What counts in grammatical number agreement?

Laurel Brehm *, Kathryn Bock

Department of Psychology, University of Illinois at Urbana-Champaign, USA
Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, USA

ARTICLE INFO

ABSTRACT

Both notional and grammatical number affect agreement during language production. To explore their workings, we investigated how semantic integration, a type of conceptual relatedness, produces variations in agreement (Solomon & Pearlmutter, 2004). These agreement variations are open to competing notional and lexical–grammatical number accounts. The notional hypothesis is that changes in number agreement reflect differences in referential coherence: More coherence yields more singularity. The lexical–grammatical hypothesis is that changes in agreement arise from competition between nouns differing in grammatical number: More competition yields more plurality. These hypotheses make opposing predictions about semantic integration. On the notional hypothesis, semantic integration promotes *singular* agreement. On the lexical–grammatical hypothesis, semantic integration promotes *plural* agreement. We tested these hypotheses with agreement elicitation tasks in two experiments. Both experiments supported the notional hypothesis, with semantic integration creating faster and more frequent singular agreement. This implies that referential coherence mediates the effect of semantic integration on number agreement.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Many things begin with an idea, speech included. Before we talk, we usually have a thought to express, a conceptualization of an event we want to recount, or a desire to make something happen. The nonlinguistic ideas that give rise to spoken sentences are sometimes called *messages*. Messages are a sort of distillation from mental models of how referents relate to each other in causality, intentionality, space, or time (e.g. Johnson-Laird, 1983), along with other conceptual and pragmatic properties that drive what we say. Among these properties is information about the *notional number*, or numerosity, of particular referents. In many languages, English included, notional number is vital even when it is immaterial to a speaker's communicative intention, because the syntax of sentences requires it for setting *grammatical number*: Nouns and other words must be grammatically specified as either singular (e.g. *dog*) or plural (e.g. *dogs*), with no in-between options. Grammatical number in turn controls the necessary operations of subject–verb agreement that speakers duly implement. The question in the present investigation was how variations in the mental representations of numerosity affect the creation of number agreement during language production.

The question is not easy to answer, because notional number does not perfectly predict grammatical number. Notional number discriminates singletons (single things) from aggregates (multiple things), but this is not always reflected in grammatical number. Objects called *clothing* can be notionally plural, but the word *clothing* is grammatically singular. An object called *scissors* is notionally singular, but the word *scissors* is grammatically plural. These grammatical specifications are what matter most in subject–verb agreement, with the grammatical singular–plural distinction for nouns tied to the grammatical singular–plural

* Corresponding author. Address: Beckman Institute, 405 N Mathews Avenue, University of Illinois, Urbana, IL 61801-2300, USA. Tel.: +1 217 244 5494.

E-mail addresses: lbrehm3@illinois.edu (L. Brehm), jkbock@illinois.edu (K. Bock).

distinction for verbs (e.g. singular *is*, plural *are*). So, English speakers use singular verbs with nouns like *clothing* and plural verbs with nouns like *scissors*, even though the notional number of the nouns' referents is typically the reverse.

There are other wrinkles in the relationship between notional and grammatical number that complicate agreement considerably. Although the grammatical number of a subject noun (the *head* noun) is a highly reliable predictor of the form of an agreeing verb, there are occasions when other factors seem to determine the number of the verb. Two of these factors have been central to psycholinguistic research on the formulation of sentence structure. First, notional number sometimes trumps grammatical number, creating notional agreement (e.g. Bock, Carreiras, & Meseguer, 2012; Eberhard, 1999; Humphreys & Bock, 2005; Vigliocco, Butterworth, & Semenza, 1995; Vigliocco & Hartsuiker, 2002). This occurs in conjunctions, as seen in examples like *Her sister and best friend*. If the sister and best friend are two separate people, the phrase promotes plural verb agreement when it serves as a sentence subject (e.g. *Her sister and best friend were maids of honor at her wedding*). In contrast, if one person plays both roles then the phrase will promote singular verb agreement (. . . *was maid of honor at her wedding*). Notional information can override other grammatical agreement features, too, including gender (Franck, Vigliocco, Antón-Méndez, Collina, & Frauenfelder, 2008; Vigliocco & Franck, 1999, 2001). Second, sometimes the grammatical number of a noun other than the head noun affects verb number, as in the phenomenon of attraction (e.g. Bock & Miller, 1991). An ironic instance is "How much correction of syntactic errors *are* there, anyway?" (Bock, 2011). Here, the verb's number reflects the local plural word *errors* and not the singular head noun, *correction*.

The relationships among grammatical agreement, notional agreement, and attraction are central to the present research. One account of how they fit together is called *Marking and Morphing* (Eberhard, Cutting, & Bock, 2005). A premise of Marking and Morphing is that the referents people have in mind when they produce sentences can be construed in ways that are more or less notionally plural. The consequence is that there is sometimes uncertainty and ambiguity in the construal of numerosity. This uncertainty and ambiguity is present with *Complex Reference Objects* in semantic theory (cf. Barker, 1992; Eschenbach, Habel, Herweg, & Rehkämper, 1989). A Complex Reference Object is a set of individuated elements that can be interpreted as either singular (the set as a whole) or plural (as the individuals in the set; e.g. Albrecht & Clifton, 1998; Kaup, Kelter, & Habel, 2002; Moxey, Sanford, Sturt, & Morrow, 2004; Patson & Ferreira, 2009; Patson & Warren, 2011; Sanford & Lockhart, 1990). In the Marking and Morphing account of agreement, a probabilistic construal of referent numerosity is responsible for a notional contribution to the number of a sentence's subject (its *marking*) that in turn can affect the grammatical number properties of the subject's agreement targets (e.g. verb morphology). This can result in notional number skewing the outcome of the verb agreement process. (Note that Marking and Morphing provides an explanation of

pronoun agreement using the same mechanisms, but with an outcome where notional number dominates.)

Attraction, according to Marking and Morphing, is an aberrant product of the normal process that puts the grammatical number of the subject into an agreement relationship with a verb. When a conventionally singular subject is accompanied by a plural noun phrase (as in *correction of syntactic errors*), the agreeing verb may surface as a plural rather than the expected singular. The hypothesized source of this apparently spurious agreement is a change in the number features of the subject due to an intrusion of the plural feature into the calculation of what would otherwise be a singular.

Although the source of attraction is different from the source of notionally driven number agreement, the outcomes can look the same. In the sentence *The picture on the postcards were ugly*, the plural verb could reflect uncertainty in referent construal that is resolved in favor of multiple picture tokens (on the notionally plural postcard objects) rather than the unique picture type (the single duplicated image). Alternatively or additionally, the verb could reflect attraction from the plural noun *postcards*. In both cases, the grammatically relevant number of the subject noun phrase becomes plural, and the process of formulating or retrieving the plural verb (*morphing*) unfolds in the same way. This makes it possible to explain apparent instances of attraction as notional agreement, and apparent instances of notional agreement as attraction.

This duality forms the crux of the present research. The aim was to assess whether notional agreement or attraction is at work in an effect of *semantic integration* on agreement (Solomon & Pearlmutter, 2004). As defined by Solomon and Pearlmutter, semantic integration is a property of linkages among parts of a to-be-produced message, with variations in semantic integration (linkage strength) corresponding to coherence in the representation of elements in a speaker's mental model or discourse representation. From this we infer that integration can affect how items are mentally grouped, individuated, and enumerated. In semantic terms, the linkages create Complex Reference Objects. Solomon and Pearlmutter found variations in plural agreement associated with variations in semantic integration, and attributed this association to the role of semantic integration in attraction, not to notional number agreement. For reasons that we develop in upcoming sections, notional agreement makes a different prediction about the effect of semantic integration on agreement. The question is, when semantic integration affects number agreement, what counts?

1.1. How does semantic integration affect number agreement?

Fig. 1 illustrates the difference between the states of affairs that Solomon and Pearlmutter treated as more and less semantically integrated. In the integrated case (panel A), there is a representational relationship between a drawing and some flowers (*the drawing of the flowers*) and in the unintegrated case (panel B) the drawing and the flowers are merely contiguous in space (*the drawing with the flowers*).
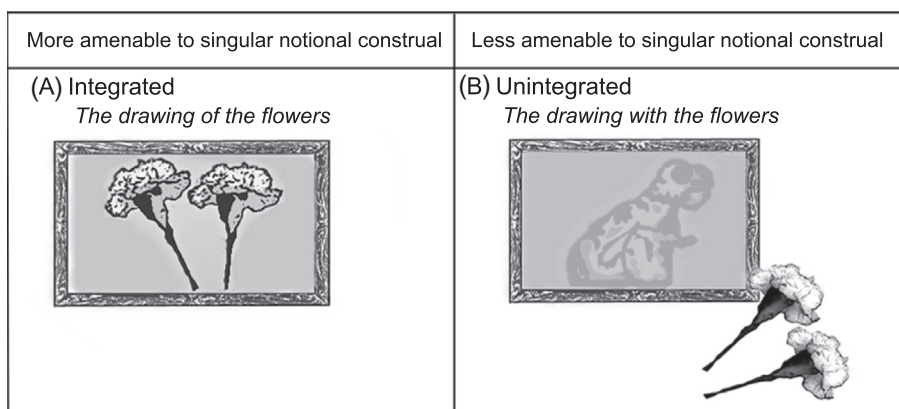
More amenable to singular notional construal	Less amenable to singular notional construal
(A) Integrated *The drawing of the flowers*	(B) Unintegrated *The drawing with the flowers*

Fig. 1. Examples of notional number variations due to semantic integration.

Table 1
Example stimuli from Solomon and Pearlmutter (2004).

Integrated	Unintegrated
The drawing of the flower(s)	The drawing with the flower(s)
The translator of the ambassador(s)	The translator with the ambassador(s)
The apple with the brown spot(s)	The apple with the fresh peach(es)
The report that described the traffic accident(s)	The report that Megan described the accident(s)

Solomon and Pearlmutter (2004) proposed that during sentence formulation, semantic integration affects the timing of lexical selection and retrieval during phrase construction. The resulting variations in timing have consequences for agreement, and in particular for the occurrence of attraction. Specifically, sentence subjects that express strongly integrated relationships (e.g. *the drawing of the flowers*) may be formulated from simultaneously active elements (e.g. *drawing*, *flowers*). Their representational integration creates more parallelism during the formulation of the subject's words (lexical selection, lexical retrieval) and phrase structure. One of the potential byproducts of parallelism is interference. With singular and plural words or phrases in the pipeline at the same time, competition between them allows the plural constituent to take control of the verb's number. The result can be an utterance like *The drawing of the flowers **were** ugly*. This is attraction. Since parallel retrieval is less likely when integration is weak, the probability of attraction goes down. The prediction, then, is that strong integration leads to more attraction than weak integration does.

Solomon and Pearlmutter's (2004) evidence for the lexical interference hypothesis came from a set of language production experiments that manipulated semantic integration in various ways. In all of the experiments, speakers produced sentence completions for subjects that varied in integration (see Table 1), and the grammatical number (singular or plural) of the verbs that the speakers used in their completions was examined. All of the experiments supported the interference hypothesis: sentence subjects like *the drawing of the flowers* were more likely to elicit attraction than subjects like *the drawing with the flowers*. In short, strong semantic integration was associated with more plural agreement than weak semantic integration, and this outcome was explained as increased attraction.

The weakness in this account stems from a general problem in research on language production. Because speakers sometimes failed to produce legitimate responses, the base rates of identifiably singular or plural agreement in the critical conditions were unequal. Across the experiments, 29% of the trials yielded uninterpretable responses (no response at all or inaccurate production of the intended subject). Of these problematic responses, nearly three-fifths occurred when sentence subjects were unintegrated, and over half of those losses were for subjects with singular heads and plural local nouns (the subjects designed to elicit attraction). As a result, interpretable agreement outcomes were more likely for integrated than for unintegrated subjects, and especially so in the attraction-inducing conditions. The upshot is that integration covaried with attraction opportunities. Because of this base-rate problem, there may have been more attraction with integrated than with unintegrated subjects because there were more occasions when attraction could be observed. This weakens the conclusion that integration affects attraction, and in particular that strong integration leads to more attraction.

