ACCOUNTING FOR SEMANTIC INTEGRATION: NOTIONAL AND GRAMMATICAL EFFECTS IN NUMBER AGREEMENT

BY

LAUREL ELLEN BREHM

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Arts in Psychology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2011

Urbana, Illinois

Master's Committee:

Professor J. Kathryn Bock, Adviser
Professor Gary Dell

ABSTRACT

Notional and grammatical number affect agreement in language production. To explore their workings, we investigated how semantic integration, a type of conceptual relatedness, produces changes in agreement (Solomon & Pearlmutter, 2004). Notional and lexical-grammatical number offer alternative accounts of these effects. 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 three experiments. All three experiments supported the notional hypothesis, with semantic integration or notional plurality creating faster and more frequent singular agreement. This implies that referential coherence mediates the effect of semantic integration on number agreement.

ACKNOWLEDGEMENTS

This research was supported by NSF BCS0843866. I would like to thank Kay Bock for enthusiasm and wisdom; Javier Ospina for technical help and moral support; Maureen Gillespie and Neal Pearlmutter for statistical advice, help with data-interpretation, and for playing devil's advocate; Benjamin Carter for data collection, coding, and data-interpretation in Experiment 3; Linda Salgado and Monica Singh for data collection and coding on Experiments 1 and 2; and Scott Fraundorf and Kristen Tooley for programming and statistics help.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

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. Notional number has to be present in messages, even when it is immaterial to a speaker's communicative intention, because the grammar of English requires it for setting *grammatical number.* Most English nouns have to be either singular (*dog*) or plural (*dogs)*, which in turn determines whether to use singular or plural verbs (e.g. *is* and *are*) and singular or plural determiners (e.g. *this* and *these*). That is, English forces its speakers to represent notional number in messages so that grammatical number, including the lexical and syntactic mechanisms that create singular and plural number-agreement, can function normally. The question in the present investigation was how variations in mental representations of numerosity affect the creation of number agreement during English language production.

Agreement in many languages reflects properties of both notional and grammatical number. At a minimum, notional number must discriminate singletons (individuals, or single things) from aggregates (multiples, or more than one thing). Grammatical number plays a major role because notional number is not predictably related to the singularity and plurality of

sentence subjects, and even is less predictably related to verb number. Instead, the systematic relationship between subject and verb number is captured in grammatical number. The grammatical singular-plural distinction for sentence subjects is tied to a grammatical singular-plural distinction for verbs, so that when nouns serve as subjects, they can trigger different verb forms (e.g. *barks* and *bark*). The result is subject-verb agreement, in which singular subjects take singular verbs (*the dog barks*) and plural subjects take plural verbs (*the dogs bark*).

Because the relationship between notional and grammatical number is imperfect, agreement cannot be explained in simple notional terms. Notional number and grammatical number conflict for many mass nouns, like *toast*, and invariant plurals, like *scissors*, which are always singular and plural, respectively. Making matters worse, notionally plural nouns can be grammatically singular even while looking like grammatical plurals. For instance, the noun *news* takes singular verbs (*The news **was** good*), singular demonstrative determiners (***this** news*), and so forth. These grammatical-number properties of words combine with the notional-number properties of referring expressions to determine subject-verb number agreement (Eberhard, Cutting, & Bock 2004).

Furthermore, some phrases can be either notionally plural or notionally singular, depending on their construal. Consider the difference in meaning between *My sister and best friend **are** in Brooklyn* and *My sister and best friend **is** in Brooklyn.* The same conjoined noun phrase is used in both sentences, but it refers to two people in the former, and one person in the latter. This works for referents that physically combine into one entity as well: *Bacon and eggs **are** full of saturated fat*, but *Bacon and eggs **is** my favorite breakfast.*

In this research, we sought to investigate the interrelations of notional and grammatical number during the formulation of utterances from multi-part messages of contrasting kinds. The

parts of messages can be more or less interrelated, reflecting mental models that are more or less referentially cohesive. In the terminology of Solomon and Pearlmutter (2004), stronger referential linkages result in messages and corresponding linguistic expressions with greater *semantic integration*. For example, *the picture of the flowers* is semantically integrated, likely to reflect things that stand in a part-whole (*flowers-picture*) relationship to each other. In contrast, *the picture with the flowers* may denote a less integrated situation in which a picture and flowers are merely proximal in space (see Figure 1).

