question accuracy for experimental items across participants and items was 96.7%, and did not differ across conditions (logistic mixed-effects model, *p*s n.s.). For singular subjects, grammatical sentences were answered with an accuracy of 95.3% ± 1.7% and ungrammatical sentences with an accuracy of 96.9% ± 1.4% (standard error is computed across participants means). For plural subjects, grammatical accuracy was 95.8% ± 1.5% and

ungrammatical accuracy was 99.0% ± 0.7%.

Self-paced reading

The results from Experiment 1 are presented in Figure 1. The subject region (R3) showed no significant differences (*F*s < 1). However, the subsequent adverb region (R4) showed a main effect of number, such that the plural conditions had longer reading times (plural mean = 480 ms; singular mean = 417 ms; 95% CI = 23 ms; $p < .001$).

The critical verb region (R5) showed a main effect of grammaticality (grammatical mean = 402 ms; ungrammatical mean = 447 ms; 95% CI = 23 ms; $p < .01$), no significant effect of number and no interaction (Fs < 1). This grammaticality effect persisted into regions 6-8, and was largest in the immediate post-verbal region (R6) (grammatical mean = 374 ms; ungrammatical mean = 477 ms; 95% CI = 18 ms; $p < .0005$). Although the singular-subject ungrammatical condition showed somewhat longer reading times than the plural-subject ungrammatical condition in the immediate post-verbal region (R6), the interaction was nonsignificant. In a pooled analysis of the three immediate post-verbal regions (R6-8), the coefficient of the grammaticality:number treatment effect did approach marginal significance (-17 ms; 95% CI [-38 ms, 6 ms]; $p < .15$).

*** INSERT FIGURE 1 HERE ***
*** INSERT TABLE 2 HERE ***

Discussion

The main effect of number that we found in the region immediately following the subject head noun suggests that some aspect of reading the plural noun does exact a significant reading time cost on the following region. This plural cost could be due to morphological processing, conceptual processing, semantic integration, or even length per se (e.g., the additional character

could cause fixations to be less optimally positioned). The current data do not allow us to discriminate between these possible explanations, which make different predictions about the duration of the effect and the extent to which the effect should be observed for nouns in other positions in the sentence.

Because number and grammaticality manipulations occur nearby in sentences used in agreement attraction experiments, these results suggest that some care must be taken in deciding whether reading time increases should be attributed to factors associated with processing nominal number alone or to factors associated with licensing an agreement relation. In the case of the simple sentences in Experiment 1, the inclusion of the adverb allowed us to see that effects of noun number preceded the critical verbal region. However, the effect of number was larger than plural effects that we have found previously (Lau, et al., 2007), perhaps related to the longer reading times and increased variability in reaction times observed, or perhaps because the critical noun appeared early in the sentence.

The main effect of grammaticality in the critical verb region shows that grammaticality effects can be found on the critical region itself in a self-paced reading paradigm, and that the differential sensitivity of singular and plural subjects to attractors is not a function of differential initial sensitivity to subject-verb agreement errors in the baseline no-attractor case. The main effect of grammaticality at the critical region for these no-attractor cases was 45 ms (control s.d.: 137 ms; Cohen's d = 0.33). The lack of a reading time difference between the two ungrammatical conditions at the verb also suggests that the form of the verb ('was'/'were') does not have a significant effect on reaction times. Interestingly, singular subjects led to a consistently larger deflection in reading times for ungrammatical conditions in later regions, but this effect was not reliable.

Experiment 2

Experiment 2 was designed to test the online effects of agreement attraction in object relative clause constructions, in which the relative clause head is a potential attractor with respect to the agreement between the relative clause subject and verb. This configuration is doubly interesting. Firstly, the plural relative clause head does not intervene, either linearly or hierarchically, between the subject and the verb. Finding agreement attraction in this configuration would present a challenge for straightforward feature percolation accounts, as the dominance path that connects the relative clause head and the relative clause subject runs opposite to that that found in PP constructions and requires both upwards and downwards percolation. Secondly, in this construction the local environment immediately preceding the verb is identical regardless of the number on the potential attractor, meaning that the comparison of grammatical cases is less susceptible to plural complexity effects.

Because previous literature in both comprehension and production has shown that singular subject/plural attractor configurations demonstrate much larger and more reliable effects than plural subject/singular attractor configurations (Eberhard, 1997, Pearlmutter, et al., 1999; cf. Pearlmutter, 2000), we tested only singular subject cases, where the attractor could be either singular or plural. If this configuration is subject to attraction effects, we should see differences between plural attractor cases, in which the attractor differed in number from the subject, and singular attractor cases, in which the number on both nouns was the same.

Participants

Participants were 30 native speakers of English from the University of Maryland community with no history of language disorders.

Materials

Experimental materials consisted of 48 sentence sets arranged in a 2 × 2 design with relative clause head number (singular/plural) and grammaticality (grammatical/ungrammatical) as factors. An example set is presented in Table 3. The first six words always contained a noun phrase modified by a relative clause, following the template *determiner-noun-'who'-determiner-noun-verb.* The agreement relation manipulated here was the agreement between the subject noun and verb inside the relative clause, and thus the head noun modified by the relative clause was considered the 'attractor'. The subject of the relative clause was always singular. In this design the noun immediately adjacent to the verb was always singular, such that effects on the verb were less likely to be conflated with effects of noun number per se. The word following the critical verb was usually a short function word and never carried agreement. The 48 sentence sets were distributed across 4 lists in a Latin Square design, and were combined with 24 items of a prepositional-phrase agreement attraction design (all grammatical; data presented below as Experiment 6) and 144 filler sentences of similar length. This resulted in 17% of the items being ungrammatical.

*** INSERT TABLE 3 HERE ***

Procedure and Analysis

The same self-paced reading procedure was used as in Experiment 1. The analysis followed the same steps. Only items for which the comprehension question was answered correctly were included in the analysis. Reading times that exceeded a threshold of 2.5 s.d. by region and condition were excluded. Two participants showed a comprehension question accuracy rate of less than 80% across all items and were thus excluded from further analysis.

The regions used for analysis consisted of single words. Data for each of the first 10 regions of the sentence were entered into a 2 × 2 repeated-measures ANOVA with attractor number and

grammaticality as factors. Regions of interest were 2-9; the attractor noun appeared in region 2, the subject noun appeared in region 5, and the critical verb appeared in region 6. Statistics for each of these 9 regions are presented in Table 4.

Results

Comprehension Question Accuracy

The mean comprehension question accuracy for experimental items across participants and items was 92.3%, and did not differ across conditions (logistic mixed-effects model, all *p*s n.s.). For singular RC head conditions, accuracy in the grammatical condition was 94.3% ± 1.6% and in the ungrammatical condition 91.7% ± 1.8%. For plural RC heads, accuracy in the grammatical condition was 93.2% ± 1.5% and in the ungrammatical condition 90.2% ± 2.2%.

Self-paced reading

As shown in Figure 2, region 2 (the attractor) showed a main effect of attractor number, such that the plural conditions had longer reading times (plural mean = 322; singular mean= 311; 95% CI = 7.7 ms, $p < .05$). Since the verb was not encountered until region 6, this effect was likely driven by additional processing costs associated with plural nouns relative to singular ones. Longer reading times for the plural-attractor conditions persisted to the critical verb (R6); the main effect was reliable at the relative pronoun (R3) and the relative clause subject noun (R5).

The critical verb region (R6) did not show a main effect of grammaticality (*F*s < 1). However, the region following the verb (R7) showed main effects of attractor number and grammaticality, and crucially, an interaction between grammaticality and attractor number. Pairwise comparisons in region 7 showed a significant grammaticality effect only when the relative clause head was singular (grammatical mean = 337 ms; ungrammatical mean = 399 ms; 95% CI = 27.2 ms, $p < .05$), but not when it was plural (grammatical mean = 331; ungrammatical

mean = 341; 95% CI = 18.7 ms, $p > .1$). No significant effects were found in region 8, but in regions 9 and 10 the main effect of attractor number again was significant, as well as the effect of grammaticality in region 10 (significant by participants, marginal by items).

Several previous studies have reported that the presence of an attractor not only reduces disruption to subject-verb mismatches when the attractor matches the verb, but also causes disruption in grammatical subject-verb match sentences when the attractor mismatches the verb. To investigate this possibility in our relative clause materials we conducted a pairwise comparison over the two grammatical conditions in the first post-verbal region (7), where we found the attraction effect in the ungrammatical cases. No significant differences were found (singular attractor mean = 337 ms; plural attractor mean = 331 ms; 95% CI = 16.3 ms; $p > .1$).

*** INSERT FIGURE 2 HERE ***
*** INSERT TABLE 4 HERE ***

Discussion

The results of Experiment 2 suggest that the head of a relative clause can act as a strong attractor for agreement. When the RC head and the subject were both singular, the region following the critical verb showed significantly longer reading times in the ungrammatical condition, where the verb was plural, than in the grammatical condition, where the verb was singular. However, when the RC head was plural, the ungrammatical singular subject-plural verb combination did not differ from the grammatical singular subject-singular verb condition in the region following the verb. Thus, we found that for subjects within RCs, an RC head attractor matching the RC verb's number significantly reduced the reading time disruption normally seen to subject-verb number mismatch. A similar reduction in the reading time disruption has previously been observed for agreement attraction caused by subject-attached PPs (Pearlmutter et al., 1999).

These results contrast with the predictions of a standard percolation account. Such accounts would predict that because the RC head noun hierarchically commands the RC subject, it should exert no attraction effect whatsoever. These data also contrast with a more liberal version of percolation, in which downwards feature percolation is also permitted. Under the assumption that the structural distance between two nodes determines the likelihood that features will spread from one node to the other, then RC attraction should be relatively weaker than in the PP-modifier case given the greater structural distance between the RC head noun and the RC subject. In fact, the attraction effect was as strong as possible, in the sense that the disruption due to ungrammaticality was eliminated entirely. One alternative explanation for this pattern is that, because of the increased processing demands of a relative clause structure, it is hard for the reader to notice ungrammaticality even in the baseline case. However, we found that the ungrammatical condition in which neither noun matched the verb's number showed a slow-down of 62 ms in the region following the critical verb (pooled s.d.: 201 ms; Cohen's *d*: 0.31), relative to the corresponding grammatical condition in which both nouns matched the verb. Both the absolute and normalized mean differences were comparable to the grammaticality effect observed in Experiment 1, despite significant differences in verb type (copula/auxiliary vs. main verb) and in experiment-wide reading times. Thus subject-verb mismatch seems equally disruptive whether it occurs in the main clause or within a relative clause.

We also found that there was no attractor effect in the grammatical conditions. In contrast to the PP-modifier configuration, in which the attractor is adjacent to the verb, the attractor was non-intervening in the RC configuration. Consequently the material adjacent to the verb was identical across singular and plural attractor conditions resulting in a better-matched baseline prior to the verb. We found that in these RC configurations the grammatical plural attractor

condition did not show significantly longer reading times than the grammatical singular attractor condition.

Experiment 3

The results of Experiment 2 showed that the presence of a plural RC-head reduced the disruption due to mismatch between a singular RC-subject and a plural RC-verb. This appears similar to previous findings of attraction effects in prepositional constructions (*The key to the cabinets are…*) in both comprehension and production. However, since the order of attractor and subject was reversed, it is conceivable that our effect represents a different process. To provide a further test of whether these phenomena represent the same underlying process, Experiment 3 tested for another property of the attraction effect found in prepositional constructions, namely the singular-plural asymmetry. Previous work has shown strong attraction effects for singular-subject/plural-attractor pairs, but not for the corresponding plural-subject/singular-attractor pairs (Pearlmutter et al., 1999). Finding the same asymmetry for the RC construction would strengthen the evidence that the attraction observed in RC structures has the same cause as classic attraction effects.

Participants

Participants were 60 native speakers of English from the University of Maryland community with no history of language disorders.

Materials

Experimental materials consisted of the same 48 sentence sets as in Experiment 2, but this time arranged in a 2 × 2 × 2 design with attractor number (singular/plural), subject number (singular/plural), and grammaticality (grammatical/ungrammatical) as factors. Although in this

design the noun immediately adjacent to the verb could be singular or plural, the comparisons of interest were between conditions in which the subject number was matched, and therefore effects of noun number were less of a concern. However, in an effort to minimize such effects the subject nouns were chosen such that on average across items they occurred equally frequently in the singular or the plural, since infrequent plural forms have been shown to have a greater effect on reading times than frequent ones (e.g. New et al., 2004; see Bock, Eberhard, & Cutting, 2004 for relevant effects in production). Counts were derived from the Cobuild Bank of English corpus (320 million words). The word following the critical verb was usually a short function word and never carried agreement. The 48 sentence sets were distributed across 8 lists in a Latin Square design, and were combined with 24 items of a prepositional-phrase agreement attraction design (half ungrammatical; data presented below as Experiment 5) and 216 filler sentences of similar length. This resulted in 13% of the items being ungrammatical.

Procedure and Analysis

The procedure was the same self-paced reading procedure described in Experiment 1, and the analysis followed similar steps. Only items for which the comprehension question was answered correctly were included in the analysis. Reading times that exceeded a threshold of 2.5 s.d. by region and condition were excluded. Due to experimenter error, the distribution of participants across the 8 lists was unbalanced. In the analysis presented here, a subset of 56 participants was analyzed such that participants were evenly balanced across the lists; however, the pattern of results did not differ from the analysis in which all 60 participants were included.

The regions used for analysis consisted of single words. Data for the first ten regions of the sentence were examined; the critical verb appeared in region 6. In this experiment, the comparisons of interest were all within a level of the subject number factor: we were interested

in whether the 4 plural subject conditions would show the same pattern relative to each other as the 4 singular subject conditions. In order to examine this question we split the design into two 2 × 2 repeated-measures ANOVAs, one for each level of subject number (singular/plural), with attractor number and grammaticality as factors. For completeness, we also computed a 2 × 2 × 2 repeated measures ANOVA with attractor number, subject number, and grammaticality as factors. Regions of interest were 2 (attractor), 3, 5 (relative clause subject), 6 (verb), 7, and 8; statistics for these regions are presented in Tables 5 and 6.

Results

Comprehension Question Accuracy

The mean comprehension question accuracy for experimental items across participants and items was 93.7%, and did not differ reliably across experimental conditions. For sentences with a singular RC subject and a singular RC head, accuracy in the grammatical condition was 93.3% ± 1.4% and in the ungrammatical condition 96.4% ± 1.1%. For sentences with a singular RC subject and a plural RC head, accuracy in the grammatical condition was 94.4 ± 1.3% and in the ungrammatical condition 94.7% ± 1.3%. For sentences with a plural RC subject and a singular RC head, accuracy in the grammatical condition was 94.4% ± 1.3% and in the ungrammatical condition 93.6% ± 1.3%. For sentences with a plural RC subject and a plural RC head, accuracy in the grammatical condition was 91.7 ± 1.7% and in the ungrammatical condition 91.7% ± 1.3%.

Inspection of the means suggested that the plural-attractor/plural-subject conditions were answered less accurately than the other conditions. A post-hoc comparison revealed a reliable effect of the presence of two plural nouns (plural attractor and subject), compared to one or none (mean of zero or one plurals = 94.5%; mean of two plurals = 91.2%; $p < .01$).

Self-paced reading

The results of Experiment 3 are plotted in Figure 3 (singular subject conditions) and Figure 4 (plural subject conditions). In regions 2 and 3, the relative clause head and relative pronoun, the omnibus ANOVA showed a main effect of attractor number, as in Experiment 2, due to slower reading times for the plural head conditions (R2: plural mean = 353 ms, singular mean = 341 ms, 95% CI = 6.5 ms, $p < .01$; R3: plural mean = 339 ms, singular mean = 331 ms, 95% CI = 4.8 ms, $p < .05$). The 2 × 2 ANOVAs, split by subject number, revealed that singular subject conditions show this effect more strongly in region 2, and plural subject conditions in region 3; this variation in timing was presumably random, as subject number was not manipulated until later in the sentence. In the by-items analysis there was a significant effect of subject number at Region 4. Region 5, the relative clause subject region, showed both main effects of attractor number and subject number (RC subject number: plural mean = 337 ms, singular mean = 329 ms, 95% CI = 6.5 ms, $p < .05$; Attractor number: plural mean = 339 ms, singular mean = 327 ms, 95% CI = 6.4 ms, $p < .05$). The effect of subject number in these regions appeared to be largely carried by an exceptional value for grammatical plural attractor/plural subject conditions, which may reflect an additional cost of having to keep track of two plural features. At the verb in region 6 there was a clear effect of RC subject number (plural mean = 365 ms, singular mean = 348 ms, 95% CI = 7.5 ms, $p < .005$). Region 7, the region following the critical verb, showed a main effect of grammaticality (mean ungrammatical = 403 ms; mean grammatical = 355 ms; 95% CI = 13.1 ms, $p < .005$), and a marginal interaction of grammaticality and attractor number. In Region 8, the 3-way interaction of grammaticality, attractor number, and subject number was significant by participants and items.

However, splitting the design by relative clause subject number revealed a pattern of

attraction similar to Experiment 2, but only for singular subjects. For singular subjects, the plural attractor conditions showed a smaller grammaticality effect than the singular attractor conditions, which was marginally significant in region 7 (Grammaticality:Ungram × RC Head:Pl Δμ= -33 ms; 95% CI = 37 ms, p < .10) and was significant in region 8 (Grammaticality:Ungram × RC Head:Pl Δμ = -50 ms; 95% CI = 32 ms, p < .01). By comparison, attractor number had no impact upon the grammaticality effect for plural RC subjects in either post-verbal region (R7 Grammaticality:Ungram × RC Head:Pl Δμ= -16 ms; 95% CI = 37 ms; R8 Δμ = 14 ms; 95% CI = 29). Note that whereas in Experiment 2 the plural attractor completely eliminated the disruption due to subject-verb mismatch, in Experiment 3 the disruption was merely reduced; however, the reason for this variation in effect size is unclear, as both experiments contained the same items in these conditions.

As in Experiment 2, we further tested for the existence of attraction effects in the grammatical conditions by conducting pairwise comparisons between singular and plural attractor conditions in grammatical sentences, for both singular and plural subject sets in the region following the critical verb (R7). We found no significant effects of attractor in either the grammatical singular-subject conditions (R7: plural mean = 356 ms, singular mean = 348 ms, 95% CI = 15.4 ms, $p > .1$; R8: plural mean = 361 ms, singular mean = 345 ms, 95% CI = 17.1 ms, p > .1) or the grammatical plural-subject conditions (R7: plural mean = 358 ms, singular mean = 355 ms, 95% CI = 14.6 ms, $p > .1$; R8: plural mean = 344 ms, singular mean = 353 ms, 95% CI = 15.2 ms, $p > .1$).

*** INSERT FIGURES 3-4 ***
*** INSERT TABLES 5-6 ***

Discussion

As in Experiment 2 we found that the presence of a plural attractor in the RC-head position

reduced the disruption due to ungrammatical subject-verb mismatch when the subject was singular. Furthermore, Experiment 3 showed that this attractor effect did not appear when the subject was plural; in this case, both ungrammatical conditions showed equal disruption relative to the grammatical baseline. This finding mirrors the same markedness pattern shown in previous literature on agreement attraction in production and comprehension of other constructions (Bock & Miller, 1991; Pearlmutter et al., 1999) and thus supports the idea that the attraction effect shown here has a similar basis, despite the different ordering of attractor and subject. Taken together, the results of Experiments 2 and 3 argue against a percolation account of agreement attraction in comprehension. Such accounts predict that the RC head should act only as a weak attractor at best, because of its structural relationship with the RC subject. However, we found strong attractor effects for this configuration in the standard agreement attraction pattern: attraction for singular-subject/plural-attractor conditions and not vice versa.

Although the finding of attraction in relative clause configurations argues against the classic percolation model, they do not argue conclusively against models that explain attraction through faulty representation of the subject. It might be possible to create a percolation model both in which percolation can occur downwards and in which the weights are arranged such that the structural distance effects observed in production studies of attraction are still predicted, but that both relatively close and relatively distant attractors can exert strong effects in comprehension. Relatedly, the model of Eberhard and colleagues (2005) leaves open the possibility for features to spread more freely than in percolation models, such that there could be a combination of parameter settings which would predict strong attraction effects even for structurally distant attractors. Such a less-constrained model might also draw on the fact that the RC head must be reactivated at the verb for thematic reasons, and could posit that the attractor features are spread

to the subject representation during this operation.

However the second major finding of Experiments 2 and 3 is much harder to explain for models that attribute attraction effects to internally inconsistent representations of the subject phrase. While in both Experiment 2 and Experiment 3 we found attractor effects in singular subject conditions, we found no parallel 'reverse' attractor effect in grammatical conditions, for either singular or plural subject conditions. In other words, the presence of an attractor noun mismatching the verb in number had no effect on reading times as long as the subject and verb did match in number. However, models of attraction in which number features spread among the previously processed elements of a sentence predict that the presence of a mismatching attractor noun should often lead to the wrong number-marking on the subject, with the consequence that a grammatical sentence should appear to the parser to be ungrammatical when the verb is encountered. The size of this increase in disruption should mirror the size of the reduction in disruption for attraction in ungrammatical sentences, because it is the same phenomenon that is assumed to drive both—feature spreading among the elements prior to the verb.