Calling the conclusion further into doubt is the paradox of the association between attraction and strong integration. Attraction is spurious plural agreement. But integration, as we see it, creates representational unity, and unity is singular. Singular agreement is thus a plausible repercussion of integration, and is what would be predicted from a notional number effect. In explicit terms, stronger integration should yield more singular construals and more singular agreement, and weaker integration should yield more aggregate construals and more plural agreement. With pronoun agreement, this is the result that has been found for Complex Reference Objects in language comprehension (e.g. Eschenbach et al., 1989; Patson & Warren, 2011) and language production (Bock, Eberhard, & Cutting, 2004), where analogs of weak integration promote plural number.

Marking and Morphing provides a foundation for the hypothesis that variations in notional number should create an effect of semantic integration on agreement. Fig. 2 illustrates the difference between the two accounts. The notional number variations associated with differences in semantic integration affect abstract number features (the number marking) of the subject noun-phrase, which in turn affect the grammatical specifications for agreement. The weaker the integration relationship (e.g. in the state of affairs described by *the drawing with the flowers*), the more likely the notional construal is to be plural. If attraction arises (due to lexical interference or some other disruption), its impact on verb number occurs when lexical and structural processes come together.

Notional plurality also adds uncertainty and difficulty to agreement. It is semantically complex (e.g. Krifka, 1989; Landman, 1989; Schwarzschild, 1992), being associated with a marked state that triggers a change in the subject from an unmarked singular value to a plural. Notional singularity is simple in comparison and can be modeled as a null or default value, extrapolating from a universal property of singular grammatical number (Greenberg, 1966). Developmentally, its semantic simplicity is reflected in the whole-object bias for word learning: Children assume that words refer to unitized entities rather than pieces within the conglomeration of entity-parts (e.g., Bloom, 2001; Markman, 1990). The default nature of singularity allows singular grammatical number to be an automatic consequence of singular notional number.

Assuming that strong integration promotes a singular construal of a complex referent and weak integration promotes a plural construal, the prediction is that problems in the formulation of agreement should be more likely to arise from weak than from strong semantic integration: Strong semantic integration (e.g., the referent of *the drawing of the flowers*) makes agreement easier. This prediction is the polar opposite of the interference hypothesis. When semantic integration *increases*, the notional hypothesis predicts higher rates of singular agreement, while the interference hypothesis predicts increases in plural agreement (as a result of attraction). As semantic integration *decreases*, the notional hypothesis predicts higher rates of plural agreement due to plural marking, while the interference hypothesis predicts lower rates of plural agreement due to fewer cases of attraction.

To pit these predictions against one another, Experiment 1 used a timed agreement production task that was designed to elicit high rates of correct (singular) agreement, allowing more powerful assessment of the underlying timing variations in agreement implementation. In the experiment, we measured the latency of initiating sentence completions after subject noun-phrases, at the point when speakers began to produce a number-specified verb. The produced verb was nearly always a singular or plural form of *to be* (*is*, *was*, *are*, *were*). The experimental materials were the same as in Solomon and Pearlmutter (2004), but the task made uniformly scorable verbs common enough to prevent covariations between the likelihood of responding and the likelihood of attraction.

The primary measure, latency to produce a number-agreeing verb, also directly taps the hypothesized consequences of semantic integration in a simple form of the
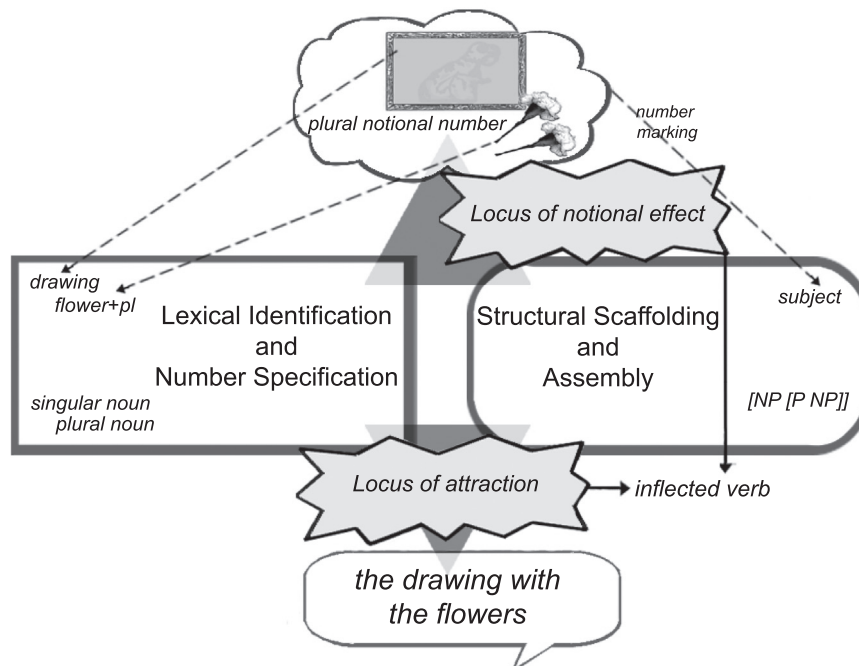

Fig. 2. Hypothesized sources of notional number and attraction effects on agreement due to semantic integration. In this sketch, information (dashed lines) passes from a conceptual level to number formulation processes. Plural notional number affects the abstract structural subject's number; grammatically plural nouns elsewhere in the sentence affect attraction. These two sources for a plural-inflected verb are represented by the solid arrows.

lexical interference account. If strong integration leads to parallel retrieval of subject nouns along with competition for control of number agreement, verb production should be slower for semantically integrated than for semantically unintegrated subjects, and the slowing of utterance production should be especially pronounced for singular verbs after plural local nouns, due to the difficulty caused by attraction. That is, lexical interference predicts an interaction between local plurality and integration, such that the difference in latencies between local plurals and singulars is largest in the integrated condition. The notional-number hypothesis instead predicts faster verb production for integrated than for unintegrated subjects, because integration makes agreement easier. Plural local nouns should slow verb production, due to the force of attraction, but to the same extent after integrated and unintegrated subjects. So, the notional-number hypothesis predicts a main effect of integration, such that integrated preambles are faster then unintegrated preambles, with no interaction between local plurality and integration. If both forces are in play, an overall benefit of strong integration on response latency should be accompanied by the larger difference between local plurals and singulars predicted from lexical interference.

Experiment 2 was aimed at extending Experiment 1 by eliciting overt attraction, as captured in the verb-number variations that speakers produced. The procedure and dependent measures were identical to those of Solomon and Pearlmutter (2004). Speakers repeated the subject noun phrases and completed them as full sentences, allowing assessment of the incidence of plural verbs.

In both experiments, the experimental materials were the same as in Solomon and Pearlmutter (2004). They consisted of semantically integrated and unintegrated subject noun-phrases, presented to participants as sentence preambles. All experimental preambles contained a singular head noun followed by a prepositional phrase, relative clause, or subordinate clause that served to modify the head. The modifying material included another noun (the *local* noun), which immediately preceded the verb. These local nouns varied in grammatical number in both of the semantic integration conditions. The simple predictions in all cases have to do with the presence or absence of interactions between local noun number and semantic integration. Relative to the singular local noun conditions, the lexical interference hypothesis predicts slower completions after local plurals for semantically integrated than for semantically unintegrated sentence subjects, and a higher incidence of plural completions. The notional number hypothesis predicts faster completions for integrated than unintegrated sentence subjects, and more singular agreement for integrated than for unintegrated subjects, regardless of local noun number.

2. Experiment 1

The first experiment was designed to examine the time course of number agreement with semantically integrated and unintegrated subjects. Verb production should be slowed under the circumstances that typically yield errors in conventional number agreement (specifically errors of attraction), even when attraction does not actually surface in speakers' utterances (cf. Staub, 2009). This reflects a common correlation between increases in correct response time and errors in the face of processing difficulty. We measured the latency to produce number-agreeing verbs using the same sentence-eliciting preambles as Solomon and Pearlmutter (2004), thereby maintaining exactly the same variations in semantic integration.

In the experiment, participants were asked to complete the preambles as full sentences, using one of four designated adjectives in their completions. This task elicits number-inflected verbs reliably but incidentally, with no explicit instructions about verb production needed for a wide majority of responses to have the conventionally correct grammatical number (i.e., singular).

2.1. Method

2.1.1. Participants

In exchange for course credit or $7.00 compensation, 94 undergraduates at the University of Illinois participated in the experiment. Participants with fewer than 80% usable experimental trials were excluded (*N* = 21; see Section 2.1.5 for criteria), and one participant was excluded due to recording failure.

2.1.2. Equipment

Stimuli were presented using PsyScope X B53 (Cohen, MacWhinney, Flatt, & Provost, 1993) on a Macintosh Mini computer with a 17-in. LCD flat-screen monitor. Audio was recorded with a Sennheiser directional microphone run through a USB button box and Tube MP preamplifier. The button box also recorded the latency of vocal responses.

2.1.3. Materials

The 100 experimental items were taken from Solomon and Pearlmutter (2004); (see Table 1 for examples). Following Solomon and Pearlmutter, there were four categories of items including 24 representational items, 24 attribute/accompaniment items, 32 relative clause/subordinate clause items, and 20 functional items. Every item had four versions, two integrated and two unintegrated, with the two versions of each integration type having different local noun numbers, singular and plural. For the functional items, Solomon and Pearlmutter created alternative integrated preambles that behaved similarly in their experiments, so we quasi-randomly selected one of the two integrated forms for each of the 20 items, balancing the incidence of the two original categories (*for* and *of*).

There were 156 filler stimuli. These were designed to increase the variety of grammatical structures in the experiment, to vary the positioning of plural and singular nouns in the sentence stems, and to balance the number of plural and singular subjects of the sentences. The fillers contained a variety of types of noun phrases (simple and complex), with singular and plural nouns in all structural positions. Counting the fillers, 114 (46%) preambles had a singular subject noun phrase and 138 (54%) had a plural subject noun phrase. Thirty of the fillers were placed at
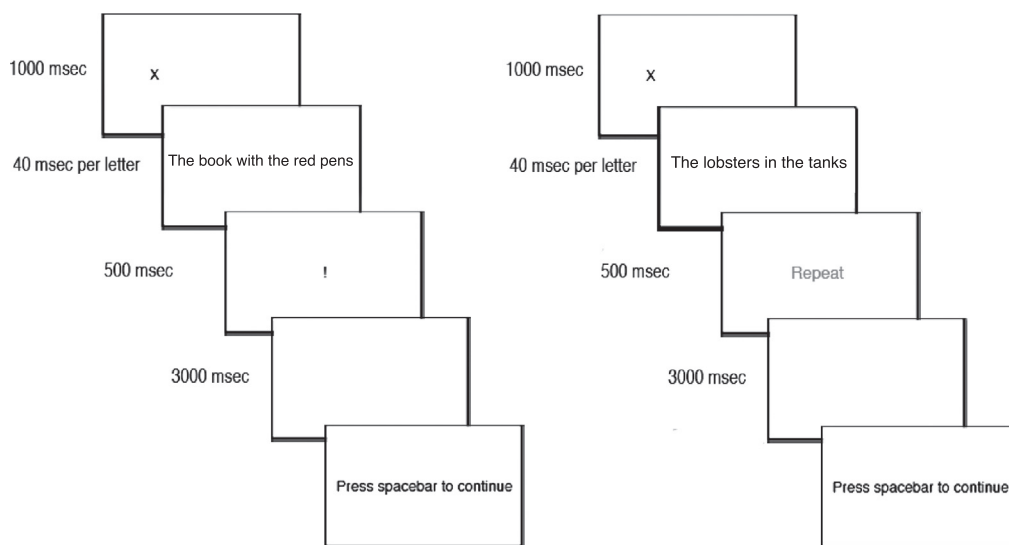
Fig. 3. Trial time-course (standard trial on left; catch trial on right).

the beginning of every experimental list as a covert practice block.

We selected four adjectives to provide plausible completions to all sentence preambles: *good*, *bad*, *ready* and *true*. To ensure that the adjective set allowed for plausible completions of the sentence stems, we collected norming data from 48 people. Participants were asked to rate on a seven-point scale the plausibility of the singular versions (integrated and unintegrated) of each experimental preamble paired with the verb *was* and each of the adjectives (e.g., *The book with the red pen was good*). The eight versions of each experimental item (two integration types by four adjectives) were divided equally among eight lists, so that participants viewed only one version of each item and viewed equal numbers of preambles with each adjective. All filler items were used in the lists along with adjectives from the set. (See Appendix A for norming instructions.) We then calculated the mean plausibility rating across preamble–adjective pairings, as well as the mean rating of the best-fitting (most plausible) adjective for each preamble. Across all pairings, the average rating was 4.66 (4.82 and 4.49 for the integrated and unintegrated preambles respectively), and the average most-plausible rating was 6.11 (6.28 integrated; 5.95 unintegrated). The most-plausible averages imply that at least one adjective offered a sensible predication for the majority of the preambles. Appendix B lists the plausibility ratings by individual preamble.

For the experimental lists, the four versions of each of the items were counterbalanced over four lists such that all lists contained one version of every item, with an equal number of items within each condition. Order within lists was quasi-random, constrained so that no more than two experimental preambles and no semantically similar items appeared consecutively. Filler trials were positioned in quasi-random positions, fixed across lists. Finally, every list

was divided into two halves, with the order of the halves counterbalanced over participants.

2.1.4. Procedure

Fig. 3 displays the time course of events on each trial. There were two types of trials. In the standard trial sequence (on the left in Fig. 3) the first event was a fixation cross on the left side of the screen, presented for 500 ms. Then the preamble appeared briefly, displayed for the greater of 1000 ms or 40 ms per character. Immediately after the preamble, the cue "!" appeared for 500 ms, prompting participants to speak. A blank screen then appeared for 2 s, giving participants a total of 2.5 s to produce a response.

The second type of trial sequence is illustrated on the right in Fig. 3. These *catch* trials were designed to maintain participants' attention to the preambles. On a catch trial, the "!" cue was replaced with the word "Repeat," shown for 500 ms. When this happened, participants were to repeat back the preamble with an ending. Otherwise, the sequence was the same as for the standard trials. All catch-trial preambles were filler items. Catch trials were pseudorandomly distributed so as to occur evenly across both halves of the experiment and never appeared consecutively.

Preambles were presented on the monitor in 36-point black Arial font on a white background, with the first character presented 10% of the width from the left margin of the screen. Participants were instructed to make these preambles into complete sentences as quickly as possible with an adjective (*good*, *bad*, *ready*, or *true*), using the most sensible choice. Participants received two explicit practice trials and were queried about the procedure before beginning the experiment. As in Solomon and Pearlmutter's studies (N. Pearlmutter & M. Gillespie, personal communication, March 3, 2011), the experimenter remained in the room

for the entire session. Appendix A gives the complete instructions.

2.1.5. Scoring

Responses used in the reaction-time analyses excluded the catch trials. Responses on all other experimental trials were scored as valid, miscellaneous, or missing, and valid trials were then scored as singular or plural with respect to verb number. To be valid, a response had to include one of the four adjectives and the verb *was*, *were*, *is*, or *are*, with the verb produced as the first word in the completion, after the cue to speak, and without disfluencies (including verb repetitions or filled pauses) or non-speech noises preceding the verb. No overtly inflected non-copular verbs occurred. Valid completions were scored as singular when the verb was *was* or *is* (6406 trials, 93%) and as plural when the verb was *were* or *are* (488 trials, 7%). Only singular responses were included in the statistical analyses of reaction time.

Scoring of verb number in the spoken completions was carried out on transcriptions of the responses. This scoring encompassed all completions where the first verb produced in the utterance was a number-specifying verb. There were 7188 trials (99%) that met this criterion. Singular verbs occurred on 6644 trials (93%) and plural verbs on 511 trials (7%).

2.1.6. Design and data analysis

Every participant received exactly one version of each of the 100 experimental preambles, 25 preambles in each of the four combinations of integration and local–noun number. Every item was presented to 18 participants in each of the cells of the design. The fixed effects in the statistical analyses were integration (integrated–unintegrated), local-noun number (singular–plural), and their interaction.

The dependent variable in the reaction time analysis of the singular verbs was the amount of time taken to initiate the verb from the onset of cue presentation to the onset of speech. Reaction times for plural verbs were not included in the statistical analyses because there were too few plural responses for meaningful interpretation. As in Solomon and Pearlmutter (2004), the dependent variable in the plural response analysis was the likelihood of a plural response on each trial, out of all number-inflected responses.

In addition to treating integration as a dichotomous independent variable, we used a graduated five-level measure of integration for each experimental preamble in order to get a finer-grained picture of integration effects. Solomon and Pearlmutter (2004) used average integration ratings to confirm a priori assignments of preambles to the integrated and unintegrated conditions, rather than making the assignments on the basis of the ratings themselves. One consequence was that rated integration sometimes departed from the a priori classifications, with some items in the integrated conditions having lower ratings than items in the unintegrated conditions (see Appendix B).

To align the effects of a priori integration with rated integration, individual preamble ratings (downloaded from http://www.psych.neu.edu/faculty/n.pearlmutter/lab/papers.html) were binned after rounding each rating to the nearest integer on the 1–7 scale. To balance the number

Table 2
Graduated integration classification of experimental items.

Graduated classification	Integration rating range	Total item versions	Dichotomous integration classification	Versions
Highest	5.51–7	46	Integrated	38
			Unintegrated	8
High	4.51–5	78	Integrated	54
			Unintegrated	24
Mid	3.51–4.50	121	Integrated	61
			Unintegrated	60
Low	2.51–3.50	101	Integrated	34
			Unintegrated	67
Lowest	1–2.5	54	Integrated	13
			Unintegrated	41

of observations at each level of integration, we combined the lowest and second-lowest bins (with a range from 1 to 2.5) and the highest and second-highest bins (ranging from 5.51 to 7). All bins contained preambles from Solomon and Pearlmutter's two a priori integration classifications, distributed as shown in Table 2. There were roughly equal numbers of items per bin containing singular and plural local nouns.

Data collected at the two levels of local plurality (singular or plural) and integration (integrated or unintegrated) were contrast-coded with the values 0.5 and −0.5. Data from the five graduated levels of integration were analyzed with Helmert contrasts, weighted to account for the number of responses at each level. This allowed us to test for differences between levels as integration increased. Inferential statistics were calculated using multi-level linear and weighted empirical logistic regression using the lme4 package in R, a statistical programming language and interface (Barr, 2008; Bates, 2005; R Development Core Team, 2005).

2.2. Results

2.2.1. Response latencies

Fig. 4 displays the mean latencies to produce singular verbs. The results with the dichotomous (integrated/unintegrated) classification of integration are in the upper panel. After integrated preambles, verbs were initiated faster than after unintegrated preambles (1037 ms against 1091 ms), and faster after singular local nouns than plural local nouns (1056 vs 1071 ms). Using a multi-level linear regression analysis (Table 3), both of these differences were significant, without a significant interaction between them. The regression model for this analysis had significantly better goodness-of-fit values (lower AIC and BIC values, tested by chi-square analyses) than an alternative that allowed random slopes for item origin by items (where item origin is the experiment in Solomon and Pearlmutter in which the item was used). The implications are that the selected model minimized variance while maximizing simplicity, and that the construct of integration did not depend on item origin.

Plural latencies followed a similar pattern. Latencies were faster after integrated than unintegrated preambles
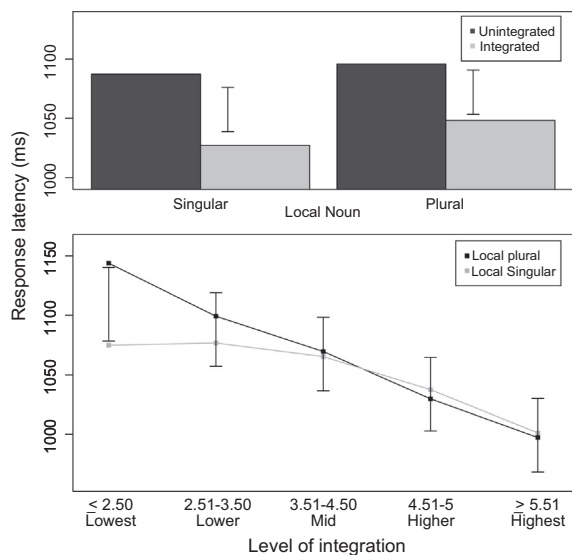
Fig. 4. Experiment 1 singular-verb production latencies (in ms) for preambles differing in semantic integration and local noun number. Upper panel shows mean latencies with the dichotomized classification of integration; lower panel shows latencies with the graduated classification. Error bars are 95% confidence intervals for differences between singular and plural local nouns, calculated from the standard errors of the means.

(709 vs 744 ms), and faster after plural than singular local nouns (711 ms vs 784 ms).

The lower panel of Fig. 4 gives the mean singular latencies using the graduated classification of integration. Overall, high integration and local singularity speeded verb production, but the difference between local nouns depended on integration. More integration elicited faster responses regardless of local noun number, with small and inconsistent differences for local singulars and plurals. At low levels of integration, plural local nouns slowed production progressively as integration decreased. At the same low levels of integration, responses were generally faster after singular than plural local nouns, but response speed changed little as a function of integration after singular local nouns. This is the opposite of what would be expected with lexical interference.

Multi-level linear regression (Table 3) confirmed the relationship between graduated integration and local plurality, with significant differences at lower levels of integration, where responses were slower after plural than after singular local nouns ($p < .05$). The model for this analysis included fixed predictors of local plurality (singular, plural) and integration (five levels spanning lowest through highest), and random participant and item intercepts. This model fit better than alternative models that were evaluated, with goodness-of-fit values (AIC and BIC) significantly better by chi-square analyses than those for an alternative model that allowed random slopes for item origin by items.

2.2.2. Production of singular and plural verbs

Of the verbs produced in valid sentence completions, only 7% had plural verbs, but the responses were distributed differently over conditions. Table 4 shows that among the singular and plural verbs produced, the proportion of plural verbs was higher after plural local nouns than after singular local nouns (.10 vs .04), reflecting attraction. There was likewise a difference between unintegrated and integrated preambles in the proportions of plural verbs, with unintegrated preambles eliciting more plurals (.09 vs .06). Fig. 5 shows that differences in attraction as a function of a priori integration were negligible, though in the same direction as in Solomon and Pearlmutter's studies (.063 and .066 at low and high integration, respectively).

To assess these differences, weighted empirical logit models of the proportions of plural responses were calculated. The logit function can be problematic for error data, as proportions approaching 0 and 1 approach negative and positive infinity respectively. The weighted empirical logit function corrects for this with weights on the cells derived from collapsing across one random factor. Following Barr (2008) and Gillespie and Pearlmutter (in press), two models were calculated, one with random intercepts for participants and another with random intercepts for items. These showed significant effects of local plurality and integration, without significant interaction between them ($ps = .28$ and .14 for participants and items, respectively; see Table 3 for parameter estimates).

Two similar models were constructed for the data broken down by graduated integration, and again the effect of local plurality was significant. In addition, at the lower and mid integration levels, there was more considerably more attraction than at the three other levels, including the two highest ones. The effect surfaced in a pattern of significant interactions between local plurality and integration (Table 3).

2.3. Discussion

The latency measure demonstrated faster production of singular verbs after semantically integrated sentence subjects than unintegrated ones. This penalty for reduced integration occurred after singular as well as plural local nouns, implying that high levels of integration simplified agreement regardless of whether the conditions for attraction were present. Plural local-noun number nonetheless slowed responding more than singular local-noun number, displaying the reaction-time analog of an attraction effect (e.g., Pearlmutter, Garnsey, & Bock, 1999; Staub, 2009; Wagers, Lau, & Phillips, 2009), without an increase due to integration.

The results are the reverse of what the lexical interference hypothesis predicts about the effect of semantic integration, supporting the notional hypothesis instead. Integrated subjects behaved like their referents were notional singletons; unintegrated subjects behaved like their referents were notional plurals, with the agreement complexity that notional plurality creates.