A potential concern about this logic is that even if feature percolation leads to a symmetric distribution of attraction in grammatical and ungrammatical sentences, the effect of attraction on mean reading times might be asymmetric. However, if we think of the reading time distribution for attractor sentences as sampled from two underlying RT distributions, one corresponding to 'perceived grammatical' and the other 'perceived ungrammatical', in proportions determined by the attraction rate, then the effect on the means should be linear and is thus expected to be symmetric. Analytically, the expectation value of two independently sampled distributions is a linear combination of the expected values of the two distributions (Brunk, 1975). We have

confirmed this to be true by simulation for the positively skewed ex-Gaussian distribution that describes reaction times (Luce, 1986; see Supplementary Materials).

Although our failure to find attraction effects in grammatical sentences presents a major challenge to theories which attribute attraction to faulty representation of subject number, some previous work has shown evidence for increased disruption in grammatical sentences with a mismatching attractor (Pearlmutter, Garnsey, & Bock, 1999; Nicol, Forster, & Veres, 1997; Häussler & Bader, submitted). These experiments differed from Experiments 2 and 3 in the configuration used, the local environment for the disruption, and in some cases, in methodology. In Experiments 4-6, we attempt to resolve these conflicting results.

Experiment 4

Experiments 2 and 3 showed an attraction effect in ungrammatical singular-subject sentences, but not in grammatical sentences. In all conditions where the subject agreed with the verb, we found no cost associated with the presence of an attractor that mismatched the verb. Only in ungrammatical sentences where the subject was singular and mismatched in number with the plural verb did the number of the attractor have an effect on reading times. However, Pearlmutter and colleagues (1999) report effects of attractor number in grammatical sentences in self-paced reading for prepositional modifier constructions.

One possible explanation for the divergence in results is that attraction effects in grammatical sentences depend for some reason on the subject-attractor-verb order of the PP-modifier construction. For example, percolation might be delayed in the relative clause constructions until the relative clause head is retrieved at the verb. If the additional assumption is made that the grammatical sentences are processed faster than the ungrammatical sentences, one could claim

that percolation proceeds too slowly to affect the grammatical conditions but just fast enough to affect the ungrammatical conditions. In order to address this and related concerns, we used the prepositional modifier construction in the remaining four experiments, in which we focus on the existence of attractor effects in grammatical sentences.

Another possible explanation for the different results is that the effect of attractor number observed by Pearlmutter et al. (1999) in the grammatical conditions may have been due to differences in the cost of processing the singular or plural attractor noun in the region prior to the verb. In Experiments 1-3 we saw significantly longer reading times associated with plural nouns, often in the region following the noun. As discussed in the Introduction, this confounding factor is a particular concern in the prepositional cases for the grammatical comparison, because the condition containing the potentially more complex plural is also the one that the percolation theory predicts to show an attraction-related slow-down, whereas in the ungrammatical conditions the predictions go in opposite directions (plural attractor reduces slow-down). The potential for confusion in interpreting these effects is exacerbated by the fact that in Experiments 2 and 3 both number and attraction effects tended to be found not on the noun or verb itself but on the subsequent word (this was also true for at least a subset of the attraction effects observed by Pearlmutter et al.,1999, in self-paced reading and eyetracking).

In Experiment 4 we examined prepositional modifier constructions in self-paced reading, using materials similar to those tested by Pearlmutter et al., 1999. However, as in Experiment 1, an adverb was inserted between the attractor noun and the verb. Although effects of plural complexity may extend beyond the adverb, depending on their source, reading times on the adverb region provide an index of effects of noun number independent of the verb.

Participants

Participants were 46 native speakers of English from the University of Maryland community with no history of language disorders.

Materials

Experimental materials consisted of 24 sentence sets arranged in a 2 × 2 design with attractor number (singular/plural) and grammaticality (grammatical/ungrammatical) as factors. As in Experiment 1, 16 of the sentences were modified versions of materials used in Pearlmutter et al. (1999); an example set is presented in Table 7. The subject of the sentence was always singular, and grammaticality was manipulated by varying the number of the verb (singular=grammatical, plural=ungrammatical). The first seven words of each experimental item always followed the same sequence: *determiner-noun-preposition-determiner-noun-adverb-verb*. Since singular and plural attractors differ in both number of characters and morphological complexity, the adverb was included to provide an index of effects of noun number onsetting before the critical verb region. In an effort to further minimize effects of noun number, the attractor nouns were chosen such that on average across items they occurred equally frequently in the singular or the plural according to the Cobuild Bank of English corpus. The word following the critical verb was either an adjective or a past participle and never carried agreement. The 24 sets of 4 conditions were distributed across 4 lists in a Latin Square design, and combined with 48 relative-clause agreement attraction sentences similar to those described in Experiments 2 and 3 (half ungrammatical) for a separate manipulation not discussed here, and 200 grammatical filler sentences of similar length. This resulted in 13% of the items being ungrammatical.

*** INSERT TABLE 7 HERE ***

Procedure

The same self-paced reading procedure was used as in Experiment 1, and the analysis

followed the same steps. Reading times that exceeded a threshold of 2.5 s.d. by region and condition were excluded. One participant showed a comprehension question accuracy rate of less than 80% across all items and was thus excluded from further analyses. Due to experimenter error, the distribution of participants across the 4 lists was unbalanced. In the analysis presented here, one further participant was excluded such that 44 participants were evenly balanced across the lists; however, the pattern of results did not differ from the analysis in which this participant was included.

Data for each of the first 10 regions of the sentence were entered into a 2 × 2 repeated-measures ANOVA with attractor number and grammaticality as factors. Regions of interest were 5-10; the attractor noun appeared in region 5, and the critical verb appeared in region 7. Statistics for each of these 6 regions are presented in Table 8.

Results

Comprehension Question Accuracy

The mean comprehension question accuracy was 89.6%. For grammatical conditions, mean accuracy for the singular attractor condition was 94.3% ± 1.5% compared to 85.2% ± 2.2% in the plural attractor condition; for ungrammatical conditions, mean accuracy for the singular attractor condition was 89.0% ± 2.2% compared to 89.8% ± 2.2% in the plural attractor condition. There was a reliable effect of attractor number ($p < .001$), and a reliable interaction with grammaticality ($p < .01$).

Self-paced reading

The results of Experiment 4 are presented in Figure 5 (Panel A). Regions 1-4 showed no significant effects. The attractor noun (R5) showed a main effect of number (singular mean = 284 ms; plural mean = 308 ms; 95% CI = 8.5 ms; $p < .001$). Since the effect preceded the critical

verb, the difference likely reflects the additional complexity of reading plurals versus singulars, as was observed in Experiments 1-3. The effect persisted in the adverb region (R6). At the critical verb (R7) there was an effect of grammaticality (grammatical mean = 332 ms; ungrammatical mean = 347 ms; 95% CI = 9.7; $p < .05$) and an interaction with attractor number. In the region following the verb (Region 8) there was a more pronounced main effect of grammaticality (ungrammatical mean = 349 ms; grammatical mean = 317 ms; 95% CI = 9.7 ms; $p < .001$), as well as again a significant grammaticality × attractor number interaction. The subsequent region (Region 9) also showed the same pattern. Finally, a main effect of grammaticality was observed at Region 11.

In order to test whether these effects reflected attraction we conducted pairwise comparisons between singular and plural attractor conditions at the verb (R7) and the region following the verb (R8) for grammatical and ungrammatical sentences separately. At the verb (R7), pairwise comparisons showed that there was no significant difference due to attractor number for the two ungrammatical conditions (singular mean = 351 ms; plural mean = 344 ms; 95% CI = 15.4; $p > .1$), but there was a difference for the two grammatical conditions (singular mean = 324 ms; plural mean = 341 ms; 95% CI = 14.5; $p < .005$). However, given that the region immediately prior to the verb also showed reading times for the plural attractor case to be significantly longer, it is difficult to unambiguously interpret this difference as an attraction effect. In contrast, in the region following the verb (R8), we found reading times in the plural attractor case to be significantly faster in the ungrammatical conditions (singular mean = 365 ms; plural mean = 333 ms; 95% CI = 15.8 ms; $p < .005$). There was no reliable difference in the grammatical conditions (singular mean = 313 ms; plural mean = 321 ms; 95% CI = 12.6 ms; $p > .1$).

*** INSERT FIGURE 5 ***

*** INSERT TABLE 8 ***

Discussion

The results of Experiment 4 replicated previous demonstrations that agreement attraction from a prepositional modifier reduces the disruption caused by an ungrammatical subject-verb mismatch. However, in contrast to some previous reports, the regions showing the attraction effect (the regions immediately following the verb) did not show a corresponding effect in the grammatical conditions. We did, however, find a modest main effect of attractor number on the attractor noun itself, which continued through the subsequent adverb region, and continued to the verb in grammatical conditions. This effect is likely due to differences in processing cost for singular and plural nouns. The size of the number effect (noun region: 24 ms; adverb region: 17 ms) was smaller than that observed in Experiment 1, but was in line with the size of effects seen in other studies (~15 ms: Lau et al. 2007).

The lack of an attraction effect for grammatical sentences in the prepositional modifier construction is consistent with Experiments 2 and 3, in which we also failed to find effects of attractor number for grammatical sentences. These findings conflict with the predictions of models discussed above in which feature-spreading causes attraction effects in creating an internally inconsistent representation of subject number. However, the findings of Experiment 4 must be considered cautiously given the previous results that did show attraction effects for grammatical sentences (Pearlmutter et al., 1999; Nicol et al., 1997). In Experiment 5, we replicate Experiment 4 using materials that are more similar to those used in previous studies.

Experiment 5

Experiment 4 showed that when an adverb intervened between the attractor and the verb in prepositional modifier constructions, a clear attraction effect was seen only in ungrammatical

sentences. We observed significant differences before the verb was ever encountered, which seemed to be due to increased processing requirements for plural nouns. These differences began at the noun and extended into the verb region. In contrast, the classic attraction effect for ungrammatical sentences did not become significant until the region following the verb, as in Experiments 2 and 3. We suggested that costs due to the increased complexity of the plural may have resulted in an increase in reading time for plurals over singulars in self-paced reading. This may sometimes create the illusion of an early-onset attraction effect for grammatical sentences, particularly when an adverb is not included. In Experiment 5, we performed a more direct replication of previous studies by examining the attractor effect in the absence of the adverb.

Participants

Participants were 60 native speakers of English from the University of Maryland community with no history of language disorders.

Materials

The experimental materials consisted of 24 sentence sets arranged in a 2 × 2 design with attractor number (singular/plural) and grammaticality (grammatical/ungrammatical) as factors. Sentences were the same as those described in Experiment 4 but without the adverb between the attractor NP and the verb. Thus, the subject of the sentence was always singular, and grammaticality was manipulated by varying the number of the verb. The first six words of each experimental item always followed the same sequence: *determiner-noun-prep-det-noun- verb*. The 24 sets of 4 conditions were distributed across 4 lists in a Latin Square design, and combined with 48 relative-clause agreement attraction sentences (half ungrammatical; described separately as Experiment 3), and 216 grammatical filler sentences of similar length. This resulted in 13% of the items being ungrammatical.

Procedure

The same self-paced reading procedure was used as in Experiment 1, and the analysis followed the same steps. Reading times that exceeded a threshold of 2.5 s.d. by region and condition were excluded. Due to experimenter error, the distribution of participants across the 8 lists was unbalanced. In the analysis presented here a subset of 56 participants were analyzed such that participants were evenly balanced across the lists; however, the pattern of results did not differ from the analysis in which all 60 participants were included.

Data for each of the first 9 regions of the sentence were entered into a 2 × 2 repeated-measures ANOVA with attractor number and grammaticality as factors. Regions of interest were 5-9; the attractor noun appeared in region 5, and the critical verb appeared in region 6. Statistics for each of these 5 regions are presented in Table 9.

Results

Comprehension Question Accuracy

Mean comprehension question accuracy was 91.8%. For grammatical conditions, the singular mean was 93.3% ± 1.3% compared to the plural mean of 89.7% ± 1.4%; for ungrammatical conditions, the singular mean was 89.7% ± 1.7% compared to the plural mean of 94.4% ± 1.2%. Neither main effect was reliable in a logistic mixed-effects model, but the interaction was reliable ($p < .05$).

Self-paced reading

As shown in Figure 6 (Panel A), regions 1-4 showed no significant effects. Region 5, the attractor region, showed a significant effect of number, but also a significant number × grammaticality interaction. Post-hoc comparisons revealed that only the grammatical conditions showed a reliable number effect (singular mean = 303 ms; plural mean = 335 ms; 95% CI = 12.8

ms; $p < .005$), but not the ungrammatical conditions (singular mean = 317 ms; plural mean = 320 ms; 95% CI = 11.2 ; $p > .1$). Because the attractor precedes the critical verb, either the presence of a number effect in one comparison or its absence in the other must be spurious. Experiments 1, 2 and 4 all showed a main effect of number on the noun itself.

In the critical verb region (R6) there was also a main effect of number, such that conditions containing a plural attractor were read more slowly. There was no reliable interaction of number and grammaticality. In the region following the verb (R7) there was a main effect of grammaticality (ungrammatical mean = 386 ms; grammatical mean = 317 ms; 95% CI = 11.4 ms; $p < .005$), as well as a significant grammaticality × attractor number interaction. The two subsequent regions (R8-9) showed the same pattern. The interaction was driven by a classic attraction pattern, specifically a smaller difference between the grammatical and ungrammatical conditions in the plural attractor conditions than in the singular attractor conditions. Pairwise comparisons at the region following the verb (R7) for the grammatical and ungrammatical conditions separately showed significant differences due to attractor number for the two ungrammatical conditions (singular mean = 416 ms; plural mean = 358 ms; 95% CI = 24.5; $p < .005$), but not for the two grammatical conditions (singular mean = 314 ms; plural mean = 321 ms; 95% CI: 10.1 ms; $p > .1$). In regions 8 and 9, the effect of attractor number in the two grammatical conditions was somewhat larger and marginally significant (R8: singular mean = 315 ms; plural mean = 329 ms; 95% CI = 11.4 ms; $p = 0.10$; R9: singular mean = 320 ms; plural mean = 332 ms; 95% CI = 10.9 ms; $p = 0.10$). In the ungrammatical conditions, however, there were no effects of attractor number in these regions (R8: singular mean = 357 ms; plural mean = 351 ms; 95% CI =14.5 ms; $p > .1$; R9: singular mean = 351 ms; plural mean = 336 ms; 95% CI = 12.4 ms; $p > .1$).

*** INSERT FIGURES 6 ***
*** INSERT TABLES 9 ***

Discussion

The results of Experiment 5 were largely similar to those of Experiment 4. Main effects of number appeared in the attractor noun region and the subsequent verb region. In other words, although the grammatical conditions showed a significant effect of attractor on the critical verb, this effect was already present in the preceding region. Once again, effects of grammaticality and an interaction between grammaticality and attractor number appeared in the region following the verb, and paired comparisons showed a significant effect of attractor number for ungrammatical conditions only. This pattern is consistent with previous results that suggest that attractor effects are largely delayed until the region following the verb, and that effects on the verb region itself likely reflect differences due to the increased complexity of plural nouns.

In contrast to Experiment 4, a small but significant effect of attractor number was seen for the grammatical pair in later regions of the sentence (Regions 8 and 9). Although this effect was not significant in Experiment 4, there was a numerical trend in the same direction. Combined, these data raise the possibility that grammatical sentences may indeed exhibit a real attraction effect, but that this effect is much smaller than the attractor effect for ungrammatical sentences and has a delayed time course. The small but significant effect of attraction and grammaticality on comprehension question accuracy is also consistent with the idea that an attraction effect for grammatical sentences may come online later in the sentence. At the same time, because the difference between the two grammatical conditions begins before the verb is encountered and remains at about the same size throughout the sentence, even the effects seen in the later regions may be due to baseline differences, if the plural cost observed stems from longer-lasting semantic/conceptual processes.

Previous region mixed-effect models analysis of Experiments 4 and 5

Experiments 4 and 5 showed that there was a large facilitation in reading times for ungrammatical sentences in the post-verbal regions when the sentence contained a plural attractor. In the same post-verbal regions there was no corresponding slow-down for grammatical sentences, counter to the predictions of a feature percolation account. However a slow-down was observed at the verb for these sentences. We argued however that this slow-down reflected increased difficulty associated with reading a plural that was sustained in subsequent regions. In the following analyses, we test this claim by estimating region-by-region mixed-effects models that incorporate the previous region reading times as factors. It is therefore possible to effectively remove the RT dependency between adjacent regions and thus estimate the RT effects independently introduced by the critical verbal region (Vasishth, 2006).

Procedure

Raw reading times, trimmed according to the criteria reported above, were entered into linear mixed-effects models. Separate models were estimated for each experiment and each region beginning with R3. In addition to the experimental factors (grammaticality, attractor number, and their interaction), the reading times at the two previous regions, RT_{n-1} and RT_{n-2}, were treated as fixed effects. Participants and items were treated as random effects. Previous region reading times were also nested under participants, which provided a better fitting model. By means of likelihood ratio tests (Baayen, Davidson, & Bates, 2008), we found that nesting the other experimental factors under either participants or items did not improve the fit of the model in the regions of interest.

Trials were removed if either RT_n, RT_{n-1} or RT_{n-2} were outside the 2.5 s.d. region-by-condition exclusion criterion. Consequently slightly more data were removed from this analysis. On average by region, 97% of the trials analyzed in Experiment 4 were entered into this analysis; and 96% of the trials analyzed in Experiment 5 were analyzed here. Excluding these additional trials was not found to alter the pattern of results or reliability reported in Experiments 4-5 for raw reading times (that is, prior to considering previous region RT).

Results

Results are reported in two ways. Firstly, the specific treatment contrasts below come from linear mixed-effect models incorporating the two experimental factors and previous region RTs. Secondly, the B Panels in Figures 5 and 6 provide a way of visualizing the specific impact of the previous region RT regression by plotting mean residual reading times of just a previous-regions RT model (that is, residualized without experimental factors). All RT_{n-1} and RT_{n-2} effects were significant, across regions and experiments. This is unsurprising since RTs within a sentence are expected to be correlated. RT_{n-1} and RT_{n-2} coefficients ranged from 0.17 to 0.77 (mean: 0.33; s.d.: 0.14).

In the Experiment 4 previous-regions RT analysis, a reliable slow-down for ungrammatical sentences occurred at the verb (R7; 13 ms; 95% CI = 10 ms, $p < .05$) and in the subsequent region (Region 8). This effect was largest in R8 (27 ms; 95% CI = 10 ms, $p < .001$). In the same region there was a reliable contrast between ungrammatical singular-attractor sentences and ungrammatical plural-attractor sentences, the latter being read significantly faster (31 ms; 95% CI = 18 ms, $p < .05$). There was a trend in the same direction already at the verb (Region 7), though it was not significant in the full model of the data. In a pairwise comparison

restricted to just the ungrammatical conditions, the effect is significant (22 ms; 95% CI = 16 ms, $p < .05$). Crucially, neither in the verb region (R7) nor in any post-verbal region was there a reliable effect of the attractor for grammatical sentences (R7: 3 ms; 95% CI =12 ms, n.s.; R8: -2 ms; 95% CI = 11 ms, n.s.). In the attractor region (R5) a reliable effect of number was detected, such that plural attractor sentences were read more slowly than singular attractor sentences (17 ms; 95% CI = 9 ms, $p < .05$). This effect was confined to R5 and was not observed on the adverb region (R6).

In the Experiment 5 analysis a reliable slow-down for ungrammatical sentences occurred in the immediate post-verbal region (R7; 63 ms; 95% CI = 10 ms, $p < .001$) and in the subsequent region (R8; 19 ms; 95% CI = 8 ms; $p < .005$). In R7, where the largest grammaticality effect occurred, there was a reliable contrast between ungrammatical singular-attractor sentences and ungrammatical plural-attractor sentences, the latter being read much faster (33 ms; 95% CI = 21 ms; $p < .05$). Once again there were no significant effects of a plural attractor in grammatical sentences in any verbal or post-verbal region. There was a marginal effect of the plural attractor in R7, such that plural-attractor grammatical sentences were read faster than singular-attractor grammatical sentences (-18 ms; 95% CI = 15 ms; $p < .10$). In the attractor region (R5), a reliable effect of number was observed (12 ms; 95% CI = 7 ms; $p < .05$), but as in the raw reading time analysis this effect was confined to the comparison between grammatical sentences (28 ms; 95% CI = 9 ms; $p < .005$).

Discussion

The previous region RT analysis preserved some patterns observed in the raw RT analysis. Specifically, in the verbal and post-verbal regions there was a robust slow-down for

ungrammatical sentences and an attenuation of this slow-down in ungrammatical sentences containing a plural attractor; in the attractor region there was a slow-down for plural attractors (in 3 of 4 pairwise comparisons). However, notably absent from this analysis was any indication in the verbal or post-verbal regions that grammatical sentences were read more slowly when they contained a plural attractor. The comparison between Figures 5 and 7 is helpful in determining the impact of incorporating previous region RTs in a model of region-by-region reading times. Both figures show that the slow-down first observed in the attractor region (R5) was present in both raw and residual datasets. In the raw reading times, this slow-down persisted to the adverb region (R6). However, once the correlations with previous region RTs were removed, no slow-down was evidenced in that region. Similarly no slow-down was present two regions from the attractor in grammatical conditions. Comparing Figures 6 and 8 shows the same pattern. There was no difference between grammatical conditions beyond the attractor region.

We conclude that reading the verb does not introduce any independent difficulty for grammatical sentences when there is also a plural attractor in the sentence. This conclusion runs counter to the symmetry prediction of accounts that attribute attraction to misvaluing of the subject number, like feature percolation.

Experiment 6

Since the presence or absence of agreement attraction in grammatical sentences has significant consequences for theories of how agreement is computed in comprehension, in Experiment 6 we conducted a further test for a grammatical attraction effect in self-paced reading for the prepositional modifier construction. In Experiment 6 we repeated the attraction manipulation of Experiment 4 (prepositional modifier constructions with the 'buffer' adverb) but in this study we tested only the grammatical conditions and correspondingly doubled the number

of items in each condition, reducing the within-subject variability. If agreement attraction can occur in grammatical sentences, then we expect to see a reading time slow-down in the plural attractor condition relative to the singular attractor condition.

Participants

Participants were 30 native speakers of English from the University of Maryland community with no history of language disorders.

Materials

Experimental materials consisted of 24 grammatical sentence pairs with singular subjects that varied in whether the attractor number matched or mismatched the number of the singular verb. Sentences were the same as those used in Experiment 4. The first seven words of each experimental item always followed the same sequence: *determiner-noun-preposition-determiner-noun-adverb-verb*. The 24 sets of 2 conditions were distributed across 2 lists in a Latin Square design, and were combined with 48 relative-clause agreement attraction sentences (half ungrammatical; described separately as Experiment 2), and 144 grammatical filler sentences of similar length. This resulted in 17% of the items being ungrammatical.

Procedure

The same self-paced reading procedure was used as in Experiment 1, and the analysis followed the same steps. Reading times that exceeded a threshold of 2.5 s.d. by region and condition were excluded. Two participants showed a comprehension question accuracy rate of less than 80% across all items and were thus excluded from further analysis.

Data for each of the first 10 regions of the sentence were entered into a repeated-measures ANOVA with attractor number as the factor of interest. Regions of interest were 5-10; the attractor noun appeared in region 5, and the critical verb appeared in region 7. Statistics for each

of these 6 regions are presented in Table 10.

Results

Comprehension Question Accuracy

Mean comprehension question accuracy was 92.0%. There was a main effect of attractor number (singular mean = 94.3% ± 1.2%; plural mean = 89.6% ± 2.0%; $p < .05$).

Self-paced reading

The results of Experiment 6 are presented in Figure 7. Regions 1-4 showed no significant differences (Fs < 1.7), although there was a marginal effect in region 2 in the participants analysis ($F_1(1,27) = 3.2$; $p < .10$). This effect must be spurious, since the materials were identical at this and preceding regions. Neither the attractor noun region (R5) nor the adverb region (R6) showed a main effect of number. However, a pooled analysis of regions 5-6 showed a reliable effect of attractor number (plural mean = 334 ms; singular mean = 319ms; 95% CI: 9.0 ms, $p < .05$). Crucially, neither the critical verb region (R7) nor any of the regions following the verb (R8-10) showed a significant difference between the singular and plural attractor conditions (R7 plural mean = 339 ms, singular mean = 338 ms, 95% CI = 12.0 ms; $p > .1$; maximal difference: R10 plural mean = 306 ms, singular mean = 300 ms, 95% CI = 8.1 ms; $p > .1$).

In order to increase the number of items per condition we did not test the corresponding ungrammatical sentences in Experiment 6, and thus there might be a concern that this particular group of participants happened not to be sensitive to attraction at all. However, this concern is mitigated by the fact that the relative-clause manipulation described in Experiment 2 was tested concurrently with the same participants in the same session, and there participants did show attraction effects in the ungrammatical relative-clause sentences.

*** INSERT FIGURE 7 ***
*** INSERT TABLE 10 ***

Discussion

In Experiment 6 we tested attraction in grammatical prepositional modifier constructions only and found no effects of attractor number on reading times at or after the critical verb. By itself, of course, this experiment is not sufficient to make a strong case against grammatical attraction, as it did not include the two ungrammatical attractor conditions to contrast with the null effect found in the grammatical conditions. However, when taken together with the previous four experiments these data provide strong evidence that grammatical sentences are not subject to attractor effects. The data from Experiment 6 replicate the results of Experiments 4 and 5 in which we also failed to find an effect of attractor number on reading times in the verb+1 region for grammatical attraction constructions of the same type, as well as the results of Experiments 2 and 3 in which we failed to find an effect of attractor number on reading times following the verb in grammatical relative-clause attraction constructions. The only evidence in favor of an online attraction effect in grammatical sentences was the significant effect seen in later regions (verb+2 and verb+3) in Experiment 4; however, this effect was later and much smaller than the attraction effect for ungrammatical sentences. Moreover the previous region RT analysis apportioned this effect to baseline differences in previous regions. The results of these five experiments thus converge to support the claim that agreement attraction either does not occur in the comprehension of grammatical sentences, or has a much different time course than attraction in ungrammatical sentences.

Although the effect of number at the attractor noun and the regions following the noun in Experiment 6 seemed to be weaker than in previous experiments, the plural conditions still showed a tendency for longer reading times in these regions, and the effect of number was significant when Regions 5 and 6 were combined. Thus, like the previous five experiments

presented, this experiment provides evidence that noun number can exert small but reliable effects on reading times both on the noun itself and on the subsequent word, although the mechanisms underlying these effects are currently unclear.

Experiment 7

In Experiments 4 and 5 the additional processing complexity due to the plural attractor itself made it more difficult to determine whether slow-downs in the grammatical condition were due to a number effect or an attraction effect. Even when an adverb was inserted between the attractor and the verb, a main effect of noun number still extended as far as the verb region. In contrast to the RC experiments, in which the plural attractor was far from the verb and we were able to clearly show the absence of a grammatical attraction effect, it was more difficult in the PP-modifier experiments to unequivocally conclude from the reading time data that there was no grammatical attraction effect. In Experiment 7 we used an end-of-sentence judgment task to try to reduce the impact of number complexity on our measure. Agreement attraction should have the effect of increasing or decreasing sentence acceptability, but the processing complexity associated with plural nouns should have very little effect on acceptability. Furthermore, several previous studies have reported grammatical attraction effects in offline measures (Nicol, Forster & Veres, 1997; Häussler & Bader, submitted), which raises the possibility that agreement attraction does occur in grammatical sentences, but that its effects are for some reason delayed relative to ungrammatical sentences, and thus might be better captured with offline measures. Experiment 7 allowed us to examine this alternative possibility.

Participants

Participants were 16 native speakers of English from the University of Maryland community with no history of language disorders.

Materials

Experimental materials consisted of 24 grammatical sentence quadruplets using the prepositional modifier construction with singular subjects that varied in whether the attractor number matched or mismatched the number of the singular verb. Sentences were the same as those used in Experiment 5 (i.e., with no adverb between the attractor and the verb). The 24 sets of 4 conditions were distributed across 4 lists in a Latin Square design and were combined with 48 relative-clause agreement attraction sentences from a different experiment not reported here (half ungrammatical), 24 wh-dependency sentences from a different experiment (half ungrammatical) and 32 filler sentences of other types (half ungrammatical), for a total of 128 items.

Procedure

Sentences were presented one word at a time in the center of the screen with rapid serial visual presentation reading (RSVP; Potter, 1988) at a rate of 300 ms per word. At the end of each sentence, a response screen appeared for 2 seconds in which participants made a yes/no response by button press. Participants were instructed to read the sentences carefully and to judge whether they sounded like an acceptable sentence that a speaker of English might say. Participants were instructed to make their response as quickly as possible following the end of the sentence; if they waited longer than 2 seconds to respond, they were given feedback that their response was too slow.

Results

Figure 8 shows the proportion of 'yes' judgments for the four experimental conditions in Experiment 7. The results of a 2 × 2 logistic mixed-effects model confirm what the figure suggests: there was a main effect of grammaticality (model coefficient, in logits, β: -4.0 ± 1.0, $p < .001$) and an interaction of grammaticality with attractor number (β: 1.7 ± 1.3, $p < .01$). Participants were more likely to accept ungrammatical sentences when an attractor was present: acceptance rates more than doubled (N_{SG}-N_{SG} ungrammatical: 25% ± 7%; N_{SG}-N_{PL} ungrammatical: 55 ± 6%). However, participants were not more likely to reject grammatical sentences, when there was an attractor (N_{SG}-N_{SG} grammatical: 93% ± 3.0%; N_{SG}-N_{PL} grammatical: 91% ± 3.7%; β: -0.26 ± 1.1, n.s.).

*** INSERT FIGURE 8 ***

Discussion

The results of Experiment 7 corroborate the results of Experiments 4-6. Consistent with the self-paced reading results, ungrammatical sentences containing a mismatch between a singular subject and a plural verb were more likely to be accepted in the presence of a plural attractor, and were judged acceptable 55% of the time, compared to 25% when a singular attractor was present. If attraction reflected errors in the encoding of subject number, then in grammatical sentences with plural attractors the singular verb would sometimes be perceived as mismatching the plural subject, and overall acceptability ratings would fall. However, no attractor effect was observed in acceptability judgments to the grammatical conditions. These data support our argument that the slow-downs observed immediately following the plural attractor in the self-paced reading

experiments were due to increased processing cost of the plural and not to perceived subject-verb mismatch.

One possible alternative account of the self-paced reading experiments on the prepositional modifier construction presented above would be that attraction effects appear earlier in grammatical sentences than in ungrammatical sentences. This would explain why only grammatical sentences showed an effect of attractor number at the critical verb region. However, although the timecourse of attraction is different for grammatical and ungrammatical sentences under this account, attraction should nevertheless be evident to the same extent for both in offline measures. Thus, the results of Experiment 7 provide evidence against this alternative.

General Discussion

The results from the six subject-verb agreement attraction experiments presented here provide two types of evidence against accounts of agreement attraction in comprehension that rely on percolation, in particular, and on faulty representations of the subject, in general. In the two relative-clause experiments (Experiments 2 and 3), we found that even when a noun does not intervene between the subject and the verb and is in fact hierarchically superior to the subject noun, it can still exert a strong attraction effect, reducing or even eliminating reading time disruption due to an ungrammatical subject-verb mismatch. Furthermore, across all five self-paced reading attraction experiments we found no evidence for a corresponding attraction effect in grammatical sentences in the same time window. Although the interpretation of reading times is complicated in the prepositional modifier construction by simple effects of noun number, we found that the effects observed on the verb could be completely explained by reading times on the previous regions. We found the same grammatical-ungrammatical attraction asymmetry in an

RSVP acceptability judgment task, and we find no trace of an attraction effect on reading times in the grammatical relative clause cases, which do not suffer from the same confound. In the following we first review the implications for previous accounts of each type of evidence and then we discuss alternative accounts that could better capture the pattern of data we observe.

Attraction in a non-intervening configuration

Our results demonstrate that a seeming quirk of agreement, first pointed out by Kimball and Aissen (1971; e.g., "The people who Clark think are in the garden ..."), is reflected in moment-to-moment processing. Kimball and Aissen speculated that the ability of a plural relative clause head to agree with the verb embedded within the relative clause was a matter of dialect, and, specifically, a dialect localized to the Northeastern US or Boston area. The one subsequent study to examine acceptability ratings for the Kimball and Aissen paradigm also likely drew largely on participants from this region (Clifton, Frazier & Deevy, 1999). However, in our study, participants largely came from the Chesapeake Bay region and surrounding Mid-Atlantic states. It therefore seems less likely that agreement with the head of a relative clause is a fact about a regional dialect. Given how closely its processing profile matches that of the canonical complex subjects that induce agreement attraction, our results favor the view that the facts Kimball and Aissen brought to our attention, and which have subsequently been analyzed by Kayne (1989) and den Dikken (2001), reflect properties of the real-time processing architecture rather than a structurally-conditioned grammatical process.

The finding that the number of the relative clause head reduces the subject-verb mismatch effect within the relative clause is notable for several reasons. First, the subject-verb mismatch involves a linearly adjacent subject and verb (*...that the reviewer praise...*), and one might

expect this configuration to make the mismatch so obvious that other parts of the sentence could have no opportunity to have any ameliorating effects. Second, the relative clause head is not within the immediate structural domain of the subject NP, so feature movement from one to the other, as hypothesized by the feature percolation model, would have to be indirect. Third, despite the fact that the relative clause head attractor was structurally and linearly further from the verb than the attractor in the prepositional modifier paradigm, the reduction of the ungrammaticality effect was of approximately the same magnitude (see Table 11).

Although our demonstration of attraction in this configuration provides important evidence against the percolation model, it does not argue conclusively against models that explain attraction through faulty representation of the subject. A feature percolation model might still be able to explain these findings if the parameters were set correctly; for example, if features could percolate downwards from a number source, and if the distance between the attractor and the subject and between the subject and the verb were weighted fairly low, the differences in distance between the two constructions would not predict large differences in the magnitude of the RT effect. The challenge in pursuing this approach is to make sure that distance is still weighted high enough that it can account for the structural distance effects found in production studies, one of the key motivations for feature percolation models in the first place.

In the case of relative clause processing, there is another potential source for feature confusion, which stems from the active completion of the filler-gap dependency that is built between the phrase containing the relative clause head and relative clause verb (Stowe, 1986; Frazier & Flores D'Arcais, 1989; Traxler & Pickering, 1996). Attraction inside relative clauses could thus be brought about by a different mechanism from attraction in prepositional modifier constructions. One might suppose that the relative clause head is reactivated at the RC verb

position in the same time frame as agreement is checked, and that features that are simultaneously co-active are liable to interference, or confusion (cf. Clifton, Frazier, Deevy, 1999; Franck et al., 2006). This view shares the crucial property of the standard feature percolation accounts (developed for complex subject NPs) in predicting symmetrical effects of attraction for grammatical and ungrammatical agreement alike, which we failed to find. An alternative means of capturing these facts, which we discuss in more detail below, is to assume that the attractor exerts its effect by means of a cue-directed retrieval in content-addressable memory.

Timing of agreement effects in self-paced reading

In the six self-paced reading experiments presented above, effects of subject-verb mismatch and attraction sometimes appeared on the verb itself and sometimes on the following word, even when reading times were corrected for the residual effect of reading times on previous regions. In Experiments 1 and 4, effects of mismatch and attraction were observed in reading times on the verb itself, while in Experiments 2, 3, and 5, effects of mismatch and attraction were observed in reading times only on the regions following the verb. One property that may be responsible for this variability in timing is the inclusion of an adverb in the region preceding the verb: Experiments 1 and 4 were the two experiments in which we included an adverb region to serve as an index of effects due to the number of the preceding noun. The pattern we observe suggests that the inclusion of the adverb also may have had the unexpected effect of causing effects of verb number to show up earlier than otherwise. We hypothesize that the occurrence of an adverb may strongly increase expectation for the upcoming verb, such that relationships that depend on the verb can be processed more quickly when the verb is actually encountered. Thus, although including an adverb region can be a useful tool for identifying the existence of effects prior to the verb, our results suggest that the presence of an adverb can exert effects of its own, which researchers must be aware of.

Lack of attraction in grammatical sentences

The lack of attraction in grammatical sentences is problematic for all models that claim that attraction in comprehension is due to changes in the structural representation of the subject. As discussed in the Introduction, any account in which features from other nouns can spread to the head noun or subject node predicts that this faulty percolation should occur in grammatical and ungrammatical sentences equally, because the percolation mechanism applies independently of the properties of the verb. The same prediction is shared by any syntactic account that assumes that the attractor noun makes itself visible to the agreement system independently of the verb; for example, Den Dikken (2001) proposes that in complex subjects a noun phrase contained within the subject moves to a position that c-commands the subject as a species of generalized quantifier raising. However, in all six attraction experiments, we never found an attraction effect in the grammatical sentences that was comparable to the ungrammatical sentences. Table 11 directly compares the effects of subject-verb grammaticality and attraction across all the self-paced reading experiments and shows the same asymmetrical pattern, with fairly similar values, in each.

*** INSERT TABLE 11 HERE ***

Pearlmutter and colleagues (1999) did show attraction effects for grammatical sentences in three experiments using materials that were almost identical to our Experiment 5. Their Experiment 1 using self-paced reading showed a pattern similar to our Experiment 5: they found the predicted symmetrical attraction effect (ungrammatical faster with a plural attractor, grammatical slower with a plural attractor) only in the region following the verb, but in the verb region itself they found a main effect of local noun number (plurals read more slowly) in both ungrammatical and grammatical sentences. Thus, although the grammatical conditions continued

to differ in the regions following the verb, the presence of the same difference in the preceding region provides a basis for the alternative interpretation that this difference represents an extended number effect. Pearlmutter et al.'s Experiment 2 used the same materials but with eyetracking fixation times as the dependent measure. In the first pass measures they found a more puzzling pattern. There were no effects in the verb region, in the first post-verbal region reading times were much faster in the grammatical singular attractor condition than in the other three conditions, and only in the second post-verbal region did a small attraction effect for ungrammatical sentences appear. These results are hard to compare with our own, since we consistently find large attraction effects in the first post-verbal region in ungrammatical sentences, but one possibility is that the unexpected pattern in their first post-verbal region reflects an interaction between the cost of processing plurals and the reduction of disruption due to ungrammaticality in the ungrammatical attraction condition. The results of their Experiment 3, which tested only grammatical sentences, are the most difficult to attribute to a plural complexity effect, since the critical verb region showed an effect of attractor number in only one of the experimental pairs. However, other aspects of the Experiment 3 results are also hard to explain under any traditional agreement account, as they also found a reverse attractor number effect for plural subject sentences, which are not normally subject to any form of agreement attraction (Pearlmutter, 2000, shows a similar pattern).

Several studies have examined agreement processing in comprehension using offline measures and have found reliable attractor effects in grammatical sentence pairs (e.g. Nicol, Forster, & Veres, 1997; Häussler & Bader, submitted). Although we failed to show a grammatical attraction effect in our speeded acceptability judgment task, in our Experiments 4 and 5 comprehension question accuracy did show a significant attractor effect in grammatical

sentences, although the effect was small. While it seems less likely that the increased complexity of plurals would result in processing costs that extend to the end of the sentence, it is possible that late processes like semantic integration of the subject and the predicate might be sensitive to pure number effects (although a control experiment by Nicol et al., 1997 suggests that noun number could not have been solely responsible for their results). Another possibility is that some offline tasks might engender some sort of late reprocessing or regeneration of the surface form of the sentence, with the result that both grammatical and ungrammatical agreement might become susceptible to the same kind of interference discussed above for production (cf. Potter & Lombardi, 1992).

Although we cannot currently provide a complete account for the previous experiments that have shown online grammatical attraction effects, we were unable to replicate these effects across five self-paced reading experiments using two separate constructions, and we would argue that the cost of processing plural nouns per se that we repeatedly observed in these experiments makes it difficult to unambiguously interpret the existing findings of grammatical attractor number effects in prepositional-modifier constructions. In comparison, the reduction of disruption in ungrammatical sentences is robust and relatively consistent in its time course across different studies and configurations. Although further replication is needed, we argue that this contrast already provides a strong argument against theories that predict symmetrical attraction effects in grammatical and ungrammatical sentences.

Retrieval-based mechanisms for attraction in comprehension

Based on the two lines of evidence outlined above, we have argued that agreement attraction effects in comprehension are not driven by a property of the structural representation of number

in the subject itself. We propose that such effects are due to properties of a content-addressable retrieval mechanism that is initiated by the verb. The fact that only sentences in which subject and verb number mismatched showed a reliable attractor effect suggests that information from the verb (or auxiliary) plays a necessary role in attraction effects online. A content-addressable-retrieval mechanism that uses the information on the verb could naturally give rise to the observed pattern of attraction (Gillund & Shiffrin 1984; Lewis & Vasishth 2005). The core idea behind such a retrieval mechanism is that features in the current input are used as cues to query the contents of memory simultaneously, much like keywords are used in Internet search engines to find matching websites. The results of such a query could match all keywords or only a subset of them, and this degree of match may affect the likelihood of retrieving the results from memory.

Retrieval processes could play a role in agreement processing in two different ways: 1) normal agreement processing could be instantiated by a cue-based retrieval process in which verb number is used as a cue to search for the subject in memory; 2) a mismatch between the verb and subject number could initiate a reanalysis process that makes use of cue-based retrieval to find a matching antecedent to fix the mismatch problem. In other words, the subject's number feature is either always retrieved during verb processing, or it is triggered by an error signal. We discuss each of these possibilities in turn.