With graduated variations in integration, the distribution of latencies was more complex but similarly inconsistent with the hypothesis that interference emerges as integration increases. There were bigger effects of local noun plurality at lower than at higher levels of integration, and at higher levels of integration the impact of plural local

Table 3

Parameters for Experiments 1 and 2 latency and proportion plural-response models.

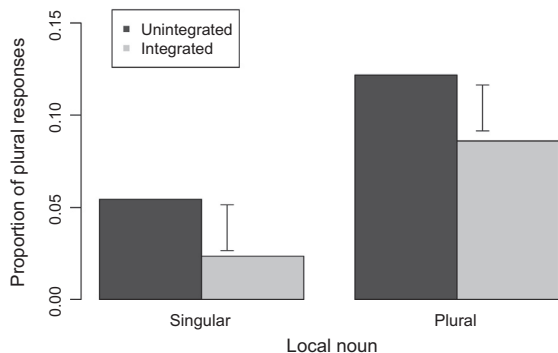| Parameters	Experiment 1 Singular response latencies 95% CI			Experiment 1 Proportion of plural responses						Experiment 2 Proportion of plural responses					
				By participants 95% CI			By items 95% CI			By participants 95% CI			By items 95% CI		
	Estimate	Lower	Upper	Estimate	Lower	Upper	Estimate	Lower	Upper	Estimate	Lower	Upper	Estimate	Lower	Upper
Local plurality × dichotomized integration															
Intercept	1069.06**	1027.4	1108	−1.82**	−1.87	−1.71	−2.28**	−2.17	−1.99	−2.22**	−2.28	−1.92	−1.99**	−1.88	−1.7
Local plurality	−15.66*	−30.9	0.51	0.57**	0.48	0.68	0.77**	0.66	1	0.87**	0.67	1.26	0.59**	0.49	0.85
Integration	−56.73**	−84.91	−28.15	−0.31**	−0.41	−0.21	−0.34**	−0.61	−0.26	0.93**	0.68	1.26	0.57**	0.57	0.94
Local plurality × integration	−7.86	−39.61	23.49	0.11	−0.08	0.32	0.16	−0.1	0.59	−1.3**	−1.95	−0.82	−0.65**	−1.21	−0.48
Local plurality × graduated integration															
Intercept	1069.16**	1028.49	1110.37	−1.46**	−1.5	−1.4	−2.28**	−2.15	−1.97	−1.78**	−1.85	−1.63	−1.98**	−1.8	−1.6
Local plurality	−14.17†	−30.36	1.07	0.36**	0.31	0.43	0.73**	0.58	0.93	0.45**	0.31	0.61	0.52**	0.34	0.73
Integration (Helmert contrast coding)															
Higher vs highest	6.82	−24.51	37.44	0	−0.08	0.09	−0.09	−0.32	0.23	−0.07	−0.28	0.15	−0.2	−0.3	0.29
Mid vs higher–highest	29.86*	2.96	59.1	−0.06	−0.14	0.02	0.19*	0.02	0.48	−0.2†	−0.39	0.01	0	−0.18	0.34
Lower vs mid–higher–highest	22.82	−11.95	63.58	−0.07	−0.18	0.05	0.11	−0.09	0.48	0.02	−0.22	0.27	0.13	−0.09	0.55
Lowest vs lower–mid–higher–highest	50.22*	17.86	83.42	0.01	−0.08	0.1	−0.15†	−0.5	0.03	0.22*	0.03	0.4	−0.04	−0.24	0.29
Local plurality × integration (Helmert contrast coding)															
Plurality × integration higher vs highest	49.27†	−6.66	103.16	−0.44**	−0.6	−0.25	−0.27	−0.86	0.23	−0.24	−0.68	0.18	−0.07	−0.81	0.4
Plurality × integration mid vs higher–highest	−7.12	−52.76	42.02	0.58**	0.43	0.75	0.61*	0.25	1.15	0.2	−0.19	0.61	0.06	−0.5	0.54
Plurality × integration lower vs mid–higher–highest	−67.58*	−127.7	−5.94	0.06	−0.15	0.29	−0.21	−0.67	0.46	−0.03	−0.51	0.48	−0.38	−0.76	0.5
Plurality × integration lowest vs lower–mid–higher–highest	−46.22†	−97.52	2.99	−0.35**	−0.53	−0.18	−0.15	−0.75	0.32	−0.22	−0.58	0.16	0.06	−0.63	0.44

* $p < .05$.
** $p < .001$.
† $p < .10$.

Table 4
Numbers of responses in Experiments 1 and 2 by integration condition and local noun number.

Experiment	Integration level	Local noun	Singular	Plural	Un-inflected	Miscellaneous	Missing	Proportion plural
1	Integrated	Singular	1755	42	0	0	3	0.023
		Plural	1645	155	0	0	0	0.086
	Unintegrated	Singular	1698	99	0	2	1	0.055
		Plural	1576	218	0	2	4	0.122
2	Integrated	Singular	785	2	304	90	19	0.003
		Plural	756	77	257	86	24	0.092
	Unintegrated	Singular	701	74	302	98	25	0.095
		Plural	687	107	266	116	24	0.135

Fig. 5. Experiment 1 proportions of plural responses by integration and local plurality. Error bars are 95% confidence intervals around differences between condition means, calculated from the interaction term in an empirical logit model.

nouns was in fact little different from that of singular local nouns.

One conjecture about the interaction between latencies and the graduated measure of integration has to do with properties of particular item groups. Consider preambles containing the preposition *with*. These preambles could be taken to express a comitative relationship (cf. McNally, 1993), a type of assemblage that is notionally plural and in some languages requires plural agreement. Two things undercut this possibility. First, in previous experimental comparisons, English comitative *with* subjects were no more likely to take plural agreement than subjects having other kinds of prepositional phrases (e.g. *for* phrases; Lorimor, 2007). Second, in the present study, the *with* items had the same rates of agreement as other-preposition items (9% plural for both), similar attraction rates (8% vs 10%, respectively), and similar integration ratings (4.12 and 4.37, respectively).

The only item-type difference of note was for the clausal-complement preambles. Compared to prepositional–phrase preambles, the clausal items had lower average integration ratings (3.28 vs 4.20), smaller average integrated–unintegrated differences (.63 vs 1.08), lower rates of plural agreement (3 vs 9%), and lower rates of attraction (1 vs 9%). This pattern suggests weaker effects overall in the clausal than in the phrasal items, consistent with pre-

vious findings on clausal modulation of agreement (e.g. Bock & Cutting, 1992).

The materials from our study were identical to those used by Solomon and Pearlmutter (2004). However, our task and our primary dependent variable differed from theirs, perhaps making certain features of our procedure responsible for the change in outcomes. Instead of free completion, speakers in Experiment 1 had to choose and produce a predicate from a designated set of adjectives, which might alter the message representation behind the sentence. This could affect the implementation of agreement. Alternatively, the high levels of plausibility associated with the adjective completions at high levels of integration may have blunted the impact of integration on attraction at one end of the scale. We were able to evaluate the role of plausibility using the best-fit plausibility ratings that we collected for every adjective with every preamble. The correlation between these ratings and Solomon and Pearlmutter's ratings (carried out on the preambles alone) was $r(198) = 0.35$, $p < .001$. This tends to discount the speculation about plausibility. Still, there is room for the contention that participants had to produce completions that made less sense to them than unconstrained completions would have. Experiment 2 addressed this concern by eliciting spontaneous, unrestricted sentence completions.

3. Experiment 2

The verb-production latencies in Experiment 1 implied that notional number has a substantial impact on agreement, without a contribution from strong integration to attraction. This is consistent with the existence of a notional component in the effects of semantic integration, and inconsistent with what would be expected from a lexical-interference-driven increase in attraction. Nonetheless, the latency data came from singular-verb production in a task that differed from Solomon and Pearlmutter's (2004). This makes it important to assess whether the implications of Experiment 1 remain the same when speakers produce unconstrained sentence completions. Experiment 2 therefore used Solomon and Pearlmutter's task to find out whether and how agreement changes when speakers produce full sentences with completions they devise themselves.

3.1. Methods

3.1.1. Participants

Undergraduates at the University of Illinois participated in this experiment in exchange for course credit or $7.00 compensation. Out of 50 participants, one was excluded because of recording failure and another because of inability to read the preambles in the allotted time, leaving a total of 48 participants.

3.1.2. Equipment

All equipment was the same as in Experiment 1.

3.1.3. Materials

The 100 experimental preambles were the same as in Experiment 1. The same eight lists were used, except that the implicit practice block at the beginning of the experiment was shortened from 30 to 6 trials in order to reduce the duration of the experiment. This left 228 trials, 128 fillers and 100 experimental trials, of which 53% were plural.

3.1.4. Procedure

As in Experiment 1, participants read the preambles silently from a computer screen, with instructions to repeat the preambles and complete them as full sentences as quickly as possible, using a completion that made sense to them. There were no other restrictions on the predications. Participants were given one explicit practice trial, and were queried about the procedure before beginning the experiment. (See Appendix A for the instructions and explicit practice trials.) The time course of trial events was identical to the catch trials in Experiment 1, except that the *Repeat* prompt was replaced by 500 ms of blank screen.

3.1.5. Scoring

Verb number in the spoken sentences (singular/plural) was coded in the same manner as Experiment 1. Unambiguously singular or plural verbs occurred in 66% of the responses, and of these, 92% were singular.

3.1.6. Design and data analysis

The design was identical to Experiment 1. As in Experiment 1 and Solomon and Pearlmutter (2004), the dependent variable was the proportion of plural verbs produced in the sentence completions, out of all singular and plural verbs.

The statistical analyses used weighted empirical logistic regression from the lme4 package in R carried out with the factors of local plurality and either dichotomous integration or graduated integration as fixed factors. Dichotomous integration was contrast-coded with the values of 0.5 and −0.5, and graduated integration was analyzed with weighted Helmert contrasts.

3.2. Results

Table 4 lists the number of singular- and plural-agreeing verbs produced in each scoring category by dichotomous integration and local plurality, and Fig. 6 shows the proportions of plural verbs in each condition. Overall, more
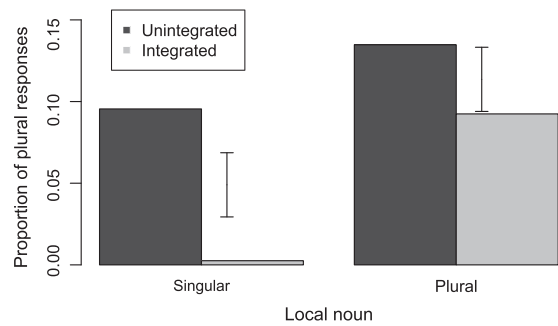

Fig. 6. Experiment 2 proportions of plural verbs after unintegrated and integrated preambles with singular or plural local nouns. Error bars are 95% confidence intervals around differences between means, calculated from the interaction term in an empirical logit model.
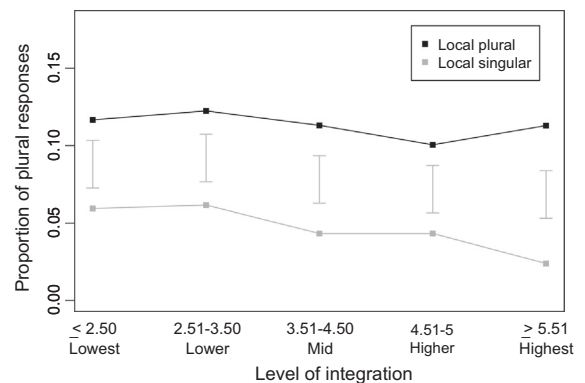

Fig. 7. Experiment 2 plural responses by graduated integration and local plurality. Error bars are 95% confidence intervals around differences between means, calculated from the standard errors of the means.

plural verbs were produced after unintegrated than integrated preambles (proportions of .12–.05) and after plural than singular local nouns (proportions of .11–.05). Notably, attraction (the difference in the proportions of plural verbs after singular and plural local nouns) was larger for integrated than for unintegrated sentence subjects, .09 vs .04.

These effects were captured statistically in weighted empirical logit regression, with models using participants and items as random intercepts (see Table 3 for parameters). The interaction between integration and local noun number was significant, given the larger difference in plural production between singular and plural local nouns for integrated than for unintegrated sentence subjects. There were also significant differences in the production of plural verbs between the integrated and unintegrated sentence-subject conditions and between the singular and plural local-noun conditions.

The second set of analyses used the graduated measure of integration instead of the dichotomous measure as the independent variable. This yielded a different and more detailed perspective on the connection between integration and attraction (see Fig. 7). For attraction, there was a substantial effect across all levels of integration. The effect remained stable across the four bottom levels of integration

(averaging .06), with the only noticeable change coming at the highest integration level (where attraction measured .09). Numerically, more than half of this difference was due to the reduced use of plural verbs after singular local nouns, not to an elevation in the use of plural verbs after plural local nouns.

This trend can be more simply characterized in terms of singular verb use (i.e., the complement of the data in Fig. 7): Increases in integration went along with increases in the production of singular agreement. What this implies for the assessment of attraction is that the higher attraction rates found with stronger integration (on the dichotomous categorization) did not reflect an increase in plural verb use in the attraction-inducing plural-local-noun condition, but a decrease in plural agreement in the baseline singular-local-noun condition.

The statistical evaluation of these trends was once again done with weighted empirical logit models, using graduated semantic integration as a factor. As reported in Table 3, there was a significant effect of local-noun plurality. For semantic integration, the only significant contrast was the one between the weakest level of integration and the four stronger levels, reflecting the higher rate of plural verbs that accompanied low integration. Differences in attraction due to semantic integration, which would be captured in an interaction between integration and local noun number, were not significant in any of the contrasts.

3.3. Discussion

In Experiment 2, just as in Experiment 1, the more integrated a sentence subject was, the more likely speakers were to produce singular agreement. Likewise, the less integrated a sentence subject was, the more likely speakers were to produce plural agreement. Integration made no systematic contribution to attraction. These results run counter to the predictions from an interference account of semantic integration. However, they align well with the notional hypothesis: Perceptually and conceptually coherent states of affairs are more compatible with singular than with plural number. The implication is that the overarching effect of semantic integration on agreement has less to do with attraction and more to do with a speaker's sense of notional number.

In one crucial respect, the results of Experiment 2 were consistent with Solomon and Pearlmutter's data and the predictions of the interference hypothesis. Given an a priori, dichotomous categorization of integration, there was significantly more attraction for integrated than for unintegrated preambles. But measured in terms of the empirical integration ratings on the full integration scale, the effects are hard to square with an interference account. With the graduated measure, what comes to the fore is that the appearance of increased attraction at the highest level of integration was the result of a *decrease* in the production of plural verb agreement in the baseline condition: Sentence subjects with singular heads, singular local nouns, and singular notional number elicited fewer plurals than any other condition. The question that remains is how strong integration could create interference that is resolved in favor of more plural attraction (as interpreted by Solomon & Pearlmutter, 2004, from the dichotomous integration measure) or in favor of more singular agreement with an unchanged attraction rate (as observed here, with the graduated integration measure).

One possibility is that using the actual values of the integration ratings to assign the four preamble versions of each item to integration conditions (as we did here) created disparities with respect to Solomon and Pearlmutter's integration classification. Solomon and Pearlmutter judged preambles a priori as integrated or unintegrated, and then used set-wise average ratings to confirm their classifications. As a result, 37 preambles that occurred in the integrated conditions in Solomon and Pearlmutter's experiments were classified as unintegrated in the present experiments, and vice versa in the unintegrated condition (for a total of 74 switched item-versions out of the 400). These reversals were rare enough to have little impact on the results of Experiment 2. (See Appendix C for plots of the item data using both Solomon and Pearlmutter's a priori classification and the empirical classification derived from their ratings.)

A different and more likely explanation is that the graduated measure of integration served to distribute the agreement outcomes more evenly across the range of integration values. That is, along the integration scale, the differences between tightly and loosely integrated preambles could vary substantially. Suppose that the empirical ratings of sentence preambles that had been judged in advance to be loosely integrated were clustered toward the lower end of the integration scale, with their tightly integrated counterparts spanning a wider range (as was in fact the case). Because of measured differences in notional number, the rates of plural agreement among the a priori integrated items could then vary more than the rates in the a priori unintegrated items.

Put differently, if loose integration promotes plural agreement regardless of local noun number, there will be more opportunities for plural agreement to come into play for items that fall lower along the graduated integration scale. This includes items that were classified a priori as (relatively) integrated despite low empirical integration ratings. By separating items according to their actual rated values instead of the dichotomous a priori division, a clearer picture emerges of the relationships among agreement, attraction, and integration. The clearer picture suggests separable effects, one due to notional number and the other due to grammatical number.

Another factor to be reckoned with, noted in the introduction, has to do with rates of defective, hence unscorable, responses. With defective responses being more common in the unintegrated condition, where notional plurality makes attraction harder to observe, reduced rates of scorable agreement outcomes shrink the power to detect attraction even more. Defective responses were more common in Experiment 2 (with free completion) than in Experiment 1, 10 and 1% respectively, but both rates were much lower than in Solomon and Pearlmutter's experiments (29%). Accordingly, we had more opportunities to observe plural agreement in both the integrated and unintegrated conditions and to better approximate its rate differentially across integration levels.

Ultimately, what the interference account lacks is an explanation for the facilitation (Experiment 1) and increased incidence of singular agreement (Experiment 2) when integration was strong, and for the increase in plural agreement when integration was weak. In contrast, the Marking and Morphing theory (Eberhard et al., 2005) predicts these differences in addition to the accompanying rates of attraction. In the General Discussion, we lay out what this alternative account has to say about how the notional properties of semantic integration change the implementation of agreement.

4. General discussion

The findings from these experiments point to notional number as a major force in how semantic integration affects the implementation of subject–verb agreement. The more coherent, more internally cohesive, or more integrated a complex referent in a mental model is, the more likely the referent is to be construed as singular, and the more rapidly and frequently speakers will be to generate singular agreement. Conversely, the less coherent, less cohesive, or less integrated a complex referent is, the slower speakers become at producing singular agreement, and the more likely they are to produce plural agreement.

The connection between semantic integration and the speed of producing verbs emerged in Experiment 1. Verb production latency was measured from the offset of a briefly presented sentence subject that was completed as a full sentence using a predicate adjective. Less-integrated subject noun phrases led to slower production of singular verbs regardless of whether integration was treated as a dichotomous or graduated property. The presence of grammatically plural, attraction-inducing local nouns further slowed the production of singulars, and especially for *less* integrated subjects.

Experiment 2 used an unconstrained production task (Bock & Miller, 1991; Gillespie & Pearlmutter, 2011; Solomon & Pearlmutter, 2004) to more closely examine the likelihood of singular and plural verb production. When speakers were allowed to formulate their own sentence continuations freely, higher levels of integration were again associated with usage of singular verbs, paralleling the increased speed of singular-verb production in Experiment 1.

These patterns show that semantic integration affects the overall rates of producing singular and plural agreement in a manner consistent with the effects of notional number. However, in existing research on semantic integration, the focus is on attraction. As an aberration from typical agreement, attraction changes the rate of plural relative to singular agreement under certain conditions. The question in the present work was how semantic integration contributes as a notional factor to singular and plural number agreement.

Viewed with an eye toward the occurrence of attraction only, the puzzling result of Experiment 2 was this: When integration was treated as a binary property, without regard to graded differences in strength, the relative proportion of plural verbs rose after plural local nouns, and more so for tightly integrated than for loosely integrated subjects. This is the same result obtained by Solomon and Pearlmutter (2004). They interpreted the finding as a reflection of variations in formulation and retrieval set in motion by semantic integration. This interpretation, despite its information-processing plausibility, is hard to square with the agreement patterns that emerged in our data when integration was treated as a graded property rather than a discrete one. In the next section we sketch an alternative account.

4.1. Marking, morphing, and semantic integration

The Marking and Morphing model (Eberhard et al., 2005) can parsimoniously account for the effect of semantic integration on agreement without calling on the mechanisms of attraction. The model handles semantic integration in the same way it handles other notional factors. The broad claim is that the number features of entire subject noun phrases control verb number agreement, not the features of isolated words or phrases alone. Subject number comes about from the combination and reconciliation of notional and lexical influences. The reconciled features of the subject can determine verb number without changes in the subject's structure (as phrasal clashes might create) or agreement with an isolated word or phrase (as lexical clashes might create).

In the model, there is a marking of syntactically relevant number features on the subject noun phrase that derives in part from notional number. The influence of lexical number specifications comes about during the reconciliation of lexical and notional number sources, with lexical sources carrying more weight and thereby reducing the impact of notional number (Berg, 1998; Foote & Bock, 2012; Lorimor, Bock, Zalkind, Sheyman, & Beard, 2008). To a large extent, the outcome of number reconciliation is rooted in a property that may be universal in language and cognition: Singularity is an unmarked default. Accordingly, when the referent of a subject is notionally singular, and the subject phrase is singular, verb agreement can unfold without specific input from the unmarked subject.

If a plural head noun (like *scissors*) enters the formulation process, the number reconciliation that is needed for agreement to proceed will almost always yield a marked plural subject: The plural lexical specification becomes the subject's number. Such takeovers are a general property of language processing that Stemberger identified as "something replacing nothing" (Stemberger, 1991). In the case of grammatical number, "something" is plurality and "nothing" is the singular default (Berent, Pinker, Tzelgov, Bibi, & Goldfarb, 2005).

In attraction, this same "something replaces nothing" principle of reconciliation is at work, but with a much lower probability of success. Local nouns within a subject phrase are not excluded from reconciliation in principle, but in practice their impact is minimal: By definition, local nouns reside deeper than head nouns in a sentence subject's structure, so their typical effect is barely discernible. However, when a plural feature from a local noun intrudes on the reconciliation, it can create attraction via the same reconciliation process that is ordinarily responsible for a plural head noun creating a plural subject noun phrase.

For present purposes, there are two predictions from Marking and Morphing for overt agreement production, as was observed in Experiment 2. The crucial prediction is that if there is no grammatically plural noun to affect a subject's marking, the only consistent force that can lead to plural agreement is notional number. This means that notional-number contributions from variations in semantic integration should surface when grammatically singular head and local nouns constitute the sentence subject, and that under these conditions, decreasing integration should be associated with increasing plurality. This prediction was confirmed: The production of plural verbs after singular subjects with singular local nouns went up as integration went down. The second prediction is with regard to attraction and the specific effect of grammatically plural local nouns. Marking and Morphing predicts an elevation in the rate of plural agreement that should not vary systematically across gradations in integration. This was also confirmed. So, what differed between more and less integrated subjects was the rate of notionally driven plural agreement, not the rate of attraction.

4.2. Interactions between attraction and integration

In order to confidently measure attraction relative to notionally driven plurality, a valid assessment of notional number is essential. Here, the ratings-based measure of integration constituted one assessment. Using these ratings, we obtained evidence that the integration judgments captured something about notional number: There was a consistent decline in the incidence of plural agreement as integration increased, when no triggers for grammatical plurality were present. When a trigger for grammatical plurality *was* present, more plural agreement occurred, but the increase was fairly stable across the range of rated integration. Thus, attraction was separable from the effect of integration in our measurements, emerging when the rates of plural agreement were aligned with respect to integration ratings, and the numbers of observations under conditions relevant to both notional agreement and attraction (i.e. with and without local plurality) was large enough for the measures of agreement to be stable.

It is impossible to be fully certain, but the discrepancies between our results and those of Solomon and Pearlmutter are likely to be the consequence of differences in success at meeting these conditions for separating the effects of notional number and attraction. We benefited from being able to call simultaneously on all of the sentence preambles that Solomon and Pearlmutter composed, rather than confronting the need for separate sets of materials with variable properties (including variable ranges of low and high integration). We also benefited from a higher incidence of scorable agreement responses. Together, these provided distributions of values that allowed systematic variations in notional agreement to be gauged.

4.3. Semantic integration, notional number, and competition for control of agreement

Our results can thus be taken to mean that semantic integration plays a more all-encompassing role in number

agreement than the interference account implies, a role that affects singular and plural agreement regardless of attraction. Beyond any impact that semantic integration could have on the time-course of lexical retrieval, well-integrated referents are notional units and poorly integrated referents are notional multiples. When present in messages at the outset of sentence formulation, units predispose grammatical singularity and multiples predispose grammatical plurality. In the Marking and Morphing model, these factors in turn can disclose or camouflage attraction.