One way to capture the pattern of effects observed in these comprehension experiments is to assume that subject-verb agreement is always computed through a cue-based retrieval process that is engaged when the verb is encountered (Lewis & Vasishth, 2005; McElree, 2006). Numerous sources of information provided by the verb may form the cues for retrieval. For the configurations we have considered, we assume that the retrieval cues consist of (privatively

specified) agreement features, like [Number:Pl], structural cues, like [Case:Nominative] or [Role:Subject], and clause-bounding cues. The crucial property that is needed for this kind of model to capture the ungrammatical-grammatical asymmetry in attraction effects is that a partially-matching NP is (almost) never retrieved if a fully-matching NP is present. If cue combination rules are supra-linear, a standard feature of content-addressable memory models (Gillund & Shiffrin, 1984; Hintzmann, 1988, inter alia), then full matching constituents will strongly out-compete partial matches. Alternately, the system may ignore weakly selective superordinate cues like [Category:NP], which apply to multiple constituents. In this way the system would almost always retrieve the correct, number-matching subject in the grammatical sentences, so there is little chance of retrieving another non-number-matching NP in the wrong structural position. On the other hand, when neither of the NPs matches the combined cue, as in the ungrammatical sentences, the number-matching non-subject is sometimes the best match. Wagers (2008) provides an extensive discussion of the possible cue structures that can capture the pattern of results we observe.

The other possibility is that the attractor effect in comprehension is specifically a reanalysis effect. On this view, agreement computation is always carried out correctly on the first-pass, but when this computation fails a reanalysis mechanism can check back to see if an error was made. The initial computation could be instantiated as a predictive process: when a subject NP is encountered, a verb marked with the correct agreement features can immediately be predicted. When the verb is encountered its number features can be checked against the predicted features, and if they match, nothing more needs to be done; in particular, there is no need to retrieve material from the prior context. However, if the bottom-up features of the verb mismatch the top-down predicted features, then a cue-based-retrieval may be deployed to check whether the

correct feature was somehow missed in the syntactic context during the first pass. It is in this 'rechecking' stage that the attractor NPs might sometimes be mistakenly retrieved. The fact that attractor effects are mainly seen in ungrammatical sentences is thus a natural consequence of this view.

The data presented here do not choose between these options and each has benefits and drawbacks. Either account relies on a cue-based retrieval for the number information in the subject. The accounts diverge on the issue of whether that retrieval occurs whenever the verb is processed or if the retrieval is limited to cases where the expectation for the verb's number marked is unmet. On the one hand, there are several reasons to favor an account that limits retrievals to instances of reanalysis. First, it makes retrieval of the number feature unnecessary in the normal case. Second, given that English agreement paradigms for lexical verbs are largely syncretic, it may be necessary in the first place to use top-down information, like the number of the subject head, to identify the number features of the verb. Finally, McElree, Foraker, & Dyer (2003) concluded on the basis of speed-accuracy tradeoff time-course dynamics that adjacent subjects and verbs can simultaneously occupy the focus of attention. As a consequence, they can be integrated without necessitating a retrieval. We observe in Experiment 2-3 that adjacent subjects and verbs are nonetheless liable to attraction when subject and verb mismatch. The existence of agreement attraction effects in adjacent subject-verb sequences can be reconciled with the focus of attention findings if no number retrieval is required unless there is an error signal. On the other hand, there are potential costs involved in making the prediction assumed in the alternative account, depending on the implementation. For example, it may not be possible to maintain the expected verb feature while intervening material is being processed (e.g., while the PP is being attached in N-PP structures), such that a retrieval should usually be necessary.

Finally, it is worth noting that the two alternative accounts can be viewed on a continuum: whether top-down versus bottom-up processes predominate may depend on precise characteristics of the constituents and structures involved. Further research will be necessary to disentangle the interaction between prediction and retrieval.

The scope of interference

We have offered a unified analysis of PP and relative clause agreement attraction in terms of retrieval-based interference. This unification is based on similar patterns of attraction involving singular/plural markedness and also the grammatical/ungrammatical asymmetry. However, there may still be differences between the two structures. Bock & Miller (1991) found that in a production study involving relative clause preambles the majority of errors occurred for singular embedded subjects when the relative clause head was plural and, crucially, animate. This finding contrasted with N-PP preambles, in which the animacy of the local noun did not affect rates of attraction. This production contrast might motivate distinguishing the two constructions, such that comprehenders are genuinely confused about the identity of the subject in relative clause attraction but not in PP attraction. Our findings resist this distinction in one respect: the fact that attraction only occurs for ungrammatical strings implies that comprehenders are not generally uncertain about the relative clause subject. If they were, the pattern of RTs would be similar to that predicted by the percolation account. In a similar vein, Wagers (2008) found that relative clause attraction in ungrammatical strings was present both when the relative clause was attached to a main clause subject and when it was attached to a main clause object. This finding provides further evidence that attraction in comprehension is not contingent upon how 'subject-like' the attractor is.

However it seems reasonable that comprehenders could become uncertain about the identity of the subject after detecting a subject-verb mismatch. To put this more generally, satisfying the agreement requirements of the verb with a grammatically-inappropriate head could lead to realignment of the grammatical relations closely related to agreement, like subject-hood or conceptual number. In some recent research, our group has found no evidence that attraction leads to reassignment of grammatical roles (Lau, Wagers, Stroud & Phillips, 2008) or conceptual number (Lau, Wagers, & Phillips, 2009) for N-PP attraction in comprehension. However the effects of attraction on the interpretation may be more subtle. We believe that investigating this question is an important future area of research, not only to provide a better account of agreement attraction but also to provide a better understanding of the role retrieval plays in assembling combinatorial representations. In particular, determining whether or not it is possible to retrieve and process some features independently of others could shed light on the granularity of the constituent encodings that support comprehension.

Production versus Comprehension

What is to be made of the link between comprehension and production? One attractive aspect of the feature percolation model is that it can be easily stated independently of the task, because the error derives from the syntactic encoding itself. Consequently, the explanation for agreement attraction in comprehension and production is fundamentally the same. It is less clear that a model based on retrieval has the same extensibility, but we would like to argue that it does. Following ideas in Solomon and Pearlmutter (2004), we suggest that the key element that speech planning and retrieval in comprehension share is that they are sequence management tasks in which elements may be simultaneously active. In Solomon & Pearlmutter (2004)'s model of

agreement in production, what drives the attraction effect is the simultaneity of the head nouns during the production planning process. The attractor can wrest away control of agreement by being co-active with the grammatical controller when conceptual structure is mapped onto syntactic structure. Because both heads are simultaneously accessible, the process of selecting the correct verb form sometimes spuriously pays attention to the attractor's features. The more accessible the attractor, the more likely it is to be spuriously selected. In their model simultaneity is determined by the tightness of the semantic relation between heads: roughly, the degree to which the attracting head characterizes the subject head. In the comprehension model proposed here, simultaneity again plays a role. Because retrieval of constituents is cue-driven and not serial, there is a possibility that multiple constituents will be rendered accessible to comprehension processes. The choice of which set of constituents is made accessible depends on the match between the retrieval cues and constituent features.

Recent theoretical work by Badecker and Lewis in the ACT-R framework (2007), and recent empirical work by Badecker & Kuminiak (2007) on gender attraction in Slovak, has sought to explain production errors of agreement in nearly exactly the same fashion that we are using to explain the comprehension errors: attraction arises when cue competition leads to a partial match. The production process is posited to take place in the same kind of architecture as comprehension, one in which most operations must retrieve information from recent memory. Verb marking occurs after the subject has been constructed, and it must retrieve the subject to inspect its number properties. Both the subject phrase and the embedded attractor share a category feature, as well as a nominative case feature (in the planning process). The whole subject is distinguished from the embedded attractor because its dominating category is also encoded. Therefore a set of cues exists that will converge upon the whole subject, but will also

partially activate the local noun. Consequently the production system retrieves the whole subject in the majority of cases. Embedded plurals lead to attraction to a greater extent than embedded singulars because the cue structure includes a variable value cue: NUM:*var*. The system is thus biased to return explicitly number marked constituents. The proposal by Badecker and Lewis is compatible with either of our accounts for comprehension, because both of our accounts attribute the attraction effect to similarity-based interference grounded in the feature composition of related constituents. Although there is no clear analog of reanalysis in production, in our view reanalysis is simply one of many factors that govern the timing of the retrieval operations that are common to both production and comprehension.

In either Solomon & Pearlmutter (2004), Badecker & Lewis (2007), or our own account of comprehension, the encoding of nominal number is assumed to be accurate but the process of accessing a number feature is error-prone. The question then arises whether the structural distance effects found in production have an analogue in comprehension, and if so, what account can be given. Pearlmutter (2000) provides some evidence for distance effects in comprehension from self-paced reading experiments involving stacked PP modifiers. The evidence is somewhat equivocal, as those results do not replicate the plural markedness effect and are subject to our more general concerns over the processing costs of plurals. However, assuming that hierarchical effects do hold in comprehension, we propose the explanation must be found in patterns of access. In production, Badecker and Lewis provide one mechanism for hierarchically modulating access to constituents in terms of activation: hierarchically more dominant categories have higher base rates of activation, because they have undergone more processing (i.e., more modification). Higher rates of activation are proposed in their framework, ACT-R, to lead to higher retrieval likelihood. It is possible to posit a similar mechanism in comprehension (Lewis & Vasishth,

2005). Some caution must be observed since modulation of activation, as a function of processing depth, will only heuristically and imperfectly recapitulate hierarchical order. However at present we would simply emphasize the need for more work in comprehension, in order to better sample the logical space of subject head and attractor configurations.

Plural complexity effects in self-paced reading

Our reading time data add to a growing body of evidence that factors known to impact lexical decision times – in this case, the effect of noun number – are also relevant in the self-paced reading technique (Niswander, Pollatsek, & Rayner, 2000; Bertram, Hyönä, & Laine, 2000; Lau, Rozanova, & Phillips, 2007). In Table 12 the effect of reading the plural attractor is compared to its singular counterpart in each of the six experiments, analyzed one region downstream. There is a consistent elevation of reading times in the plural condition. The size of the effects varies across the different experiments, which may be due to differences in structure, ordinal position, and participants. In particular, the effect in Experiment 1 was much larger numerically than in the other experiments and previous self-paced reading studies, perhaps because of the increased variability and longer reading times seen for this study, or perhaps because of the increased anticipation for the noun created by the prenominal adjective included in these materials. However, within the experiments that used the same structural environment (i.e., relative clauses in 2-3; complex subjects in 4-6), the effect size was fairly consistent. More work is needed to determine whether the plural cost is due to morphological processing, semantic integration, length, or some other factor.

It is important to note that controlling for length by means of a residual reading time analysis is insufficient to control for increased plural reading times. First, as Table 12 demonstrates, the

effect is persistent and shows up in reading times on the subsequent word, which are otherwise matched for average word length. Second, the size of the effect is larger than the typical millisecond-per-character slopes that we observed in by-participant regressions in this data set and across a broad range of unrelated studies — namely around 3 to 7 ms per character. Since the length difference between singulars and plurals is only one character for regular nouns, correction of reading times based on a word length regression should shrink the difference between the two forms by only 3 to 7 ms. In contrast, the previous-region-RT covariate analysis we presented for Experiments 4 and 5 seems to provide a satisfactory means of controlling for extended complexity effects (cf. Vasishth, 2006).

*** INSERT TABLE 12 HERE ***

Conclusion

We have demonstrated properties of agreement attraction in comprehension that are inconsistent with accounts in which attraction effects are driven by the erroneous representation of subject number. A relative clause configuration in which the attractor was hierarchically superior to the subject and did not intervene between subject and verb showed strong attraction effects. Number marking on non-subject nouns had a large and robust impact on processing ungrammatical agreement, but no effect on processing grammatical agreement. These findings suggest that the representation of subject number is not influenced by the mere presence of nearby nouns with conflicting number. We conclude that attraction errors reflect how structure-building processes gain access to information during the course of comprehension or production and, in particular, are consistent with a cue-based retrieval mechanism for accessing and comparing previously processed constituents.

Acknowledgments

We would like to thank the members of the Department of Linguistics and the Cognitive Neuroscience of Language Laboratory for their helpful feedback on this project. We thank Norbert Hornstein and Carson Schütze for valuable discussions. We thank Katya Rozanova and the members of the Undergraduate Psycholinguistics Seminar,Fall 2006, for assistance in running the experiments. EFL was supported by an NSF Graduate Research Fellowship. CP was supported in part by NSF #DGE-0801465.

Supplementary data

Appendix A: Experimental Materials

Experiment 1

The experiment contained 24 sets of 4 conditions.

Conditions were as follows:

Condition a:	singular subject, grammatical
Condition b:	plural subject, grammatical
Condition c:	singular subject, ungrammatical
Condition d:	plural subject, ungrammatical

An example of a full set:

1a. The new slogan unsurprisingly was designed to get attention for the company's signature designer bath products.
1b. The new slogans unsurprisingly were designed to get attention for the company's signature designer bath products.
1c. The new slogan unsurprisingly were designed to get attention for the company's signature designer bath products.
1d. The new slogans unsurprisingly was designed to get attention for the company's signature designer bath products.

2. The wedding picture(s) definitely was taken at a village church in the south of France, and there seemed to be herds of sheep in the background.
3. The programming mistake(s) certainly was disastrous for the small software company, although their competitors couldn't have been happier.
4. The red label(s) probably was meant to serve as a warning about the hazardous chemicals that were being stored inside the vault.
5. The morale problem(s) ultimately was solved by firing the custodian, who never came in on time and was always complaining about the workload.
6. The product name(s) probably was tested on a number of focus groups before being selected by the prominent advertising firm.
7. The shocking crime(s) doubtlessly was a reflection of both the desperation of the population and the failure of law and order in the wartorn country.
8. The major defect(s) likely was unknown to consumers and government regulators, although the company executives had been aware of the issue for years.
9. The erroneous citation(s) unfortunately was responsible for many wasted hours in the library, but the research assistant eventually realized the error.
10. The enormous advertisement(s) typically was visible from almost a mile away, but when it was raining it was harder to see.
11. The young traitor(s) regrettably was passing information to the enemy through various secret methods, one of which involved hiding notes in the dirty laundry.
12. The logical flaw(s) usually was clear to everyone who heard the theory, but the young students were more influenceable and had a misguided faith in their professor.
13. The back door(s) accidentally was left unlocked by the cleaning service, and by the next morning all of the expensive equipment had gone missing.
14. The urgent memo(s) unfortunately was about the delinquent tax return, which would soon be investigated by the Internal Revenue Service.

15. The bad check(s) apparently was signed by a well-known con artist who made his living by convincing old ladies to invest in his business enterprises.
16. The old key(s) unsurprisingly was rusty from many years of disuse, but it would still open the door to the attic if you turned it just right.
17. The important letter(s) allegedly was received in San Francisco in late March, but the prosecutor argued that it had been received three weeks earlier.
18. The beautiful entrance(s) evidently was hard to clean because of all of the delicate carvings and intricate lettering.
19. The sudden warning(s) unsurprisingly was a major shock to the residents, because they had always been told that their city was one of the safest in the world.
20. The small bridge(s) apparently was about half an hour off the main highway, but it was the only opportunity to cross the river for a hundred miles.
21. The unexpected gift(s) apparently was from a charitable organization, but unbeknownst to the recipient, thousands of dollars worth of illegal drugs were hidden inside.
22. The winding path(s) happily was lined with lovely trees and flowers, which were cared for by a host of well-trained gardeners.
23. The recent message(s) unfortunately was about several serious crimes which had occurred off-campus and whose perpetrators had not been apprehended.
24. The massive gate(s) evidently was falling down due to disrepair, and the historical society didn't hold out much hope that it could be saved.

Experiments 2 and 3

The experiments contained 48 sets of 4 (Exp. 2) or 8 (Exp. 3) conditions. Conditions a-d were included in both experiments, while conditions e-g were included in Experiment 3 only.

Conditions were as follows:

Condition a:	singular subject, singular attractor, grammatical
Condition b:	singular subject, plural attractor, grammatical
Condition c:	singular subject, singular attractor, ungrammatical
Condition d:	singular subject, plural attractor, ungrammatical
Condition e:	plural subject, singular attractor, grammatical
Condition f:	plural subject, plural attractor, grammatical
Condition g:	plural subject, singular attractor, ungrammatical
Condition h:	plural subject, plural attractor, ungrammatical

An example of a full set:

1a. The runner who the driver sees during the commute every morning always waves to say hi.
1b. The runners who the driver sees during the commute every morning always wave to say hi.
1c. The runner who the driver see during the commute every morning always waves to say hi.
1d. The runners who the driver see during the commute every morning always wave to say hi.
1e. The runner who the drivers see during the commute every morning always waves to say hi.
1f. The runners who the drivers see during the commute every morning always waves to say hi.
1g. The runner who the drivers sees during the commute every morning always waves to say hi.
1h. The runners who the drivers sees during the commute every morning always waves to say hi.

2. The industrialist(s) who the activist(s) criticize(s) always put(s) profit above concern for the environment.
3. The player(s) who the coach(es) like(s) the best is/are the one who always gets to practice on time.
4. The volunteer(s) who the director(s) relies on completely is/are dedicated to her work.
5. The author(s) who the librarian(s) recommend(s) for summer reading is/are very popular in India.
6. The pilot(s) who the smuggler(s) employ(s) in Brazil must be discreet and good at thinking fast.

7. The colleague(s) who the researcher(s) spend(s) the most time with come(s) up with crazy but inventive ideas.
8. The patient(s) who the doctor(s) check(s) on daily is/are undergoing an experimental treatment.
9. The Marine(s) who the officer(s) want(s) to promote is/are always ready to risk his/their life/lives for someone in need.
10. The sculptor(s) who the donor(s) support(s) so generously has/have held several acclaimed exhibitions.
11. The caterer(s) who the hostess(es) suggest(s) to her/their friends is/are excellent but outrageously expensive.
12. The senator(s) who the voter(s) listen(s) to patiently at all the local political events just never seem(s) trustworthy.
13. The teenager(s) who the farmer(s) hire(s) to pick fruit every summer usually work(s) pretty hard.
14. The goalie(s) who the fan(s) hold(s) the greatest confidence in is/are reliably calm under pressure.
15. The quarterback(s) who the recruiter(s) watch(es) from the bleachers every weekend might win a lucrative contract.
16. The publicist(s) who the politician(s) use(s) during a crisis is/are expected to do damage control.
17. The tenant(s) who the landlord(s) strongly disapprove(s) of stay(s) out late on weekends and never mow(s) his/their lawn.
18. The skateboarder(s) who the kid(s) admire(s) passionately always sign(s) autographs for fans.
19. The chef(s) who the gourmet(s) visit(s) regularly always come(s) up with daring new recipes.
20. The firefighter(s) who the ranger(s) keep(s) on hand during the summer months has/have special training in forest fires.
21. The overseer(s) who the worker(s) distrust(s) the most has/have been firing anyone who mentions unions.
22. The heiress(es) who the bachelor(s) pursue(s) so relentlessly is/are clever and won't be easily fooled.
23. The actor(s) who the producer(s) consider(s) the safest bet is/are the one who has/have been around for a while.
24. The accountant(s) who the administrator(s) depend(s) on to balance the books is/are very well paid.
25. The orphan(s) who the nun(s) tutor(s) in algebra lost his/their parents in the ongoing civil war.
26. The translator(s) who the diplomat(s) need(s) for communication was/were trying furiously to facilitate peace talks.
27. The dancer(s) who the host(s) applaud(s) so frequently has/have been performing since she/they was/were (a) young girl(s) in Ecuador.
28. The supervisor(s) who the counselor(s) report(s) to directly is/are supposed to provide advice and support.
29. The commentator(s) who the viewer(s) trust(s) the most has/have a very slick hairstyle.
30. The journalist(s) who the editor(s) appreciate(s) the most always turn(s) her stories in promptly.
31. The customer(s) who the waitress(es) hate(s) the most is/are the one who complain(s) about everything and leave(s) a lousy tip.
32. The therapist(s) who the survivor(s) meet(s) with weekly is/are trying to help her/them recover from her/their awful ordeal.
33. The cheerleader(s) who the choreographer(s) work(s) with usually train(s) all year for the national competition.
34. The cowboy(s) who the rancher(s) know(s) the best has/have thirty years of experience on the range.
35. The surgeon(s) who the nurse(s) prefer(s) most to assist has/have an impeccable record in the operating room.
36. The bartender(s) who the patron(s) gossip(s) with after work always has/have new stories to tell.
37. The columnist(s) who the reader(s) complain(s) about frequently give(s) extreme opinions without backing them up.
38. The anthropologist(s) who the villager(s) respect(s) the most has/have taken great pains to learn the local customs.
39. The bricklayer(s) who the contractor(s) assign(s) to important projects often doesn't show up on time.
40. The comedian(s) who the prisoner(s) find(s) so entertaining can find humor in the worst situations.
41. The judge(s) who the criminal(s) fear(s) the most hand(s) down tough sentences for minor offenses.
42. The student(s) who the bull(ies) tease(s) during lunch is/are too afraid to tell the teachers.

43. The philosopher(s) who the scientist(s) discuss(es) during the radio program rarely express(es) his/their ideas in plain language.
44. The lawyer(s) who the candidate(s) consult(s) every morning tell(s) him what campaign money he/they is/are allowed to accept.
45. The musician(s) who the reviewer(s) praise(s) so highly will probably win a Grammy.
46. The receptionist(s) who the boss(es) give(s) the biggest projects never fail(s) to do a stellar job.
47. The actress(es) who the designer(s) create(s) award show dresses for is/are extremely fickle and demanding.
48. The programmer(s) who the manager(s) oversee(s) at the software firm is/are eccentric and difficult to control.

Experiments 4-7

The experiments contained 24 sets of 2 (Exp. 6) or 4 (Exps. 4-5,7) conditions. Experiment 6 contained conditions a and c only. All experimental sentences used a singular subject. The pre-verbal adverb was not included in Experiment 5.

Conditions were as follows:

Condition a:	singular attractor, grammatical
Condition b:	plural attractor, grammatical
Condition c:	singular attractor, ungrammatical
Condition d:	plural attractor, ungrammatical

An example of a full set:

1a. The slogan on the poster (unsurprisingly) was designed to get attention.
1b. The slogan on the posters (unsurprisingly) was designed to get attention.
1c. The slogan on the poster (unsurprisingly) were designed to get attention.
1d. The slogan on the posters (unsurprisingly) were designed to get attention.

2. The picture on the flier(s) (definitely) was/were of a village church in the south of France.
3. The mistake in the program(s) (certainly) was/were disastrous for the small software company.
4. The label on the container(s) (probably) was/were a warning about the hazardous chemicals inside.
5. The problem in the store(s) (ultimately) was/were solved by firing the custodian.
6. The name on the billboard(s) (probably) was/were of a prominent local businessman.
7. The crime in the suburb(s) (doubtlessly) was/were a reflection of the violence in today's society.
8. The defect in the appliance(s) (likely) was/were unknown to consumers and government regulators.
9. The citation on the notecard(s) (unfortunately) was/were quite difficult to track down using the library's limited resources.
10. The advertisement on the skyscraper(s) (typically) was/were hard to read except from a long distance away.
11. The traitor in the camp(s) (regrettably) was/were passing information to the enemy through various secret methods.
12. The flaw in the argument(s) (usually) was/were clear to everyone in the lecture hall after just five minutes.
13. The door to the laborator(ies) (accidentally) was/were left unlocked by the cleaning service.
14. The memo from the supervisor(s) (unfortunately) was/were about the delinquent tax return.
15. The check from the stockbroker(s) (happily) was/were a dividend on a long-term bond.
16. The key to the cell(s) (unsurprisingly) was/were rusty from many years of disuse.
17. The letter from the investigator(s) (allegedly) was/were received in San Francisco in late March.
18. The entrance to the exhibit(s) (evidently) was/were hard to locate on the diagram.
19. The warning from the expert(s) (unsurprisingly) was/were a shock to the residents of the city.
20. The bridge to the island(s) (apparently) was/were about ten miles off the main highway.

21. The gift from the visitor(s) (apparently) was/were a beautiful clock from Germany.
22. The path to the monument(s) (happily) was/were lined with lovely trees and flowers.
23. The message from the engineer(s) (regrettably) was/were about the recent weakening to the building's structural supports.
24. The gate to the pasture(s) (evidently) was/were falling down due to disrepair.

Footnotes

1. The following standard linguistic category symbols are used throughout: N, noun; NP, noun phrase; P, preposition; PP, prepositional phrase; V, verb. RC stands for ‘relative clause.’ A category or word’s grammatical number is indicated in subscript as either SG, singular, or PL, plural.

2. References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language.* In press.

Badecker, B. & Lewis, R. (2007). A new theory and computational model of working memory in sentence production: Agreement errors as failures of cue-based retrieval. Paper presented at the 20th Annual CUNY Sentence Processing Conference. La Jolla, CA: University of California, San Diego.

Bayer, S., & Johnson, M. (1996). Features and Agreement. In *Proceedings of the 33rd Annual Meeting on Association for Computational Linguistics.* Morristown, NJ: Association for Computational Linguistics, 70-76

Bertram, R., Hyona, J. & Laine, M. (2000). Word morphology in context: Effects of base and surface frequency. *Language and Cognitive Processes, 15,* 367-388.

Bock, J. K. & Cutting, J. C. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language, 31*, 99-127.

Bock, J. K. & Eberhard, K. M. (1993). Meaning, sound, and syntax in English number agreement. *Language and Cognitive Processes, 8*, 57-99.

Bock, J. K., Eberhard, K. M., & Cutting, J. C. (2004). Producing number agreement: How pronouns equal verbs. *Journal of Memory and Language, 51,* 251-278.

Bock, J. K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology, 23*, 45-93.

Brunk, H.D. (1975). *An Introduction to Mathematical Statistics*, 3rd ed. Boston: Ginn.

Byram, M. (2007). Double objects in left to right syntax. Unpublished honors thesis, University of Arizona, Tucson.

Chomsky, N. (1970). Remarks on Nominalization. In RA Jacobs, PS Rosenbaum, eds., *Readings in English Transformational Grammar,* Waltham, MA: Ginn, 184-221.

Clark, H. (1973). The language-as-a-fixed-effect fallacy: critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12*, 335-359.

Clifton, C., Jr., Frazier, L., & Deevy, P. (1999). Feature manipulation in sentence comprehension. *Rivista di Linguistica, 11*, 11-39.

den Dikken, M. (2001). ìPluringularsî, pronouns and quirky agreement. *The Linguistic Review, 18*, 19-41.

Eberhard, K. (1997). The marked effect of number on subject-verb agreement. *Journal of*

Memory and Language, 36, 147-164.

Eberhard, K. M., Cutting, J. C., & Bock, J. K. (2005). Making syntax of sense: Number agreement in sentence production. *Psychological Review, 112*, 531-559.

Francis, W. N. (1986). Proximity concord in English. *Journal of English Linguistics*, *19*, 309-317.

Franck, J., Lassi, G., Frauenfelder, U., & Rizzi, L. (2006). Agreement and movement: a syntactic analysis of attraction. *Cognition, 101*, 173-216.

Franck, J., Vigliocco, G., & Nicol, J. (2002). Attraction in sentence production: The role of syntactic structure. *Language and Cognitive Processes, 17(4)*, 371-404.

Frazier, L. & Flores d'Arcais, G. (1989). Filler driven parsing: A study of gap filling in Dutch. *Journal of Memory and Language, 28*, 331-344.

Gazdar, G., Pullum, G.K., Klein, E., Sag, I.A. (1985). *Generalized phrase structure grammar.* Cambridge, MA: Harvard University Press.

Gillund, G. & Shiffrin, R.M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91,* 1-65.

Gouvea, A., Phillips, C., Kazanina, N., & Poeppel, D. (submitted). The linguistic processes underlying the P600.

Hagoort, P., Wassenaar, M. & Brown, C. (2003). Syntax-related ERP effects in Dutch. *Cognitive Brain Research, 16,* 38-50.

Hartsuiker, R. J., Antón-Méndez, I., & Van Zee, M. (2001). Object attraction in subject-verb agreement construction. *Journal of Memory and Language, 45*, 546-572.

Haskell, T. & MacDonald, M.C. (2003). Conflicting cues and competition in subject-verb agreement. *Journal of Memory and Language, 48*, 760-778.

H‰ussler, J. & Bader, M. (submitted). Agreement checking and number attraction in sentence comprehension: Insights from German relative clauses.

Hemforth, B. & Konieczny, L. (2003). Proximity in agreement errors. In *Proceedings of the 25th Annual Conference of the Cognitive Science Society*, Boston, MA.

Hintzman, D. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review, 95*: 528-551.

Jackendoff, R. S. (1977). *X′-Syntax: A Study of Phrase Structure*. Cambridge, MA: MIT Press.

Jespersen, O. (1924). *The philosophy of grammar.* London: Allen & Unwin.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General, 3*, 228-238.

Kaan, E. (2002). Investigating the effects of distance and number interference in agreement processing: An ERP study. *Journal of Psycholinguistic Research, 31*, 165-193.

Kayne, R. (1984). *Connectedness and Binary Branching*. Dordrecht, Netherlands: Foris.

Kimball, J., & Aissen, J. (1971). I think, you think, he think. *Linguistic Inquiry, 2,* 241-246.

Konieczny, L., Schimke, S., & Hemforth, B. (2004). An activation-based model of agreement errors in production and comprehension. In *Proceedings of the 26th Annual Meeting of the Cognitive Science Society, Chicago, IL.*

Lau, E., Rozanova, K., & Phillips, C. (2007). Syntactic prediction and lexical surface frequency effects in sentence processing. *University of Maryland Working Papers in Linguistics*, *16*, 163-200.

Lau, E., Wagers, M., Stroud, C., & Phillips, C. (2008). Agreement and the subject of confusion. Paper presented at the 21st Annual CUNY Sentence Processing Conference. Chapel Hill, NC: University of North Carolina.

Lau, E.F., Wagers, M., & Phillips, C., 2009. The consequences of number agreement on number interpretation. Poster presented at the 22nd Annual Meeting of of the CUNY Conference on Human Sentence Processing, Davis, CA: UC Davis.

Lehtonen, M., Wande, E., Niska, H., Niemi, J. & Laine, M. (2006). Recognition of inflected words in a morphologically limited language: frequency effects in monolinguals and bilinguals. *Journal of Psycholinguistic Research, 35(2),* 121-146.

Lewis, R. L. & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science, 29*, 1-45.

Luce, R. D. (1986). *Response times*: *their role in inferring elementary mental organization.* New York: Oxford University Press.

McElree, B. (2006). Accessing Recent Events. In B.H. Ross (ed.), *The psychology of learning and motivation,* Vol. 46. San Diego: Academic Press.

McElree, B., Foraker, S., Dyer, L. (2003) Memory structures that subserve sentence comprehension. *Journal of Memory and Language, 48,* 67-91.

New, B., Brysbaert, M., Segui, J., Ferrand, L., & Rastle, K. (2004). The processing of singular and plural nouns in French and English. *Journal of Memory and Language*, *51*, 568-585.

Niswander, E., Pollatsek, A., & Rayner, K. (2000). The processing of derived and inflected suffixed words during reading. *Language and Cognitive Processes, 15*(4/5), 389-420.

Nicol, J., Forster, K., & Veres, C. (1997). Subject-verb agreement processes in comprehension. *Journal of Memory and Language, 36*, 569-587.

Nunberg, G. (2003, March 9). Machines make moral judgments, selectively. *The New York Times,* p. 5.

Osterhout, L. & Mobley, L. A. (1995). Event-related potentials elicited by failure to agree. *Journal of Memory and Language, 34*, 739-773.

Pearlmutter, N. (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.

Plate, T. (2003). *Holographic Reduced Representation*. Stanford: CSLI Publications.

Pollard, C. & Sag, I. (1994). *Head-driven Phrase Structure Grammar*. Chicago: CSLI Publications.

Potter, M.C. (1988). Rapid serial visual presentation (RSVP): A method for studying language processing. In D. E. Kieras and M.A. Just (eds.), *New methods in reading comprehension research.* (Hillsdale, NJ: Erlbaum Press).

Potter, M.C. & Lombardi, L. (1992). The regeneration of syntax in short-term memory. *Journal of Memory and Language, 31,* 713-33.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language.* New York: Longman.

Raaijmakers, J. G. W., Schrijnemakers, J. M. C., & Gremmen, F. (1999). How to deal with "The language-as-fixed-effect fallacy": Common misconceptions and alternative solutions. *Journal of Memory and Language, 41*, 416-426.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*, 510-532.

Selman, B. & Hirst, G. (1985). A rule-based connectionist parsing system. In *Proceedings of the 7th Annual Conference of the Cognitive Science Society*, Hillsdale, NJ: Erlbaum. pp 212-221.

Shieber, S. M. (1986). *An introduction to unification-based approaches to grammar.* CSLI Lecture Notes No. 4. Stanford: CSLI.

Solomon, E. & Pearlmutter, N. (2004). Semantic integration and syntactic planning in language production. *Cognitive Psychology, 49*, 1-46.

Stowe, L. (1986). Parsing WH-constructions: Evidence for on-line gap location. *Language and Cognitive Processes, 1(3)*, 227-245.

Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language, 48(3)*, 542-562.

Tabor, W., Galantucci, B., & Richardson, D. (2000). Effects of merely local syntactic coherence on sentence processing. *Journal of Memory and Language, 50*, 355-370.

Thornton, R. & MacDonald, M. C. (2003). Plausibility and grammatical agreement. *Journal of Memory and Language, 48*, 740-759.

Traxler, M. J. & Pickering, M. J. (1996). Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language 35*, 454-475.

Trubetzkoy, N.S. (1939). *Grundzüge der Phonologie*. Christiane Baltaxe, trans. *Principles of Phonology*, Berkeley: University of California Press, 1969.

Van der Velde, F. & de Kamps, M. (2006). Neural blackboard architectures of combinatorial structures in cognition. *Behavioral and Brain Sciences, 29*, 37-70.

van Riemsdijk, H. C. & Williams, E. (1986). *Introduction to the Theory of Grammar*. Cambridge, MA: MIT Press.

Vasishth, S. (2006). On the proper treatment of spillover in real-time reading studies: consequences for psycholinguistic theories. In *Proceedings of the International Conference on Linguistic Evidence*. Tübingen, Germany.

Vigliocco, G. & Nicol, J. (1998). Separating hierarchical relations and word order in language production. Is proximity concord syntactic or linear? *Cognition, 68*, 13-29.

Wagers, M. (2008). The structure of memory meets memory for structure in linguistic cognition. Unpublished doctoral dissertation, University of Maryland, College Park.

Table 1
Sample set of experimental items for Experiment 1

singular subject/ grammatical	The old key unsurprisingly was rusty from many years of disuse…
plural subject/ grammatical	The old keys unsurprisingly were rusty from many years of disuse…
singular subject/ ungrammatical	The old key unsurprisingly were rusty from many years of disuse…
plural subject/ ungrammatical	The old keys unsurprisingly was rusty from many years of disuse…

Table 2 **Experiment 1 Omnibus Repeated Measures Analysis of Variance**

ANOVA Tests reliable at $\alpha = 0.05$ highlighted in bold.

	By participants				By items				MinF'		
	df	MS_{effect}	F_1	*p*	df	MS_{effect}	F_2	*p*	df	minF'	*p*
Region 3 (subject)											
Grammaticality	1,31	3724	0.18	0.67	1,23	4105	0.15	0.70	1,51	0.08	0.78
Subject number	1,31	216	0.01	0.92	1,23	446	0.01	0.92	1,53	0.01	0.92
Number x gram.	1,31	14652	0.83	0.37	1,23	21556	0.91	0.35	1,53	0.43	0.51
Region 4 (adverb)											
Grammaticality	1,31	16047	0.64	0.43	1,23	9537	0.14	0.72	1,33	0.11	0.74
Subject number	**1,31**	**664627**	**14.5**	**<0.001**	**1,23**	**689556**	**9.34**	**0.01**	**1,48**	**5.68**	**0.02**
Number x gram.	1,31	18589	0.65	0.43	1,23	18982	0.47	0.50	1,49	0.27	0.61
Region 5 (verb)											
Grammaticality	**1,31**	**376432**	**8.97**	**0.01**	**1,23**	**353467**	**6.94**	**0.01**	1,50	3.91	0.05
Subject number	1,31	2595	0.09	0.77	1,23	3312	0.14	0.71	1,54	0.05	0.82
Number x gram.	1,31	617	0.02	0.88	1,23	256	0.01	0.91	1,43	0.01	0.92
Region 6 (verb+1)											
Grammaticality	**1,31**	**1922696**	**59.2**	**<0.001**	**1,23**	**1957355**	**50.3**	**<0.001**	**1,51**	**27.2**	**<0.001**
Subject number	1,31	101120	3.77	0.06	1,23	112310	4.15	0.05	1,53	1.97	0.17
Number x gram.	1,31	27296	0.74	0.4	1,23	21992	0.48	0.49	1,48	0.29	0.59
Region 7 (verb+2)											
Grammaticality	**1,31**	**210563**	**16.44**	**<0.001**	**1,23**	**221568**	**7.47**	**0.01**	**1,42**	**5.14**	**0.03**
Subject number	1,31	25372	2.72	0.11	1,23	20459	1.50	0.23	1,45	0.97	0.33
Number x gram.	1,31	9116	0.68	0.42	1,23	10458	0.84	0.37	1,54	0.38	0.54
Region 8 (verb+3)											
Grammaticality	**1,31**	**50907**	**5.37**	**0.03**	1,23	58906	4.07	0.06	1,50	2.32	0.13
Subject number	1,31	23238	2.06	0.16	1,23	16298	1.00	0.33	1,43	0.67	0.42
Number x gram.	1,31	8670	0.92	0.35	1,23	11122	1.15	0.29	1,54	0.51	0.48
Region 9 (verb+4)											
Grammaticality	1,31	7005	0.87	0.36	1,23	5772	0.36	0.56	1,41	0.25	0.62
Subject number	1,31	20268	2.31	0.14	1,23	16710	0.83	0.37	1,39	0.61	0.44
Number x gram.	1,31	4617	0.53	0.47	1,23	3756	0.36	0.55	1,48	0.22	0.64

Table 3

Sample set of experimental items for Experiments 2 & 3

Singular subject items (Experiments 2 & 3)	
singular attractor/ grammatical	The musician who the reviewer praises so highly will probably win a Grammy.
plural attractor/ grammatical	The musicians who the reviewer praises so highly will…
singular attractor/ ungrammatical	The musician who the reviewer praise so highly will…
plural attractor/ ungrammatical	The musicians who the reviewer praise so highly will…
Plural subject items (Experiment 3)	
singular attractor/ grammatical	The musician who the reviewers praise so highly will…
plural attractor/ grammatical	The musicians who the reviewers praise so highly will…
singular attractor/ ungrammatical	The musician who the reviewers praises so highly will…
plural attractor/ ungrammatical	The musicians who the reviewers praises so highly will…

Table 4 **Experiment 2 Omnibus Repeated Measures Analysis of Variance**

ANOVA Tests reliable at $\alpha = 0.05$ highlighted in bold.

	By participants				By items				MinF'		
	Df	MS_{effect}	F_1	*p*	df	MS_{effect}	F_2	*p*	df	minF'	*P*
Region 2 (RC head)											
Attractor number	**1,27**	**3642**	**5.7**	**0.02**	1,47	7715	2.84	0.09	1,74	1.89	0.17
Grammaticality	1,27	1866	1.68	0.21	1,47	3506	2.15	0.14	1,63	0.94	0.34
Number x gram.	1,27	307	0.5	0.49	1,47	4547	2.42	0.12	1,38	0.41	0.53
Region 3 ('who')											
Attractor number	**1,27**	**4109**	**6.49**	**0.02**	1,47	5004	1.97	0.16	1,69	1.51	0.22
Grammaticality	1,27	704.1	1.23	0.28	1,47	397	0.25	0.62	1,64	0.21	0.65
Number x gram.	1,27	1471	1.56	0.22	1,47	2914	1.87	0.17	1,65	0.85	0.36
Region 4 ('the')											
Attractor number	1,27	400	1.75	0.20	1,47	1058	0.79	0.38	1,73	0.54	0.46
Grammaticality	1,27	977	2.31	0.14	1,47	284	0.29	0.59	1,58	0.25	0.62
Number x gram.	1,27	187	0.22	0.64	1,47	243	0.13	0.72	1,74	< 0.1	0.78
Region 5 (RC subject)											
Attractor number	**1,27**	**4001**	**6.26**	**0.02**	**1,47**	**10474**	**4.8**	**0.03**	1,73	2.71	0.1
Grammaticality	1,27	282	0.61	0.44	1,47	11	< 0.1	0.94	1,48	< 0.1	0.92
Number x gram.	1,27	51	< 0.1	0.81	1,47	1020	0.51	0.47	1,33	< 0.1	0.82
Region 6 (verb)											
Attractor number	1,27	3051	2.06	0.16	1,47	7341	2.04	0.15	1,69	1.02	0.32
Grammaticality	1,27	270	0.33	0.57	1,47	335	< 0.1	0.78	1,66	< 0.1	0.79
Number x gram.	1,27	2722	0.32	0.58	1,47	321	0.12	0.73	1,71	< 0.1	0.77
Region 7 (verb+1)											
Attractor number	**1,27**	**28444**	**8.24**	**0.01**	**1,47**	**49733**	**12.2**	**<0.01**	**1,60**	**4.91**	**0.03**
Grammaticality	1,27	**38706**	**5.31**	**0.03**	**1,47**	**56498**	**9.87**	**<0.01**	1,55	3.45	0.07
Number x gram.	1,27	**17078**	**6.92**	**0.01**	**1,47**	**31842**	**10.6**	**<0.01**	**1,59**	**4.18**	**0.05**
Region 8 (verb+2)											
Attractor number	1,27	2846	2.37	0.14	1,47	4518	1.09	0.30	1,73	0.75	0.39
Grammaticality	1,27	3312	1.97	0.17	1,47	4587	1.32	0.25	1,74	0.79	0.38
Number x gram.	1,27	869	1.27	0.27	1,47	1102	0.3	0.58	1,66	0.24	0.63
Region 9 (verb+3)											
Attractor number	**1,27**	**3845**	**4.26**	**0.05**	**1,47**	**7502**	**4.24**	**0.04**	1,69	2.12	0.15
Grammaticality	1,27	0.6	< 0.1	0.98	1,47	490	0.1	0.75	1,27	< 0.1	1
Number x gram.	1,27	590	0.88	0.36	1,47	1264	0.91	0.34	1,68	0.45	0.5
Region 10 (verb+4)											
Attractor number	**1,27**	4451	3.75	0.06	**1,47**	**7104**	**3.8**	**0.05**	1,68	1.89	0.17
Grammaticality	**1,27**	5062	3.3	0.08	**1,47**	**11269**	**5.91**	**0.02**	1,56	2.12	0.15
Number x gram.	1,27	188	0.227	0.64	1,47	1198	0.503	0.48	1,51	0.16	0.69

Table 5 **Experiment 3 Omnibus Repeated Measures Analysis of Variance**

ANOVA Tests reliable at $\alpha = 0.05$ highlighted in bold.