A pivotal difference between our account and the one proposed by Solomon and Pearlmutter (2004), (further elaborated by Gillespie & Pearlmutter, 2011) is in the role of simultaneity in processing. We suspect that parallel processing and timing factors in general play a large role in producing agreement, and Marking and Morphing does not exclude the possibility that the scope of planning and variations in timing matter. The question remains whether the model can predict timing effects without changing other properties (for a dissenting view see DiBattista & Pearlmutter, 2011; Gillespie & Pearlmutter, 2013, 2011) and whether lexical interference or competition in some form contributes to attraction. For the moment, the point is that semantic integration appears to affect the formulation of agreement more by way of notional number than by way of attraction.

It is certainly conceivable that effects of lexicalization and notional number on response latencies are both lurking in our data, particularly if one process takes much longer than the other (cf. McClelland, 1979). That is, if conceptual encoding of notional number is more protracted than resolving competition between lexical items, lexical effects could be hidden. The simple lexical competition hypothesis that we tested in Experiment 1 assumed a penalty from competition that would slow singular-verb responses after plural local nouns more so in integrated than in unintegrated conditions. The results failed to show this interaction. In fact, they showed the opposite, with longer latencies after plural local nouns only at the *lowest* levels of integration. The latency and response data together pointed in the same direction: Integration makes agreement easier. Though this does not rule out lexical competition, it does suggest that different measures of agreement are keyed to the same production mechanisms.

4.4. Retrieval processes in language production and language comprehension

In emphasizing the role of notional number in what semantic integration does, we do not intend to discount the contribution of memory mechanisms to agreement. Retrieving, maintaining, and reconciling agreement features are essential for coping with long-distance dependencies, in ways that are beginning to be illuminated (Badecker & Kuminiak, 2007; Lewis & Badecker, 2010; Lewis & Vasishth, 2005; Van Dyke & McElree, 2006; Wagers et al., 2009). However, much of this research is on language comprehension (or on the production of lexically controlled gender agreement). The impact

of lexically specified agreement features comes relatively early during the understanding of sentences, but relatively late during formulation for production. The consequence is significant tradeoffs between maintaining and retrieving information that vary by modality, as well as variations in the kinds of information that are retrieved and maintained.

Viewed more broadly, it is obvious that interactions among semantic integration, notional number, and agreement matter in both comprehension and production. In research on language comprehension, there is considerable evidence that variations in notional number affect the understanding of pronouns (Albrecht & Clifton, 1998; Garnham, 2001; Garnham, Oakhill, Ehrlich, & Carreiras, 1995; Gernsbacher, 1991; Koh & Clifton, 2002; Moxey et al., 2004; Oakhill, Garnham, Gernsbacher, & Cain, 1992) in ways that parallel the production of pronouns (Bock et al., 2004, 2006; Moxey et al., 2004). Differences in the mental representation of complex referents, in the integration or individuation of their parts, entail differences in notional plurality that can change parsing decisions (Ferreira & McClure, 1997; Patson & Ferreira, 2009; Patson & Warren, 2011). The products of comprehension likewise exhibit effects of integration, as in classic work on mental models showing that spatially associated, semantically integrated referents are accessed faster in memory than spatially dissociated, semantically unintegrated referents (Glenberg, Meyer, & Lindem, 1987). It remains to be seen whether integration, as a fundamental conceptual property, enters into the syntax of understanding and speaking in the same ways.

4.5. Conclusion

The present results affirm Solomon and Pearlmutter's (2004) view that semantic integration is important to agreement. However, our findings suggest that semantic integration affects agreement in ways that require a different explanation. Semantic integration reflects basic features of the mental representations of events and objects in the world, and how speakers enumerate objects within those representations. What emerges from this enumeration, notional number, lays the groundwork for grammatical number and its role in the fortunes and misfortunes of grammatical agreement.

Appendix A. Instructions for norming, Experiment 1 task, and Experiment 2 task

A.1. Norming instructions

In this experiment, you will see a sentence on the screen. Each sentence will contain one of four adjectives: "good", "bad", "ready", and "true." Please judge the plausibility of this sentence on a scale from 1 to 7 (implausible to plausible) by pressing the number keys.

For example:

The casserole in the oven was ready	Implausible 1 2 3 4 5 6 7 Plausible
This is very plausible, so you would press 7.	
The casserole in the thimble was ready	Implausible 1 2 3 4 5 6 7 Plausible
This is very implausible, so you would press 1.	
After you press a number key, you will get the next sentence.	
Do you have any questions?	

A.2. Experiment 1 task instructions

In this experiment, you will be seeing the beginnings of sentences and then completing them. For the completions, you can choose any one of four adjectives: good, bad, ready, true. The way a trial works is like this: First you will see "X" on the screen for about a second. This is a warning signal, telling you a phrase is about to appear. Then you will see the phrase, which you should treat as the start of a sentence. It will be followed immediately by an exclamation point. This is your signal to speak. For example, if you saw "The fire engine" and then saw "!", You could say "was ready." Now let's try an example. Press the space bar to begin: X/Cookie monster/!

Sometimes something a little different will happen. Instead of an exclamation point, the word "Repeat" will appear. When this happens, you should first REPEAT the phrase, and THEN complete it. For example, if you had just seen "The fire engine", and then you saw the word "Repeat", you would repeat the phrase aloud along with your completion. Like, "The fire engine was good." Your job here is to repeat the phrase, exactly as you saw it, and make it into a complete sentence using one of the four adjectives (good, bad, ready, true). Press the space bar for an example: X/The rollercoasters/Repeat.

Let's review: What are the 4 adjectives you can use? What do you do when you see "!" ? What do you do when you see "Repeat" ? Any questions? When you are ready, press space bar to begin.

A.3. Experiment 2 task instructions

In this experiment, you will be seeing the beginnings of sentences and then completing them. The way a trial works is like this: First you will see "X" on the screen for about a

second. This is a warning signal, telling you a phrase is about to appear. Then you will see the phrase, which you should treat as the start of a sentence. Please repeat the beginning part and then make it into a complete sentence.

For example, if you saw "The fire engine" you could say "The fire engine was ready." If you saw "The rollercoasters at Six Flags" you could say "The rollercoasters at Six Flags were fun."

Now let's try an example. Press the space bar to begin: X/Cookie monster.

Any questions?

Appendix B.

Ratings for preamble integration and plausibility from Solomon and Pearlmutter (2004); (preamble integration and plausibility ratings shown for preamble versions representing the two integration categories and two local noun numbers, in the following order: integrated-singular, unintegrated-singular, integrated-plural, unintegrated-plural) and for adjective-completion plausibility ratings from present experiments (for integrated-singular and unintegrated-singular versions)

Preamble (integrated/unintegrated)	Integration	Plausibility	Mean plausibility	
			All adjectives	Best-fitting adjective
The drawing (of/with) the flower(s)	5.0/2.9/ 3.4/2.6	6.6/6.1/ 6.1/4.6	5.6/5.4	6.8/6.4
The picture (of/with) the gem(s)	4.3/3.3/ 4.8/3.3	5.5/3.8/ 5.5/3.9	5.1/4.8	6.2/6.8
The sculpture (of/with) the key(s)	4.0/2.6/ 3.8/2.0	3.5/2.5/ 3.1/2.8	5.4/4.7	6.8/6.2
The sketch (of/with) the bookcase(s)	4.2/2.3/ 4.1/2.7	5.2/3.0/ 4.3/3.2	5.4/4.9	7.0/6.6
The painting (of/with) the costume(s)	3.3/2.6/ 3.0/2.0	4.9/2.6/ 3.9/3.8	5.5/4.7	6.8/6.8
The statue (of/with) the bird(s)	3.3/3.4/ 4.6/2.7	5.3/4.9/ 4.8/4.5	5.7/4.2	6.6/5.6
The tape (of/with) the record(s)	5.3/3.9/ 5.6/4.6	4.1/3.1/ 4.5/4.8	5.4/5.7	6.2/6.8
The photo (of/with) the document(s)	4.3/3.6/ 4.9/3.4	4.6/4.7/ 3.7/4.6	6.3/5.6	6.6/6.2
The xerox (of/with) the memo(s)	4.6/4.5/ 4.6/4.3	5.9/4.4/ 5.8/4.5	5.2/5.5	6.8/6.2
The video (of/with) the puppet(s)	3.2/3.3/ 3.6/3.1	4.0/3.7/ 5.0/4.5	5.0/6.4	6.4/7.0
The illustration (of/with) the map(s)	5.3/5.1/ 4.4/4.4	4.9/5.1/ 5.1/5.0	5.7/5.5	6.8/6.4
The photocopy (of/with) the article(s)	4.4/3.6/ 3.9/3.6	6.5/4.7/ 5.9/4.0	6.5/5.3	7.0/6.4
The reproduction (of/with) the antique(s)	4.2/3.5/ 3.6/2.8	5.4/3.1/ 5.3/4.2	4.8/4.1	6.4/4.4
The fax (of/with) the blueprint(s)	3.5/3.5/ 5.6/4.3	4.8/4.6/ 4.9/4.9	5.8/5.2	7.0/5.6
The telecast (of/with) the movie(s)	5.2/4.3/ 3.7/4.0	5.0/3.8/ 4.3/3.8	5.8/5.8	6.2/6.8
The snapshot (of/with) the letter(s)	3.2/2.5/ 4.1/2.9	3.9/4.5/ 2.4/4.5	5.5/5.3	6.8/6.6
The description (of/with) the CD(s)	3.7/3.9/ 2.8/3.4	5.9/4.5/ 5.5/4.7	5.8/5.5	6.8/7.0
The slide (of/with) the magazine(s)	3.2/2.8/ 3.8/2.2	3.4/2.9/ 3.1/3.1	4.6/4.8	5.4/5.6
The broadcast (of/with) the show(s)	5.8/5.3/ 5.5/5.7	6.2/3.3/ 5.9/4.4	6.1/4.9	7.0/6.2
The draft (of/with) the report(s)	5.2/5.0/	6.4/5.2/	6.7/5.6	7.0/6.8

(continued on next page)

(*continued*)