	By participants				By items				MinF'		
	df	MS_{effect}	F_1	*p*	df	MS_{effect}	F_2	*p*	df	minF'	*p*
Region 2 (RChead)											
Grammaticality	1,55	643	0.35	0.56	1,47	2416	2.08	0.16	1,73	0.3	0.59
Attractor number	**1,55**	**20084**	**8.18**	**0.01**	**1,47**	**14098**	**5.89**	**0.02**	1,96	3.42	0.07
Subject number	1,55	2643	1.43	0.24	1,47	4792	2.05	0.16	1,101	0.84	0.36
Att-Num x gram.	1,55	9320	2.52	0.12	1,47	3402	1.05	0.31	1,82	0.74	0.39
Sub-Num x gram	1,55	711	0.39	0.59	1,47	580	0.26	0.62	1,100	0.14	0.71
A-num x S-num	1,55	9320	2.52	0.12	1,47	6379	2.36	0.13	1,101	1.22	0.27
3-way interaction	1,55	1042	0.53	0.47	1,47	353	0.15	0.70	1,72	0.12	0.73
Region 3 ('who')											
Grammaticality	1,55	314	0.29	0.59	1,47	75	< 0.1	0.84	1,102	0.91	0.34
Attractor number	**1,55**	**8909**	**5.52**	**0.02**	1,47	5052	3.6	0.06	1,94	2.18	0.14
Subject number	1,55	2212	1.26	0.27	**1,47**	**7988**	**4.22**	**0.05**	1,84	0.97	0.33
Att-Num x gram.	1,55	2	< 0.1	0.97	1,47	18	< 0.1	0.93	1,78	< 0.1	1
Sub-Num x gram	1,55	152	< 0.1	0.76	1,47	570	0.47	0.5	1,75	< 0.1	0.78
A-num x S-num	1,55	1876	1.01	0.32	1,47	1854	1.9	0.18	1,97	0.66	0.42
3-way interaction	**1,55**	**4427**	**4.03**	**0.05**	1,47	3437	1.93	0.17	1,86	1.3	0.26
Region 5 (RC subj)											
Grammaticality	1,55	703	0.27	0.60	1,47	2749	1	0.32	1,82	0.21	0.65
Attractor number	**1,55**	**15895**	**7.03**	**0.01**	**1,47**	**10596**	**4.75**	**0.03**	1,95	2.83	0.1
Subject number	**1,55**	**8661**	**4.34**	**0.04**	**1,47**	**9244**	**5.35**	**0.03**	1,102	2.4	0.12
Att-Num x gram.	1,55	828	0.39	0.54	1,47	3723	1.02	0.32	1,90	0.28	0.6
Sub-Num x gram	1,55	1945	0.76	0.39	1,47	2940	1.36	0.25	1,98	0.49	0.49
A-num x S-num	1,55	2941	1.21	0.28	1,47	1948	0.76	0.39	1,93	0.46	0.5
3-way interaction	1,55	3871	1.67	0.20	1,47	3319	1.21	0.28	1,96	0.70	0.4
Region 6 (verb)											
Grammaticality	1,55	942	0.3	0.59	1,47	1831	0.72	0.40	1,92	0.21	0.65
Attractor number	1,55	1443	0.61	0.44	1,47	261	< 0.1	0.80	1,57	< 0.1	0.81
Subject number	**1,55**	**35005**	**11.6**	**0.001**	**1,47**	**31055**	**11.4**	**0.001**	**1,101**	**5.76**	**0.02**
Att-Num x gram.	1,55	540	0.22	0.64	1,47	61	0.02	0.88	1,56	< 0.1	0.89
Sub-Num x gram	1,55	717	0.25	0.62	1,47	1477	0.45	0.50	1,98	0.16	0.69
A-num x S-num	1,55	1	< 0.1	0.99	1,47	12	< 0.1	0.96	1,64	< 0.1	0.99
3-way interaction	1,55	7927	2.39	0.13	1,47	6128	1.79	0.19	1,7	1.02	0.32
Region 7 (verb+1)											
Grammaticality	**1,55**	**363776**	**28.6**	**<0.001**	**1,47**	**251224**	**42.1**	**<0.001**	**1,101**	**17**	**<0.001**
Attractor number	1,55	11147	1.81	0.18	1,47	4380	1.12	0.29	1,93	0.69	0.41
Subject number	1,55	4289	0.45	0.50	1,47	1817	0.45	0.51	1,101	0.22	0.64
Att-Num x gram.	1,55	19253	3.48	0.07	1,47	13675	1.89	0.18	1,89	1.23	0.27
Sub-Num x gram	1,55	2216	0.54	0.47	1,47	141	< 0.1	0.86	1,52	< 0.1	0.86
A-num x S-num	1,55	130	< 0.1	0.89	1,47	1517	0.29	0.59	1,62	< 0.1	0.89
3-way interaction	1,55	1044	0.30	0.58	1,47	5546	1.37	0.25	1,78	0.25	0.62
Region 8 (verb+2)											
Grammaticality	1,55	5540	1.39	0.24	1,47	1566	0.47	0.50	1,77	0.35	0.56
Attractor number	1,55	2723	0.92	0.34	1,47	2785	1.04	0.31	1,102	0.49	0.49
Subject number	1,55	45	< 0.1	0.90	1,47	49	0.02	0.90	1,102	< 0.1	0.9
Att-Num x gram.	1,55	7739	2.65	0.11	1,47	9174	2.85	0.10	1,102	1.37	0.24
Sub-Num x gram	1,55	4338	1.30	0.26	1,47	2174	0.90	0.35	1,95	0.53	0.47
A-num x S-num	1,55	1501	0.61	0.44	1,47	545	0.25	0.62	1,82	0.18	0.67
3-way interaction	**1,55**	**27327**	**6.03**	**0.02**	**1,47**	**31886**	**7.68**	**0.01**	1,102	3.38	0.07

Table 6 **Experiment 3 Repeated Measures Analysis of Variance by RC subject number**

First number is 2x2 for RC Subject=singular, second number is 2x2 for RC Subject=plural. ANOVA Tests reliable at $\alpha = 0.05$ in bold.

	By participants				By items				MinF'		
	df	MS_{effect}	F_1	*p*	df	MS_{effect}	F_2	*p*	df	minF'	*p*
Region 2 (RC head)											
Grammaticality	1,55	1354 | 1	0.65 | < 0.1	0.43 | 0.98	1,47	2682 | 314	1.79 | 0.16	0.19 | 0.69	1,88 | 1,55	0.47 | < 0.1	0.49 | 0.98
Attractor numb.	1,55	**28383** | 1021	**9.4** | 0.35	**< 0.01** | 0.56	1,47	**19721** | 755	**7.39** | 0.31	**0.01** | 0.58	1,98 | 1,100	**4.14** | 0.17	**0.04** | 0.68
Number x gram.	1,55	**7950** | 1894	**4.19** | 0.84	**0.05** | 0.36	1,47	2974 | 781	0.95 | 0.31	0.33 | 0.58	1,68 | 1,79	0.78 | 0.23	0.38 | 0.63
Region 3 ('who')											
Grammaticality	1,55	451 | 15	0.35 | < 0.1	.56 | .92	1,47	115 | 530	< 0.1 | 0.34	0.79 | 0.56	1,66 | 1,56	< 0.1 | < 0.1	0.81 | 0.92
Attractor number	1,55	1305 | **9480**	0.54 | **9.1**	0.47 | **< 0.01**	1,47	393 | **6514**	0.29 | **6.35**	0.59 | **.02**	1,89 | 1,96	0.19 | 3.74	0.66 | 0.06
Number x gram.	1,55	2319 | 2110	0.35 | 1.87	0.56 | 0.18	1,47	1478 | 1978	0.84 | 0.86	0.36 | 0.36	1,92 | 1,85	0.25 | 0.59	0.62 | 0.44
Region 5 (RC subj.)											
Grammaticality	1,55	155 | 2494	< 0.1 | 0.78	0.78 | 0.38	1,47	2 | 5687	< 0.1 | 2.2	0.98 | 0.14	1,48 | 1,85	< 0.1 | 0.57	0.98 | 0.45
Attractor number	1,55	2581 | **16255**	0.98 | **7.85**	0.33 | **0.01**	1,47	1729 | 10815	0.36 | 3.86	0.36 | .06	1,100 | 1,87	0.46 | 2.59	0.5 | 0.11
Number x gram.	1,55	559 | 4140	0.27 | 1.76	0.61 | 0.19	1,47	6 | 7036	< 0.1 | 1.78	0.96 | 0.19	1,48 | 1,101	< 0.1 | 0.88	0.96 | 0.35
Region 6 (verb)											
Grammaticality	1,55	8 | 1651	< 0.1 | 0.42	0.95 | 0.52	1,47	10 | 3299	< 0.1 | 1.1	0.95 | 0.3	1,101 | 1,87	< 0.1 | 0.31	0.97 | 0.58
Attractor number	1,55	677 | 767	0.27 | 0.2	0.6 | 0.66	1,47	194 | 80	< 0.1 | < 0.1	0.84 | 0.89	1,61 | 1,55	< 0.1 | < 0.1	0.84 | 0.89
Number x gram.	1,55	6302 | 2165	2.14 | 0.74	0.15 | 0.39	1,47	3708 | 2482	1.13 | 0.82	0.29 | 0.37	1,89 | 1,102	0.74 | 0.39	0.39 | 0.53
Region 7 (verb+1)											
Grammaticality	1,55	**154603 | 211389**	**18.2 | 25.5**	**<.001 | <.001**	1,47	**131642 | 119723**	**17.8 | 38.1**	**<.001 | <.001**	**1,101 | 1,98**	**25.5 | 38.1**	**<.001 | <.001**
Attractor number	1,55	4436 | 6841	0.93 | 0.87	0.34 | 0.36	1,47	5527 | 371	1.11 | < 0.1	0.3 | 0.77	1,102 | 1,57	0.51 | < 0.1	0.48 | 0.36
Number x gram.	1,55	14631 | 5666	3.40 | 1.22	0.07 | 0.27	1,47	18319 | 902	3.84 | 0.14	0.06 | 0.71	1,102 | 1,58	1.8 | 0.12	0.18 | 0.73
Region 8 (verb+2)											
Grammaticality	1,55	37 | 9841	< 0.1 | 2.12	0.91 | 0.15	1,47	25 | 3715	< 0.1 | 1.37	0.78 | 0.25	1,73 | 1,93	< 0.1 | 0.83	0.92 | 0.36
Attractor number	1,55	4135 | 90	1.44 | < 0.1	0.24 | 0.85	1,47	2896 | 433	0.93 | 0.26	0.34 | 0.62	1,94 | 1,70	0.56 | < 0.1	0.46 | 0.86
Number x gram.	1,55	**32075** | 2991	**8.52** | 0.81	**<0.01** | 0.37	1,47	**37633** | 3427	**8.63** | 1.14	**0.01** | 0.29	**1,101** | 1,101	**4.29** | 0.47	**0.04** | 0.49

Table 7
Sample set of experimental items for Experiments 4-6. Note that the adverb region was included in Experiments 4 and 6 only.

Grammatical items (Experiments 4-6)	
singular attractor/ grammatical	The key to the cell (unsurprisingly) was rusty from many years of disuse.
plural attractor/ grammatical	The key to the cells (unsurprisingly) was rusty from many years of disuse.
Ungrammatical items (Experiments 4-5)	
singular attractor/ ungrammatical	The key to the cell (unsurprisingly) were rusty from many years of disuse.
plural attractor/ ungrammatical	The key to the cells (unsurprisingly) were rusty from many years of disuse.

Table 8 **Experiment 4 Omnibus Repeated Measures Analysis of Variance**

ANOVA Tests reliable at $\alpha = 0.05$ highlighted in bold.

	By participants				By items				MinF'		
	df	MS_{effect}	F_1	*p*	df	MS_{effect}	F_2	*P*	df	minF'	*P*
Region 5 (attractor)											
Grammaticality	1,43	9	< 0.1	0.93	1,23	209	0.15	0.70	1,47	< 0.1	0.92
Attractor number	**1,43**	**17073**	**10.9**	**< 0.01**	**1,23**	**14163**	**12.8**	**< 0.01**	**1,63**	**5.88**	**0.02**
Number x gram.	1,43	5243	2.71	0.11	1,23	766.8	0.62	0.44	1,34	0.50	0.48
Region 6 (adverb)											
Grammaticality	1,43	96	< 0.1	0.84	1,23	19	< 0.1	0.91	1,37	< 0.1	0.92
Attractor number	1,43	**17180**	**5.1**	**0.03**	1,23	**10507**	**5.76**	**0.02**	1,62	2.70	0.11
Number x gram.	1,43	1070	0.58	0.45	1,23	11	< 0.1	0.94	1,23	< 0.1	0.92
Region 7 (verb)											
Grammaticality	**1,43**	**12000**	**5.1**	**0.03**	**1,23**	**6688**	**4.55**	**0.04**	1,58	2.41	0.13
Attractor number	1,43	3	< 0.1	0.97	1,23	724	0.26	0.62	1,43	< 0.1	1
Number x gram.	**1,43**	**16963**	**7.77**	**0.01**	1,23	2540	1.5	0.23	1,32	1.26	0.27
Region 8 (verb+1)											
Grammaticality	**1,43**	**37681**	**15.3**	**< 0.001**	**1,23**	**27042**	**27.0**	**< 0.001**	**1,66**	**9.76**	**< 0.01**
Attractor number	1,43	6074	2.09	0.16	1,23	3662	2.32	0.14	1,62	1.10	0.3
Number x gram.	**1,43**	**21399**	**9.91**	**< 0.01**	1,23	7962	3.31	0.08	1,39	2.48	0.12
Region 9 (verb+2)											
Grammaticality	**1,43**	**27089**	**21.1**	**< 0.001**	**1,23**	**14408**	**23.4**	**< 0.001**	**1,62**	**11.1**	**0.001**
Attractor number	1,43	4428	2.97	0.09	1,23	1855	1.84	0.19	1,50	1.14	0.29
Number x gram.	**1,43**	**10368**	**10.3**	**< 0.01**	1,23	4302	3.08	0.09	1,37	2.37	0.13
Region 10 (verb+3)											
Grammaticality	1,43	2708	1.5	0.23	1,23	2296	2.34	0.14	1,66	0.91	0.34
Attractor number	1,43	552	0.38	0.54	1,23	104	< 0.1	0.78	1,33	< 0.1	0.79
Number x gram.	1,43	4634	4	0.05	1,23	841	0.8	0.38	1,32	0.66	0.42
Region 11 (verb+4)											
Grammaticality	**1,43**	**18134**	**17.5**	**< 0.001**	**1,23**	**12543**	**10.5**	**< 0.01**	**1,49**	**6.55**	**0.01**
Attractor number	1,43	338	0.35	0.56	1,23	532	0.61	0.44	1,66	0.22	0.64
Number x gram.	1,43	958	0.6	0.44	1,23	6	< 0.1	0.94	1,23	< 0.1	0.92

Table 9 **Experiment 5 Omnibus Repeated Measures Analysis of Variance**

ANOVA Tests reliable at $\alpha = 0.05$ highlighted in bold.

	By participants				By items				MinF'		
	df	MS_{effect}	F_1	*P*	df	MS_{effect}	F_2	*p*	df	minF'	*p*
Region 5 (attractor)											
Grammaticality	1,55	92	< 0.1	0.80	1,23	44.7	< 0.1	0.81	62	< 0.1	0.86
Attractor number	**1,55**	**27160**	**12.7**	**< 0.01**	**1,23**	**9050**	**5.85**	**0.02**	**45**	**4.01**	**0.05**
Number x gram.	**1,55**	**10609**	**5.65**	**0.02**	**1,23**	**5475.4**	**8.56**	**< 0.01**	74	3.40	0.07
Region 6 (verb)											
Grammaticality	1,55	780	0.41	0.52	1,23	614.5	0.72	0.40	76	0.26	0.61
Attractor number	**1,55**	**11293**	**5.25**	**0.03**	**1,23**	**4188.2**	**4.24**	**0.05**	59	2.34	0.13
Number x gram.	1,55	2643	1.25	0.27	1,23	2323.6	3.28	0.08	78	0.91	0.34
Region 7 (verb+1)											
Grammaticality	**1,55**	**275561**	**59.3**	**< 0.001**	**1,23**	**114352**	**56.5**	**<0.001**	**64**	**28.9**	**<0.001**
Attractor number	**1,55**	**35484**	**9.24**	**< 0.01**	**1,23**	**14903**	**7.33**	**< 0.01**	**59**	**4.09**	**0.05**
Number x gram.	**1,55**	**60373**	**13.5**	**< 0.001**	**1,23**	**23232**	**10.8**	**< 0.01**	**59**	**6.02**	**0.02**
Region 8 (verb+2)											
Grammaticality	**1,55**	**57641**	**21.8**	**0.001**	**1,23**	**24761**	**21.2**	**< 0.001**	**64**	**10.8**	**<0.01**
Attractor number	1,55	302	< 0.1	0.76	1,23	814.9	0.82	0.38	67	< 0.1	0.77
Number x gram.	**1,55**	**9226**	**3.9**	**0.05**	1,23	2623	2.67	0.12	55	1.58	0.21
Region 9 (verb+3)											
Grammaticality	**1,55**	**14617**	**8.12**	**0.01**	1,23	5837	2.81	0.11	40	2.09	0.16
Attractor number	1,55	< 1	< 0.1	0.99	1,23	11.4	< 0.1	0.89	55	< 0.1	1
Number x gram.	**1,55**	**10311**	**6.16**	**0.02**	**1,23**	**3999**	**7.27**	**0.01**	69	3.33	0.07
Region 10 (verb+4)											
Grammaticality	1,55	1104	0.79	0.38	1,23	346.1	0.49	0.49	52	0.30	0.59
Attractor number	1,55	127	< 0.1	0.76	1,23	9.9	< 0.1	0.91	29	< 0.1	0.92
Number x gram.	1,55	5556	3.2	0.08	1,23	1477.7	1.73	0.20	49	1.12	0.3

Table 10 **Experiment 6 Omnibus Repeated Measures Analysis of Variance**

ANOVA Tests reliable at $\alpha = 0.05$ highlighted in bold.

	By participants				By items				MinF'		
	df	MS_{effect}	F_1	*p*	df	MS_{effect}	F_2	*p*	df	minF'	*p*
Region 5 (attractor)											
Attractor number	1,27	2624	1.89	0.18	1,23	5281	3.12	0.09	1,49	1.18	0.28
Region 6 (adverb)											
Attractor number	1,27	7775	1.69	0.21	1,23	5804	1.39	0.25	1,48	0.76	0.39
Region 7 (verb)											
Attractor number	1,27	16	< 0.1	0.92	1,23	43	< 0.1	0.86	1,43	< 0.1	0.92
Region 8 (verb+1)											
Attractor number	1,27	209.5	0.19	0.67	1,23	126	< 0.1	0.77	1,42	< 0.1	0.81
Region 9 (verb+2)											
Attractor number	1,27	403	0.31	0.58	1,23	131	< 0.1	0.77	1,35	< 0.1	0.79
Region 10 (verb+3)											
Attractor number	1,27	831	0.71	0.41	1,23	880	0.7	0.41	1,50	0.35	0.56
Region 11 (verb+4)											
Attractor number	1,27	1058	1.18	0.29	1,23	702.8	0.74	0.40	1,45	0.45	0.51

Exp.	Baseline	Pooled s.d.	Effect		
			Grammaticality	Attraction	
				Grammatical	Ungrammatical
1	401 ms	184 ms	115 ms (86, 144)	--	--
2	317 ms	174 ms	62 ms (39, 84)	-8 ms (-30, 16)	-49 ms (-80, -16)
3	340 ms	175 ms	68 ms (45, 94)	9 ms (-15, 35)	-34 ms (-70, -1)
4	312 ms	127 ms	53 ms (33, 69)	10 ms (-11, 26)	-45 ms (-70, -17)
5	327 ms	149 ms	100 ms (80,122)	8 ms (-14, 31)	-63 ms (-92, -33)
6	309 ms	126 ms	---	1 ms (-16, 18)	---

Table 11

Grammaticality and Attraction Effects in Region Following Verb

Region of interest is one region beyond the critical verb. 95% confidence interval lower and upper bounds presented within parentheses. 'Baseline' denotes experiment-wide average reading time per word. Pooled standard deviation is computed only over the region of interest.

Exp.	Baseline	Pooled s.d.	PLURALITY EFFECT	
1	401 ms	201 ms	57 ms	(21, 100)
2	317 ms	159 ms	11 ms	(-0.1, 23)
3	340 ms	105 ms	9 ms	(2, 16)
4	312 ms	134 ms	19 ms	(3, 36)
5	327 ms	124 ms	14 ms	(3, 26)
6	309 ms	162 ms	15 ms	(-4, 37)

Table 12 Effect of Plurality in Region following the Attractor

Region of interest is one region beyond the attractor noun. 95% confidence interval lower and upper bounds are presented within parentheses. Baseline denotes experiment-wide average reading time per word. Pooled standard deviation is computed only over the region of interest.

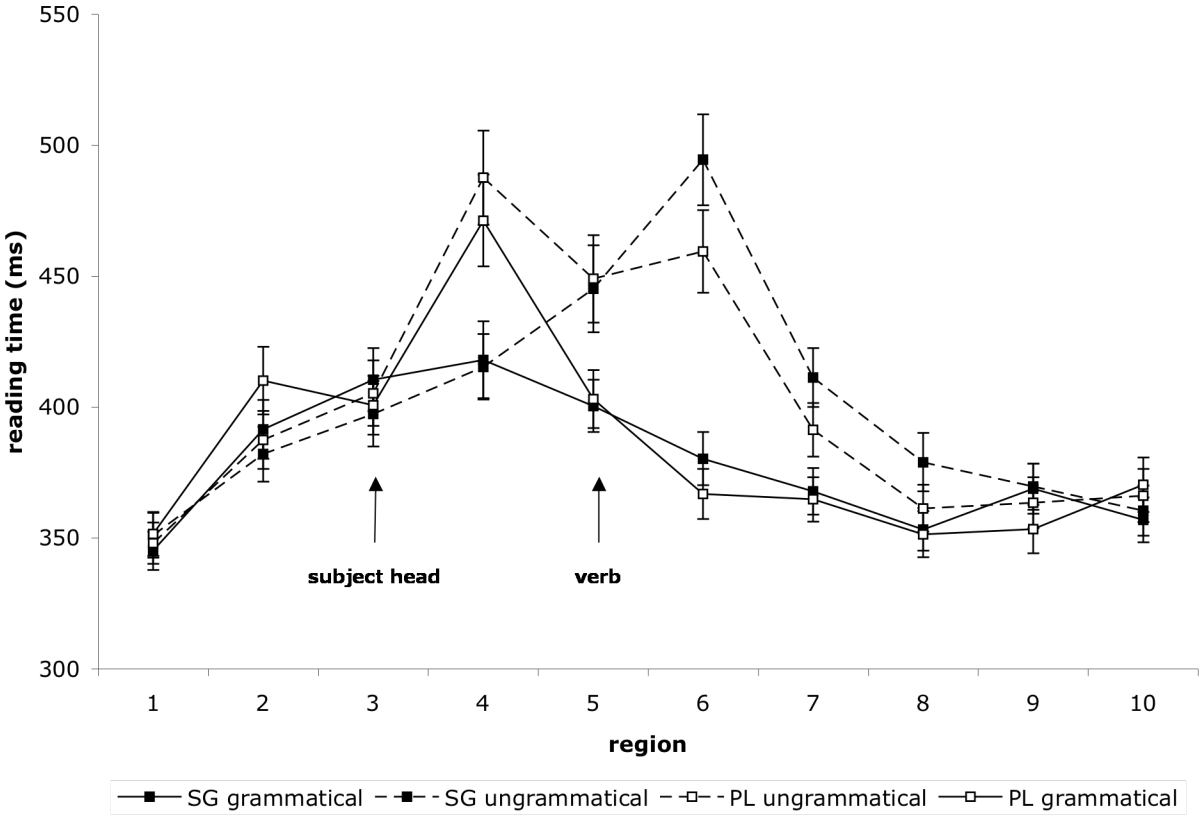

Figure 1 **Experiment 1 Self-paced Reading Results**

Region by region means according to subject number and grammaticality. Error bars indicate standard error of the mean.

Sample sentence:
The$_1$ old$_2$ key(s)$_3$ unsurprisingly$_4$ was/were$_5$ rusty$_6$ from$_7$ many$_8$ years$_9$ of$_{10}$ ...

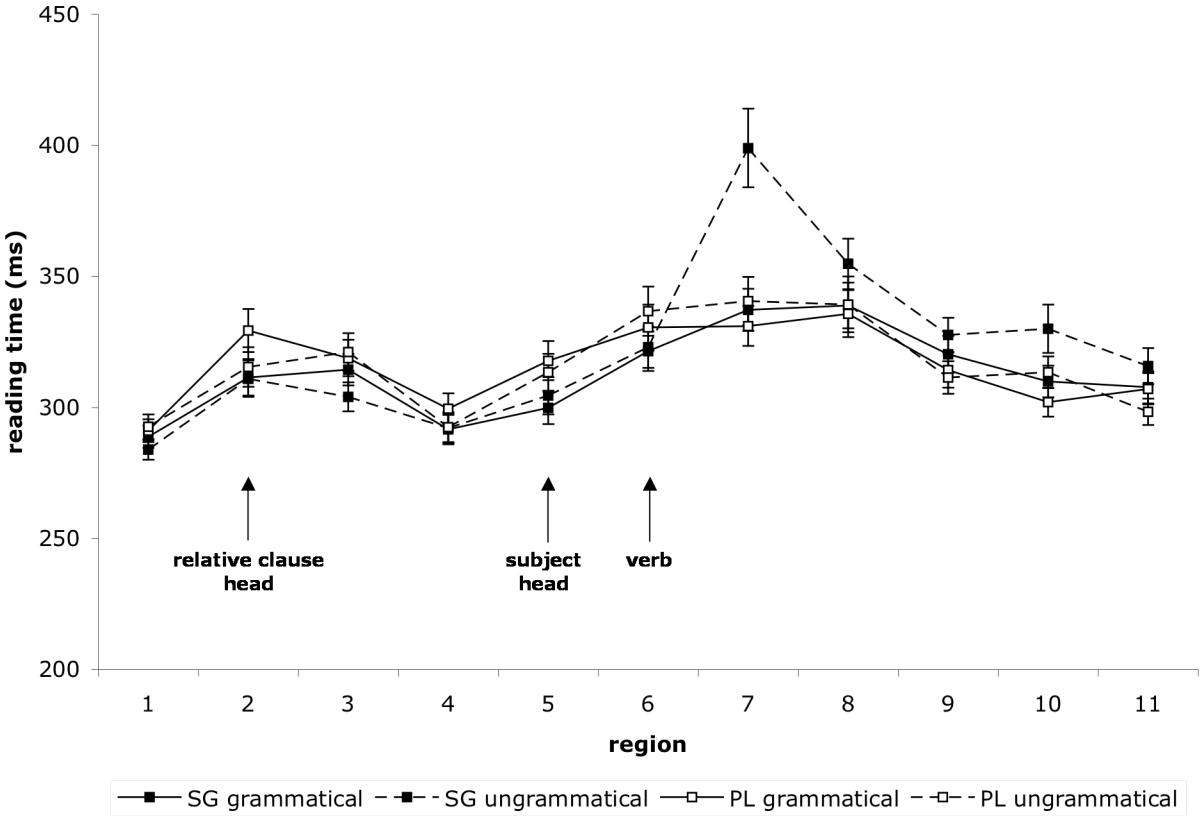

Figure 2 **Experiment 2 Self-paced Reading Results**

Region by region means segregated by relative clause head number and grammaticality. Error bars indicate standard error of the mean.

Sample sentence:
The$_1$ musician(s)$_2$ who$_3$ the$_4$ reviewer$_5$ praise(s)$_6$ so$_7$ highly$_8$ will$_9$ probably$_{10}$ win(s)$_{11}$ ...

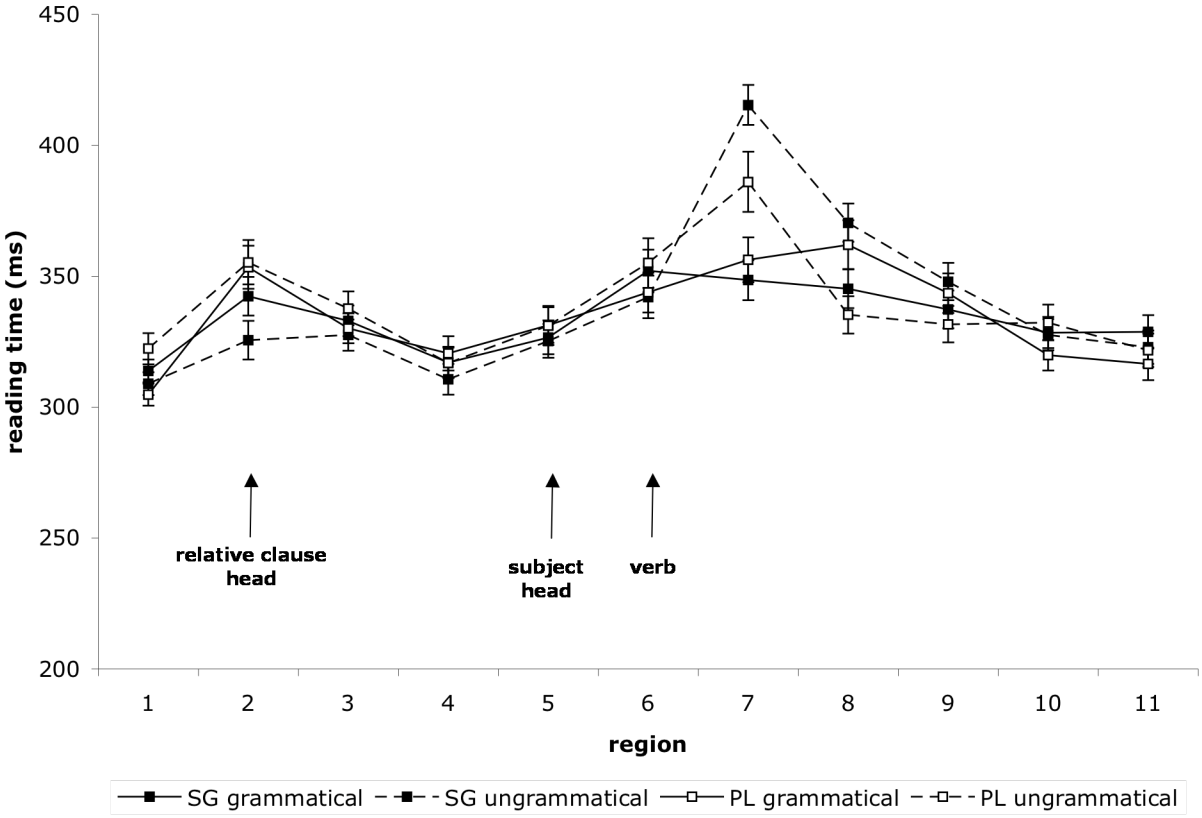

Figure 3 **Experiment 3 Self-paced Reading Results: Singular Relative Clause Subjects**

Region by region means segregated by relative clause head number and grammaticality. Error bars indicate standard error of the mean.

Sample sentence:
The$_1$ musician(s)$_2$ who$_3$ the$_4$ reviewer$_5$ praise(s)$_6$ so$_7$ highly$_8$ will$_9$ probably$_{10}$ win(s)$_{11}$ ...

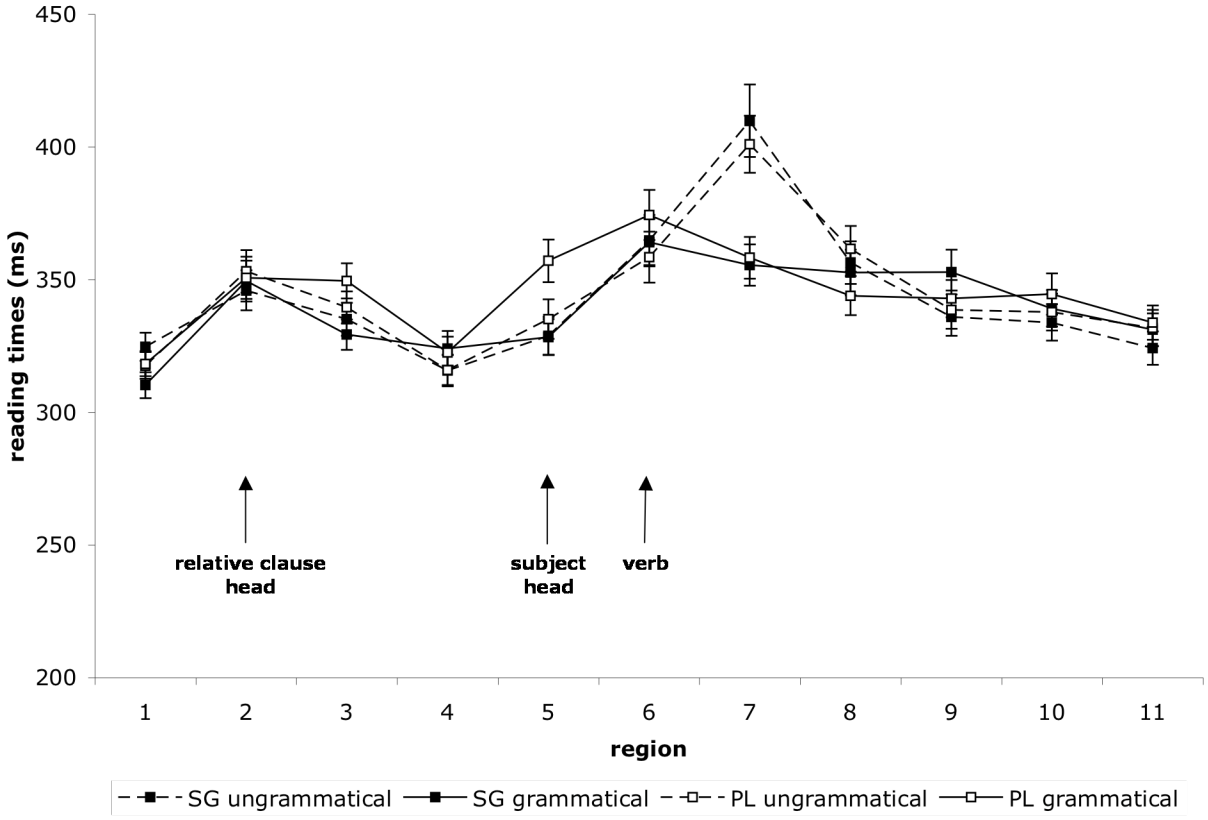

Figure 4 **Experiment 3 Self-paced Reading Results: Plural Relative Clause Subjects**

Region by region means segregated by relative clause head number and grammaticality. Error bars indicate standard error of the mean.

Sample sentence:
The$_1$ musician(s)$_2$ who$_3$ the$_4$ reviewers$_5$ praise(s)$_6$ so$_7$ highly$_8$ will$_9$ probably$_{10}$ win(s)$_{11}$ ...

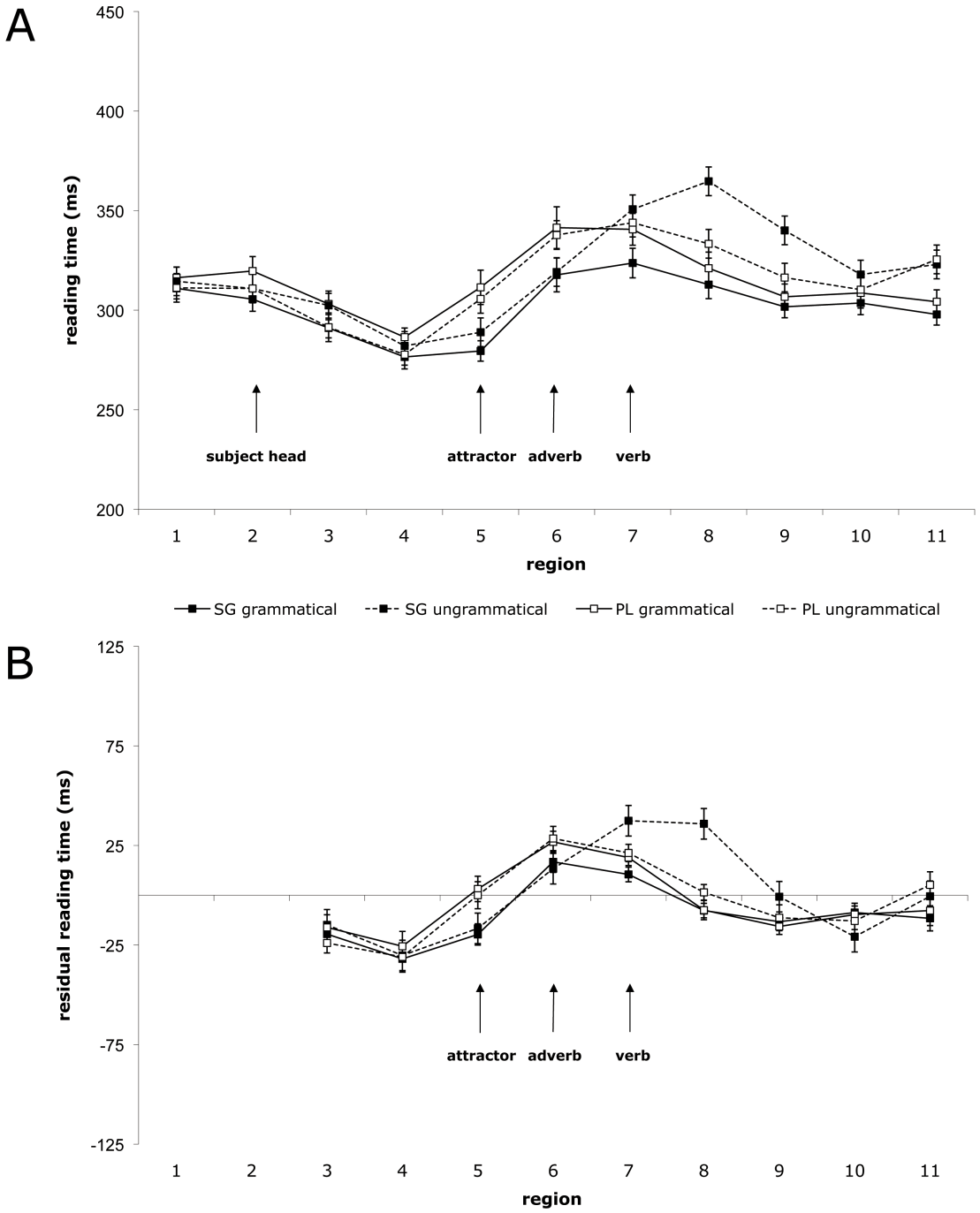

Figure 5 **Experiment 4 Self-paced Reading Results**

Panel A Region by region means segregated by attractor number and grammaticality. Error bars indicate standard error of the mean.

Panel B Region by region residual means based on a mixed-effects model that incorporated previous region reading times.

Sample sentence

The$_1$ key$_2$ to$_3$ the$_4$ cabinets(s)$_5$ unsurprisingly$_6$ was/were$_7$ rusty$_8$ from$_9$ many$_{10}$ years$_{11}$ …

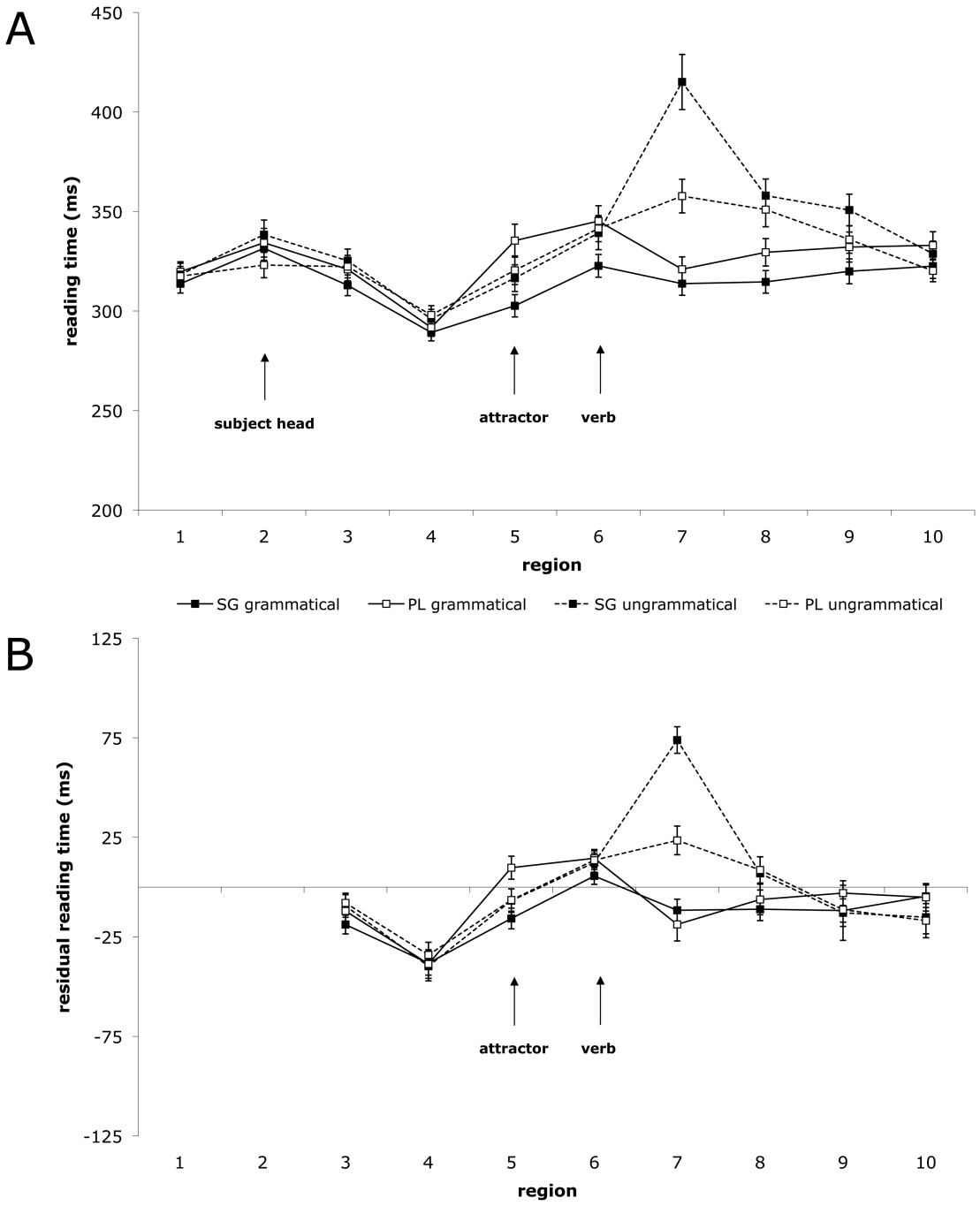

Figure 6 **Experiment 5 Self-paced Reading Results**

Panel A Region by region means segregated by attractor number and grammaticality. Error bars indicate standard error of the mean.