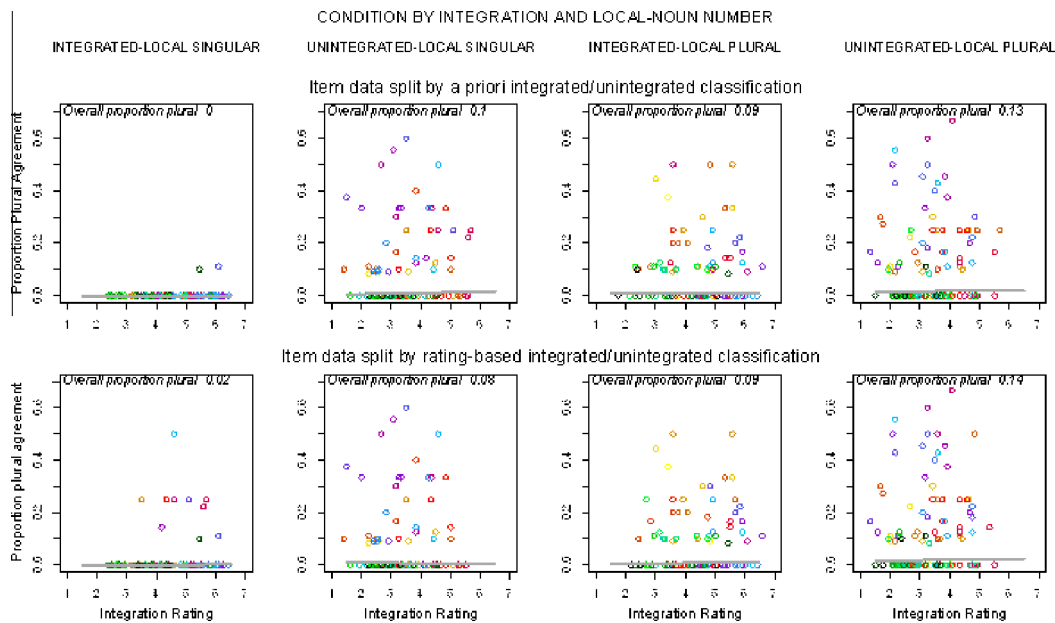| Preamble (integrated/unintegrated)	Integration	Plausibility	Mean plausibility	
			All adjectives	Best-fitting adjective
	4.8/4.9	6.2/4.9		
The sonogram (of/with) the infant(s)	5.3/5.3/ 5.3/4.7	5.8/5.0/ 5.6/4.7	6.2/4.7	7.0/6.0
The polaroid (of/with) the stamp(s)	2.9/2.3/ 2.4/1.7	3.2/2.2/ 2.8/2.5	5.3/4.0	6.6/5.8
The postcard (of/with) the shoe(s)	2.4/1.4/ 2.8/1.8	2.6/2.7/ 3.5/2.4	4.6/3.7	6.0/5.0
The portrait (of/with) the crown(s)	3.5/3.2/ 4.8/3.6	4.7/4.5/ 3.8/4.2	6.1/5.3	7.0/6.6
The assistant (for/with) the inspector(s)	4.3/3.8/ 3.6/2.8	4.9/4.5/ 6.1/4.5	5.5/4.4	7.0/6.2
The chauffeur (for/with) the actor(s)	4.2/3.2/ 2.8/2.8	5.6/4.9/ 6.4/5.0	5.1/4.9	7.0/6.6
The apprentice (for/with) the tailor(s)	5.1/4.7/ 4.5/4.1	5.2/5.5/ 5.3/4.3	5.4/4.8	6.8/7.0
The supporter (of/with) the evangelist(s)	3.8/3.0/ 3.6/3.2	5.1/4.2/ 5.0/5.1	4.9/5.1	6.2/5.8
The translator (of/with) the ambassador(s)	5.2/4.8/ 4.1/3.8	5.8/5.7/ 6.4/6.2	5.4/5.3	6.8/6.8
The secretary (of/with) the supervisor(s)	5.4/5.0/ 3.9/3.4	5.7/6.1/ 5.4/5.2	5.2/5.1	7.0/6.0
The accountant (for/with) the millionaire(s)	4.3/3.5/ 4.4/4.3	6.7/6.0/ 6.2/5.7	5.2/4.0	7.0/5.8
The nurse (for/with) the surgeon(s)	4.8/4.9/ 5.0/4.6	5.9/6.2/ 4.9/6.3	5.5/5.7	6.8/6.8
The consultant (of/with) the producer(s)	4.3/4.3/ 4.3/3.8	4.4/5.4/ 4.7/5.2	4.8/5.1	5.8/6.6
The advisor (of/with) the attorney(s)	5.3/3.3/ 4.7/4.3	4.5/5.1/ 3.3/5.6	5.7/5.3	6.8/6.4
The servant (for/with) the diplomat(s)	3.9/3.3/ 3.7/3.8	5.2/5.4/ 5.2/5.2	5.9/4.5	6.4/5.8
The manager (of/with) the band(s)	4.7/4.5/ 5.1/4.3	5.1/5.9/ 6.2/5.1	5.5/5.1	7.0/6.8
The agent (for/with) the artist(s)	4.0/3.3/ 3.5/3.7	6.3/5.0/ 5.1/4.9	5.2/4.8	6.8/6.2
The trainer (for/with) the athlete(s)	5.5/5.5/ 5.3/5.5	6.6/6.2/ 6.7/6.9	5.3/6.2	6.8/7.0
The doctor (of/with) the patient(s)	5.8/5.5/ 5.9/5.5	6.9/6.8/ 6.8/7.0	5.4/4.6	6.8/5.8
The tutor (for/with) the student(s)	5.5/5.7/ 5.4/4.8	5.7/6.9/ 6.8/5.8	5.4/4.7	6.8/6.2
The coach (of/with) the gymnast(s)	5.3/5.6/ 5.8/4.3	6.9/6.5/ 6.5/5.9	5.4/4.9	7.0/6.6
The photographer (of/with) the supermodel(s)	4.4/5.6/ 4.9/4.9	6.9/6.3/ 7.0/6.4	5.2/5.7	6.6/7.0
The promoter (for/with) the DJ(s)	4.5/4.6/ 4.6/4.1	5.6/5.6/ 6.6/5.5	5.9/5.2	6.8/7.0
The groundskeeper (of/with) the landowner(s)	4.4/4.4/ 3.6/4.5	5.4/5.6/ 5.2/5.5	4.6/5.0	6.8/6.2
The book with (the torn page(s)/the red pen(s))	6.0/3.2/ 6.1/2.9	5.9/4.3/ 6.5/3.8	4.2/4.6	6.2/6.6
The shirt with (the crazy pattern(s)/the dirty towel(s))	5.2/2.7/ 4.8/3.3	5.5/3.6/ 6.6/3.8	4.4/3.6	6.0/5.4

(continued)

Preamble (integrated/unintegrated)	Integration	Plausibility	Mean plausibility	
			All adjectives	Best-fitting adjective
The ring with (the fake diamond(s)/the gold bracelet(s))	6.1/3.8/ 5.9/3.9	4.3/6.0/ 6.5/5.4	4.5/4.9	5.8/5.6
The apple with (the brown spot(s)/the fresh peach(s))	4.0/3.1/ 4.4/3.8	5.5/5.4/ 6.5/4.9	3.1/4.6	6.6/7.0
The tie with (the hideous stripe(s)/the cotton blazer(s))	4.1/4.2/ 5.4/4.7	5.0/5.8/ 6.6/5.3	3.4/4.2	4.4/6.8
The watch with (the missing hand(s)/the black wallet(s))	6.0/3.3/ 5.1/3.2	4.2/3.9/ 5.6/3.9	3.4/3.8	5.8/5.6
The jacket with (the faulty zipper(s)/the wet umbrella(s))	6.3/3.3/ 5.9/1.6	5.5/5.0/ 5.5/3.6	3.5/2.8	5.6/4.2
The razor with (the rusty blade(s)/the empty can(s))	6.2/2.0/ 6.6/2.1	5.7/3.6/ 5.3/3.2	3.9/2.7	6.8/4.6
The key with (the jagged edge(s)/the shiny coin(s))	5.0/2.6/ 4.8/2.3	6.6/3.4/ 6.3/2.6	3.8/3.7	6.4/5.0
The bed with (the creaky spring(s)/the tall bookcase(s))	5.8/2.9/ 5.3/3.3	5.9/4.5/ 5.7/4.0	4.0/3.9	5.8/5.2
The phone with (the missing button(s)/the broken toaster(s))	5.3/1.5/ 5.3/1.3	4.5/1.6/ 3.8/1.9	3.5/2.4	6.4/3.8
The pillow with (the nasty stain(s)/the flannel sheet(s))	3.5/5.1/ 4.8/4.8	5.1/5.6/ 5.0/6.2	2.7/4.2	4.8/5.8
The lamp with (the fluorescent bulb(s)/the antique portrait(s))	6.1/2.4/ 5.7/2.2	6.3/4.2/ 6.6/4.4	4.4/3.9	6.0/6.2
The magazine with (the colorful ad(s)/the telephone book(s))	6.1/3.5/ 5.8/3.3	6.4/2.9/ 6.6/3.8	5.5/3.3	6.8/4.2
The sweater with (the tiny hole(s)/the linen suit(s))	4.5/3.8/ 4.7/3.1	5.6/3.2/ 4.2/3.6	4.0/3.6	5.6/5.6
The receipt with (the blurry price(s)/the sealed package(s))	5.8/3.7/ 5.7/3.5	5.5/4.3/ 5.1/4.6	4.9/3.9	6.6/5.0
The tree with (the dead branch(es)/the small shrub(s))	6.4/4.3/ 6.4/4.8	6.4/3.7/ 6.1/5.1	3.2/3.9	3.8/4.6
The pizza with (the yummy topping(s)/the tasty beverage(s))	6/2.6/6.4/ 3.6	6.6/5.6/ 6.5/5.1	4.6/4.2	7.0/6.4
The milk with (the extra vitamin(s)/the blueberry muffin(s))	5.6/2.8/ 4.9/3.1	5.8/5.9/ 5.4/4.4	3.6/5.5	6.0/6.6
The guitar with (the loose string(s)/the loud drum(s))	6.0/3.8/ 6.0/3.8	6.2/4.8/ 5.6/3.3	4.1/4.9	6.6/6.8
The blanket with (the soft fringe(s)/the clean skirt(s))	5.4/2.5/ 5.8/3.9	6.2/3.4/ 6.4/3.1	4.4/3.5	6.2/5.0
The glass with (the lengthy crack(s)/the crystal bowl(s))	4.6/4.6/ 4.9/4.8	5.6/4.9/ 4.3/4.0	3.6/3.6	6.0/5.4
The bike with (the bent spoke(s)/the surfboard(s))	5.9/2.2/ 6.3/2.2	6.2/3.2/ 5.0/2.5	4.1/3.3	4.6/4.6
The chair with (the wobbly leg(s)/the old table(s))	5.0/4.4/ 5.9/3.6	6.6/5.7/ 5.4/4.4	3.6/5.0	4.8/6.8
The report that (described the traffic accident(s)/Megan described the accident(s))	4.8/4.1/ 4.8/4.8	6.3/4.8/ 6.0/4.6	6.1/5.9	6.6/6.8
The confirmation that (excited the employee(s)/Phil fired the employee(s))	3.6/2.2/ 3.4/2.9	5.5/4.6/ 5.9/5.6	4.6/4.9	5.8/5.6
The opinion that (insulted the executive(s)/Claire should help the executive(s))	3.3/1.9/ 3.3/2.5	5.6/5.3/ 5.3/5.9	4.6/3.9	6.2/5.8
The fear that (consumed the prisoner(s)/Dan freed the prisoner(s))	4.3/3.1/ 4.8/3.9	5.3/4.5/ 5.3/5.3	4.1/4.0	6.4/6.2
The idea that (struck the crazy scientist(s)/the school sued the scientist(s))	4.4/2.9/ 4.4/3.5	5.4/5.9/ 5.0/6.0	5.8/4.2	6.8/6.2

(continued)