Panel B Region by region residual means based on a mixed-effects model that incorporated previous region reading times.

Sample sentence

The$_1$ key$_2$ to$_3$ the$_4$ cabinets(s)$_5$ was/were$_7$ rusty$_8$ from$_9$ many$_{10}$ years$_{11}$ …

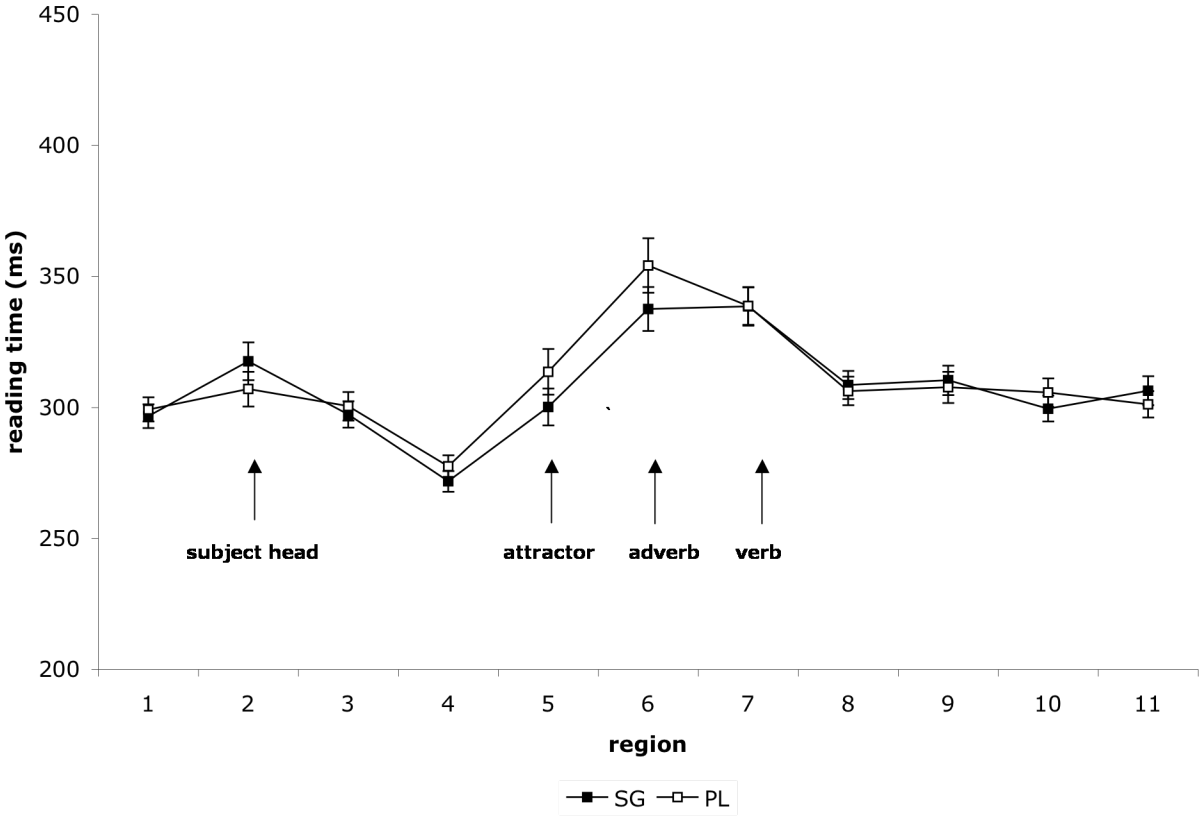

Figure 7 **Experiment 6 Self-paced Reading Results**

Region by region means segregated by attractor number. Error bars indicate standard error of the mean.

Sample sentence:
The$_1$ key$_2$ to$_3$ the$_4$ cabinets(s)$_5$ unsurprisingly$_6$ was$_7$ rusty$_8$ from$_9$ many$_{10}$ years$_{11}$ ...

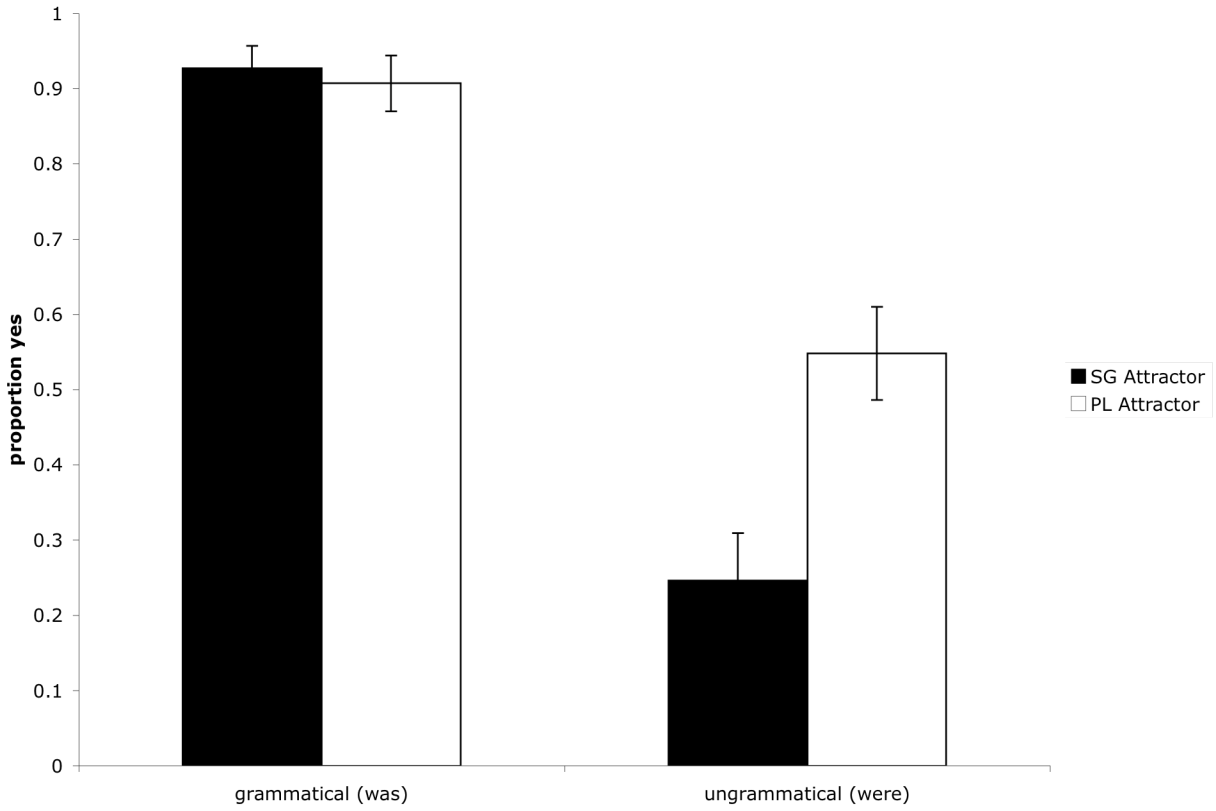

Figure 8 **Experiment 7 Speeded Acceptability Judgment Results**

Mean proportion 'acceptable' responses by grammaticality and attractor number. Error bars indicate standard error of the mean.

Sample sentence:
The key to the cabinets(s) was rusty from many years of disuse.

Supplementary Material
Click here to download Supplementary Material: supplementary.docx

SUPPLEMENTARY MATERIAL to accompany Wagers, Lau, Phillips, "Agreement attraction in comprehension: representations and processes," *Journal of Memory and Language.*

Simulations of the feature percolation symmetry prediction

Feature percolation accounts of agreement attraction predict that ungrammatical sentences containing a plural attractor should sometimes be judged as grammatical, and that grammatical sentences containing a plural attractor should just as often be judged as ungrammatical. In this appendix we test whether this symmetric shift in the perception of grammaticality leads to a symmetric response in mean RTs. Response times have a characteristically positively-skewed distribution and, for this reason, one anonymous reviewer wondered whether the shape of these distributions might mean that underlyingly symmetric shifts in grammaticality would not be reflected by symmetric shifts in mean RTs. Analytically we did not expect this to be the case, as we note on page 31 of the manuscript. However, we also simulated the size and variability of the RT response to attraction in both grammatical and ungrammatical sentences by modeling participant RTs to agreement licensing at the single trial level. The results of 3 simulation scenarios are reported below which show that symmetric changes in the underlying distribution of grammaticality lead to symmetric shifts in the mean.

We defined two random variables, **G**, corresponding to the distribution of RTs resulting from an individual's response to grammatical agreement; and **U**, corresponding to ungrammatical agreement. Both **G** and **U** were generated with ex-Gaussian density functions, a convolution of the normal and exponential density functions which accurately models RT distributions (Luce, 1986). Ex-Gaussian distributions have three parameters: μ, the normal mean; σ, the normal variance; and τ, the exponential mean. The mean of the ex-Gaussian is simply $\mu + \tau$; and its variance $\sigma^2 + \tau^2$. A population of fifty-six participants was created whose parameters μ, σ, τ were drawn from a normal distribution, constrained in particular ways in each of the three simulation scenarios below. For each participant, 24 experimental trials were simulated, corresponding to six trials for each of the four conditions in our Experiment 4: Sg attractor, Grammatical and Ungrammatical; Pl attractor, Grammatical and Ungrammatical. For trials involving a singular attractor, the RT for a single trial was drawn from that participant's **G**, if agreement was grammatical, or **U**, if it was ungrammatical. For trials involving a plural attractor, **G** and **U** were sampled at the rate of percolation, *p*. For grammatical trials, RTs were drawn from **U** *p*% of the time; for ungrammatical trials, RTs were drawn from **G** *p%* of the time.

Because the parameter space for this simulation was so large, we had to restrict our consideration to three scenarios which we considered most germane to the prediction of symmetry. In simulation one, the population parameters are estimated from actual experimental data, and a range of percolation rates is considered. In simulation two, percolation rate is held constant, and the mean parameters are manipulated to assess the effect of concentrating the grammaticality effect in the tail of the distribution. In simulation three, low percolation rates are examined. For each simulation scenario, 100 experimental runs were generated (with a new population of participants each time). The mean and standard deviation over the difference between attractor conditions is summarized below.

Simulation 1 In simulation one, population parameters for **G** and **U** were estimated from RT data collected in Experiment 5, Region 7. An ex-Gaussian function was fit to both the

singular-attractor, grammatical aggregate data, to estimate the population mean of **G**; and the singular-attractor ungrammatical aggregate data, to estimate the population mean of **U**. Variation in the population parameters was set at 20 ms standard deviation for μ and τ and 10 ms for σ. It would have been desirable to estimate distribution parameters for each participant, and thus empirically estimate both the location and the shape of the parameter distributions. Unfortunately, the number of trials collected per experimental condition in our experiment (6) and in most sentence processing experiments we are aware of are simply too few to collect robust estimates.

Five percolation rates (p) were considered, from 0.10 to 1.0. Table I reports the mean RTs for Sg attractor conditions (reported once, as they do not change with percolation rate), and then the differences between those conditions and the attractor conditions for each of the percolation rates. The results are clear: the mean RT for Pl attractor grammatical condition shifts 1:1 with the mean RT for Pl attractor ungrammatical conditions. The size of this shift corresponds to approximately p times the difference in means between **G** and **U**. Note that the distribution of ungrammatical mean differences does show greater variability. This is expected since variation in **U** is proportional to τ^2 and correspondingly the difference between variability in grammatical and ungrammatical effects depends on the number of observations drawn from **U** (i.e., p for grammatical comparisons, and (1-p) for ungrammatical ones).

Based on this simulation (and the estimates of Pearlmutter, Garnsey, & Bock, 1999), the percolation rate would have to be larger in comprehension than in production, to account for the size of the mean RT shifts observed in our experiments 4 and 5.

Simulation 1 Parameters estimated from Experiment 5, Region 7, group distributions *Experimental runs:* 100	**Mean simulation parameters** (s.d. in parentheses) *per run:* 56 participants, 6 trials/per condition				
	Grammatical RT **μ:** 218 ms (20 ms) **σ:** 47 ms (10 ms) **τ:** 96 ms (20 ms)		**Ungrammatical RT** **μ:** 181 ms (20 ms) **σ:** 21 ms (10 ms) **τ:** 229 ms (20 ms)		
Basic grammaticality effect [Sg Sg] Grammatical: 314 ± 6 ms [Sg Sg] Ungrammatical: 410 ± 14 ms					
Attraction effects	**Percolation rate (p)**				
	0.10	0.25	0.5	0.75	1.0
Grammatical: [Sg Pl]-[Sg Sg]	11 ms ± 10 ms	25 ms ± 13 ms	49 ms ± 14 ms	72 ms ± 15 ms	99 ms ± 15 ms
Ungrammatical: [Sg Sg]-[Sg Pl]	8 ms ± 22 ms	24 ms ± 19 ms	50 ms ± 19 ms	72 ms ± 17 ms	96 ms ± 15 ms

Table I **Effect of different percolation proportions on mean RT differences**
RT distribution parameters estimated from the experimental group distributions: Experiment 5, Region 7, singular attractor conditions. The standard deviation of simulation means is reported.

Simulation 2 The overall mean of an ex-Gaussian distribution depends on the means of its generating distributions, the normal and exponential means. In this simulation, we held the mean difference between **G** and **U** constant, but investigated the trade-off between locating that difference entirely in the τ parameter, or in the μ parameter. μ_{Gram}, τ_{Gram}, σ and *p* were held constant. $(\Delta\mu + \Delta\tau)$, the difference between **G** and **U** parameters, was constrained to 75 ms, thus holding the basic grammaticality effect constant.

The results are reported in Table II. As in simulation 1, the mean RT difference in both grammatical and ungrammatical attractor conditions is nearly identical, and does not depend on the size of $\Delta\tau$. Across the board the effect is approximately $p \bullet (\Delta\mu + \Delta\tau)$ (= 18.75 ms). For higher $\Delta\tau$, slightly more variability is observed in ungrammatical comparisons.

Simulation 2 Grammaticality effect constrained to be 75 ms *Experimental runs:* 100	**Mean simulation parameters** (s.d. in parentheses) *per run:* 56 participants, 6 trials/per condition $(\Delta\mu + \Delta\tau)$ = 75 ms; μ_{Gram}: 200 ms (20 ms); μ_{Ungram}: $\mu_{Gram} + \Delta\mu$ τ_{Gram}: 50 ms (20 ms); τ_{Gram}: $\tau_{Gram} + \Delta\tau$ σ: 35 ms (10 ms) Percolation rate constant: 0.25

Basic grammaticality effect [Sg Sg] Grammatical: 250 ms ± 3 ms [Sg Sg] Ungrammatical: 326 ± (3-6) ms (higher s.d.s observed for higher τ)						
Attraction effects	$\Delta\tau$					
	75 ms	60 ms	45 ms	30 ms	15 ms	0 ms
Grammatical: [Sg Pl]-[Sg Sg]	19 ms ± 6 ms	20 ms ± 6 ms	20 ms ± 6 ms	17 ms ± 5 ms	18 ms ± 5 ms	19 ms ± 4 ms
Ungrammatical: [Sg Sg]-[Sg Pl]	19 ms ± 10 ms	19 ms ± 10 ms	18 ms ± 7 ms	18 ms ± 6 ms	19 ms ± 7 ms	18 ms ± 5 ms

Table II **Effect of the RT distribution tail parameter on mean RT differences**
The size of the grammaticality effect is held constant at 75 ms by constraining the parameters the μ and τ parameters. Higher $\Delta\tau$s correspond to scenarios in which most of the difference between ungrammatical and ungrammatical distributions in borne by the tail. The standard deviation of simulation means is reported.

It is hard to say what the real balance between $\Delta\tau$ and $\Delta\mu$ might be in the population. A conclusion cannot be drawn from the group distributions. We subsequently examined the Experiment 1 data, in which more trials were collected per condition and consequently more stable parameter fits could be done on a per-subject basis. There we found great variability in whether ungrammatical conditions led to shifts in μ, τ, or both. It therefore seems reasonable to suppose that, at the individual level, **G** and **U** may differ in both parameters and that ungrammatical responses are not confined entirely to the tail.

Simulation 3 The previous two simulations revealed that ungrammatical attraction effects are the same in size to grammatical attraction effects. For low percolation rates, or high $\Delta\tau$, there is greater variability in the ungrammatical effect size. In this final simulation, we consider only low percolation rates ($p \leq 0.25$) to better assess the difference in variability. μ and σ are held constant: the difference between **G** and **U** is modeled entirely as a difference in τ, which is set here at 75 ms.

Table III reports the results. As is expected, the mean differences between grammatical and ungrammatical conditions are virtually identical. For $p \leq 0.10$, variability of in the population mean for ungrammatical comparisons is roughly 3 times higher than for grammatical comparisons; the ratio drops to 2 for $0.10 \leq p \leq 0.25$.

Simulation 3 Tail-shift only, low mixing proportions *Experimental runs:* 100	**Mean simulation parameters** (s.d. in parentheses) *per experimental run:* 56 participants, 6 trials/per condition **μ:** 200 ms (20 ms) **τ_{Gram}:** = 50 ms (20 ms); **τ_{Ungram}:** = 125 ms (20 ms) **σ:** 35 ms (10 ms)

Basic grammaticality effect
[Sg Sg] Grammatical: 250 ms ± 3 ms
[Sg Sg] Ungrammatical: 325 ± 7 ms

Attraction effects	**Percolation rate**					
	0	0.05	0.10	0.15	0.20	0.25
Grammatical: [Sg Pl]-[Sg Sg]	0 ms ± 4 ms	8 ms ± 6 ms	14 ms ± 6 ms	10 ms ± 6 ms	14 ms ± 5 ms	18 ms ± 6 ms
Ungrammatical: [Sg Sg]-[Sg Pl]	-1 ms ± 11 ms	6 ms ± 18 ms	16 ms ± 18 ms	10 ms ± 11 ms	15 ms ± 10 ms	18 ms ± 10 ms

Table III **Effect of low percolation proportions on mean RT differences**

The relationship between percolation proportion and mean RT differences is simulated. The difference between grammatical and ungrammatical distributions is borne entirely by the tail of the distribution. The standard deviation of simulation means is reported.

Summary The results of the three simulations reported show that, despite the positive skew in RT distributions, an underlying symmetric shift in distributions of grammaticality leads to a symmetric shift in mean RTs. Therefore, the prediction of symmetric RT shifts in the self-paced reading experiments in our study defines a fair test of feature percolation accounts.

For low values of p or high values of $\Delta\tau$, we did find greater variability in the ungrammatical comparisons. This raises the possibility that finding an asymmetry would be not be inconsistent with the feature percolation account (in the sense that, the wider the distribution of effect sizes, the more likely they will be discrepant). However it is important to note three things: First, it is as likely we would observe smaller effects for ungrammatical comparisons, as for larger ones. Secondly, the lower variability in the grammatical effect sizes leads us to have more confidence in expecting a non-null result for grammatical comparisons than for ungrammatical ones. Thirdly, the size of the actual ungrammatical attractor effects we observe in Experiments 2-5 and 7 suggests that the percolation rate cannot be very low.