Preamble (integrated/unintegrated)	Integration	Plausibility	Mean plausibility	
			All adjectives	Best-fitting adjective
The proof that (changed the original ruling(s)/the lawyer approved the ruling(s))	4.1/3.6/ 4.2/3.3	5.6/5.0/ 5.0/5.0	5.7/4.3	6.6/5.0
The rumor that (plagued the gorgeous dancer(s)/Ted dated the dancer(s))	2.5/2.6/ 3.0/2.6	5.5/5.9/ 4.6/5.9	4.5/4.4	6.2/7.0
The discovery that (revealed the precious gem(s)/Kim bought the precious gem(s))	3.3/3.6/ 3.7/2.9	5.6/5.4/ 5.9/4.1	4.8/3.9	5.8/5.8
The request that (pleased the energetic chef(s)/Troy please the talented chef(s))	2.9/2.8/ 3.3/2.2	5.4/4.8/ 5.8/5.8	4.6/3.4	5.4/4.4
The speculation that (disgusted the elite committee(s)/Julie disgusted the committee(s))	3.7/3.1/ 3.1/1.9	5.0/4.8/ 4.4/5.6	4.7/4.6	5.8/6.6
The announcement that (cancelled the scheduled meeting(s)/the senator moved the meeting(s))	3.4/3.4/ 4.4/3.8	6.0/6.4/ 6.0/6.0	5.0/5.9	6.6/6.8
The dream that (haunted the wealthy stockbroker(s)/Claire would assist the stockbroker(s))	2.8/2.4/ 2.1/2.7	5.1/5.3/ 4.5/4.9	4.2/4.7	5.8/6.2
The point that (calmed the anxious witness(es)/Kevin calmed the witness(es))	2.4/2.6/ 2.4/2.7	4.6/5.3/ 4.6/4.0	4.3/3.2	5.8/4.0
The threat that (mentioned the deadly bomb(s)/the thief planted the bomb(s))	4.9/4.4/ 5.1/4.9	4.8/4.1/ 5.6/5.0	4.7/3.7	6.2/5.6
The observation that (confused the brilliant scientist(s)/Sarah confused the scientist(s))	4.4/3.8/ 4.8/4.0	4.6/4.4/ 5.0/4.4	4.2/3.6	4.4/5.0
The verification that (consoled the busy worker(s)/David consoled the worker(s))	2.6/2.7/ 2.3/2.7	5.5/5.6/ 5.3/5.3	5.0/4.3	6.2/5.8
The misconception that (puzzled the boy(s)/Mark smacked the boy(s))	2.8/1.6/ 2.8/2.3	5.4/4.9/ 5.8/4.8	4.0/4.5	6.0/6.8
The suggestion that (disappointed the cranky governor(s)/the aide disappointed the governor(s))	2.9/2.3/ 2.8/2.1	4.1/4.5/ 3.6/4.3	4.4/4.4	6.0/5.6
The possibility that (interested the recent applicant(s)/Richard might interest the applicant(s))	3.9/2.4/ 2.9/2.3	5.5/4.0/ 5.6/4.9	3.8/4.6	5.2/6.3
The feeling that (troubled the little girl(s)/Curt troubled the small girl(s))	3.9/2.6/ 3.3/2.3	5.9/4.4/ 5.1/4.1	3.7/4.5	5.6/6.2
The finding that (showed the important result(s)/Beth lied about the result(s))	5.8/4.4/ 5.0/3.7	5.3/3.5/ 4.6/3.6	5.2/5.0	6.2/6.2
The myth that (amused the small town(s)/Ann amused the town(s))	2.4/2.3/ 2.3/2.1	5.8/4.5/ 4.5/2.6	4.7/3.8	7.0/5.8
The argument that (bothered the enraged striker(s)/the boss bothered the striker(s))	3.3/2.7/ 4.6/3.2	5.0/4.1/ 4.6/4.1	4.9/4.0	5.6/6.0
The remark that (insulted the good friend(s)/Seth insulted the friend(s))	2.5/2.5/ 2.9/2.7	5.5/4.3/ 5.9/4.5	5.0/3.8	6.6/5.8
The recommendation that (thrilled the avid reader(s)/Ms. Drew thrill the reader(s))	2.8/2.7/ 3.4/2.4	4.8/4.1/ 5.5/3.1	4.7/3.7	5.6/6.4
The information that (convinced the rich investor(s)/Mae convinced the investor(s))	3.8/3.0/ 4.3/3.7	5.4/4.9/ 6.0/4.8	5.6/4.0	6.4/4.2
The evidence that (convicted the criminal(s)/Dan hated the criminal(s))	5.4/4.6/ 5.4/3.6	6.0/4.4/ 5.9/5.3	5.8/4.0	7.0/5.8
The hope that (reassured the actor(s)/Bob would meet the actor(s))	3.6/2.9/ 2.4/1.5	4.4/3.8/ 5.5/5.3	4.0/3.5	5.8/4.0
The insinuation that (alluded to the affair(s)/Jane had the steamy affair(s))	4.3/3.8/ 3.6/3.2	5.0/4.6/ 5.0/5.1	4.5/5.0	6.2/7.0
The judgment that (annoyed the defending attorney(s)/Brianna annoyed the attorney(s))	4.4/3.3/ 5.1/3.6	5.3/4.4/ 5.5/3.9	4.7/3.5	6.8/4.2
The news that (encouraged the baseball team(s)/the coach encouraged the team(s))	3.4/2.7/ 2.4/2.5	6.0/5.5/ 5.6/4.5	4.7/4.6	6.0/6.2
The notice that (hurt the caring parent(s)/the son hurt the parent(s))	3.4/2.3/ 1.8/2.2	4.3/4.6/ 5.8/4.1	3.5/3.6	4.6/4.6

CONDITION BY INTEGRATION AND LOCAL-NOUN NUMBER

Appendix C.

Graphs showing distributions of plural-response proportions by dichotomous integration and local noun number using (top row) a priori classification of integration from Solomon and Pearlmutter, 2004, and (bottom row) ratings-based empirical classification of integration (using rating data from Solomon & Pearlmutter). Regression lines were calculated from back-transformed arcsine values of the plural proportions. Points represent condition-specific versions of single items ($n = 100$ per condition).

Point colors denote specific item-versions within conditions, and individual items are shown in the same color across conditions..

References

Albrecht, J., & Clifton, C. Jr., (1998). Accessing singular antecedents in conjoined phrases. *Memory and Cognition, 26*, 599–610.

Badecker, W., & Kuminiak, F. (2007). Morphology, agreement, and working memory retrieval in sentence production: Evidence from gender and case in Slovak. *Journal of Memory and Language, 56*, 65–85.

Barker, C. (1992). Group terms in English: Representing groups as atoms. *Journal of Semantics, 9*, 69–93.

Barr, D. J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language, 59*, 457–474.

Bates, D. M. (2005). Fitting linear mixed models in R: Using the lme4 package. *R News: The Newsletter of the R Project, 5*, 27–30.

Berent, I., Pinker, S., Tzelgov, J., Bibi, U., & Goldfarb, L. (2005). Computation of semantic number from morphological information. *Journal of Memory and Language, 53*, 342–358.

Berg, T. (1998). The resolution of number conflicts in English and German agreement patterns. *Linguistics, 36*, 41–70.

Bloom, P. (2001). Precis of how children learn the meaning of words. *Behavioral and Brain Sciences, 24*, 1095–1103.

Bock, J. K. (2011). How much correction of syntactic errors are there, anyway? *Language and Linguistics Compass, 5*, 322–335.

Bock, J. K., Butterfield, S., Cutler, A., Cutting, J. C., Eberhard, K. M., & Humphreys, K. R. (2006). Number agreement in British and American English: Disagreeing to agree collectively. *Language, 82*, 64–113.

Bock, J. K., Carreiras, M., & Meseguer, E. (2012). Number meaning and number grammar in English and Spanish. *Journal of Memory and Language, 66*, 17–37.

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., & 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.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 25*, 257–271.

DiBattista, A., & Pearlmutter, N. J. (2011). Effects of semantic integration on phrase and word ordering errors in production. In L. Carlson, C. Hoelscher, & T. F. Shipley (Eds.), *Proceedings of the 33rd annual conference of the cognitive science society* (pp. 2295–2300). Austin, TX: Cognitive Science Society.

Eberhard, K. M. (1999). The accessibility of conceptual number to the processes of subject–verb agreement in English. *Journal of Memory and Language, 41*, 560–578.

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

Eschenbach, E., Habel, C., Herweg, M., & Rehkämper, K. (1989). Remarks on plural anaphora. In *Proceedings of the fourth conference of the European chapter of the association for computational linguistics* (pp. 161–167). Manchester.

Ferreira, F., & McClure, K. K. (1997). Parsing of garden-path sentences with reciprocal verbs. *Language and Cognitive Processes, 12*, 273–306.

Foote, R., & Bock, J. K. (2012). The role of morphology in subject–verb number agreement: A comparison of Mexican and Dominican Spanish. *Language and Cognitive Processes, 27*, 429–461.

Franck, J., Vigliocco, G., Antón-Méndez, I., Collina, S., & Frauenfelder, U. H. (2008). The interplay of syntax and form in sentence production: A cross-linguistic study of form effects on agreement. *Language and Cognitive Processes, 23*, 329–374.

Garnham, A. (2001). *Mental models and the interpretation of anaphora.* Hove, England: Psychology Press.

Garnham, A., Oakhill, J., Ehrlich, M.-F., & Carreiras, M. (1995). Representations and processes in the interpretation of pronouns:

New evidence from Spanish and French. *Journal of Memory and Language, 34*, 41–62.

Gernsbacher, M. A. (1991). Comprehending conceptual anaphors. *Language and Cognitive Processes, 6*, 81–105.

Gillespie, M., & Pearlmutter, N. J. (2013). Against structural constraints in subject–verb agreement production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 515–528.

Gillespie, M., & Pearlmutter, N. J. (2011). Hierarchy and scope of planning in subject–verb agreement production. *Cognition, 118*, 377–397.

Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language, 26*, 69–83.

Greenberg, J. H. (1966). *Language universals*. The Hague: Mouton.

Humphreys, K. R., & Bock, J. K. (2005). Notional number agreement in English. *Psychonomic Bulletin and Review, 12*, 689–695.

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, MA: Harvard University Press.

Kaup, B., Kelter, S., & Habel, C. (2002). Representing referents of plural expressions and resolving plural anaphors. *Language and Cognitive Processes, 17*, 405–450.

Koh, S., & Clifton, C. (2002). Resolution of the antecedent of a plural pronoun: Ontological categories and predicate symmetry. *Journal of Memory and Language, 46*, 830–844.

Krifka, M. (1989). Nominal reference, temporal constitution and quantification in event semantics. In R. Bartsch, J. van Benthem, & P. van Emde Boas (Eds.), *Semantics and contextual expression* (pp. 75–115). Dordrecht: Foris.

Landman, R. (1989). Groups, I. *Linguistics and Philosophy, 12*, 559–605.

Lewis, R. L., & Badecker, W. (2010). Sentence production and the declarative and procedural components of short term memory. *Paper presented at the 23rd annual CUNY conference on human sentence processing*, New York, NY.

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science, 2*, 375–419.

Lorimor, H. (2007). *Conjunction and grammatical agreement*. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.

Lorimor, H., Bock, J. K., Zalkind, E., Sheyman, A., & Beard, R. (2008). Agreement and attraction in Russian. *Language and Cognitive Processes, 23*, 769–799.

Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science, 14*, 57–77.

McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review, 86*, 287–330.

McNally, L. (1993). Comitative coordination: A case study in group formation. *Natural Language and Linguistic Theory, 11*, 347–379.

Moxey, L. M., Sanford, A. J., Sturt, P., & Morrow, L. I. (2004). Constrains on the formation of plural reference objects: The influence of role, conjunction, and type of description. *Journal of Memory and Language, 51*, 346–364.

Oakhill, J., Garnham, A., Gernsbacher, M. A., & Cain, K. (1992). How natural are conceptual anaphors? *Language and Cognitive Processes, 7*, 193–400.

Patson, N. D., & Ferreira, F. (2009). Conceptual plural information is used to guide early parsing decisions: Evidence from garden-path sentences with reciprocal verbs. *Journal of Memory and Language, 60*, 464–486.

Patson, N. D., & Warren, T. (2011). Building complex reference objects from dual sets. *Journal of Memory and Language, 64*, 443–459.

Pearlmutter, N. J., Garnsey, S. M., & Bock, J. K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language, 41*, 427–456.

R Development Core Team (2005). *R: A language and environment for statistical computing, reference index version 2.2.1*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. <http://www.R-project.org>.

Sanford, A. J., & Lockhart, F. (1990). Description types and methods of conjoining as factors influencing plural anaphora: A continuation study of focus. *Journal of Semantics, 7*, 365–378.

Schwarzschild, R. (1992). Types of plural individuals. *Linguistics and Philosophy, 15*, 641–675.

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

Staub, A. (2009). On the interpretation of the number attraction effect: Response time evidence. *Journal of Memory and Language, 60*, 308–327.

Stemberger, J. P. (1991). Apparent anti-frequency effects in language production: The addition bias and phonological underspecification. *Journal of Memory and Language, 30*, 161–185.

Van Dyke, J. A., & McElree, B. (2006). Retrieval interference in sentence comprehension. *Journal of Memory and Language, 55*, 157–166.

Vigliocco, G., Butterworth, B., & Semenza, C. (1995). Constructing subject–verb agreement in speech: The role of semantic and morphological factors. *Journal of Memory and Language, 34*, 186–215.

Vigliocco, G., & Franck, J. (1999). When sex and syntax go hand in hand: Gender agreement in language production. *Journal of Memory and Language, 40*, 455–478.

Vigliocco, G., & Franck, J. (2001). When sex affects syntax: Contextual influences in sentence production. *Journal of Memory and Language, 45*, 368–390.

Vigliocco, G., & Hartsuiker, R. J. (2002). The interplay of meaning, sound, and syntax in language production. *Psychological Bulletin, 128*, 442–472.

Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language, 61*, 206–237.