Situational properties like these can affect how a speaker construes notional number during the formulation of a message from a mental or perceptual representation. Greater referential cohesion promotes notional singularity; less referential cohesion promotes notional plurality. Nevertheless, grammatically singular linguistic expressions can be formulated that appropriately refer to both states of affairs: The phrase *the picture of the flowers* is grammatically singular, and so is the phrase *the picture with the flowers*. When speakers formulate expressions where notional and grammatical number are incongruent, the implementation of number agreement may become more complicated. The question is how this kind of number complexity matters to agreement. When dealing with complex mental relationships between referents, what counts?

1.2 HOW DOES SEMANTIC INTEGRATION AFFECT NUMBER AGREEMENT?

Solomon and Pearlmutter (2004) proposed a role in agreement for how semantic integration affects the formulation of linguistic expressions from messages. In this proposal, integration has consequences for the process of lexicalization during the creation of phrases. The hypothesis proposed is that the nouns for phrases that express integrated relationships are more

likely to be retrieved in near-synchrony from the lexicon. This creates lexical interference, which has been shown to disrupt and slow access to words which are mentally linked (for related word pairs, see Wheeldon & Monsell, 1994).

When the phrase is a sentence's subject, and the nouns for the phrase differ in grammatical number (as in *picture* and *flowers*), one consequence of this interference is competition for control of verb number. If the "wrong" noun takes control, it may trigger a plural verb when the actual structural properties of the subject would typically yield a plural verb. For example, if *flowers* seizes control of verb number while the phrase *the picture of the flowers* is formulated, a plural verb may result, despite the fact that the head nouns of subject phrases (*picture*) ordinarily control subject-verb agreement. This kind of spurious agreement between subject phrases and verbs is called attraction (Bock & Miller, 1991), and a resulting utterance containing an attraction error might be *the picture of the flowers* ***were*** *ugly*, with a plural verb. Conversely, weak integration in messages is less likely to yield a parallel retrieval process, thereby reducing the occurrence of attraction. The prediction from this lexical interference hypothesis is that strong message integration leads to a higher incidence of spurious agreement, in the form of attraction, than weak message integration.

Solomon and Pearlmutter (2004) tested this hypothesis in an extensive set of language production experiments in which semantic integration was manipulated in five different ways (see Table 1). All of the experiments supported the lexical interference hypothesis: More semantic integration, presumably yielding more interference, resulted in more plural verb agreement, with subjects like *the picture of the flowers* more likely to create attraction than subjects like *the picture with the flowers.*

An alternative view about the involvement of semantic integration in verb agreement comes from the Marking and Morphing account (Bock, Eberhard, & Cutting, 2004, Eberhard et al., 2005). According to Marking and Morphing, a notionally driven number marking for the subject noun phrase (driven from the subject's referent) initiates the agreement process. Marking interacts in principled ways with the lexical specifications of singular and plural grammatical number to arrive at the number value that will control verb agreement. Under this account, the claim is made that there is a singular default for subject number. Some of the major predictions from this account rest on this claim, which means that in most circumstances, all that is needed for singular verb agreement to occur is that (a) notional number is not plural; and (b) none of the nouns retrieved for the subject noun phrase are plural.

This account of agreement motivates what we call the *notional number* hypothesis. Under this hypothesis, strong semantic integration is associated with notional unitization, which has singularity as its default; weak integration is associated with notional aggregation, which triggers plural marking. The effect of these variations on verb number is to raise the probability of singular agreement after well-integrated subjects and to raise the probability of plural agreement after unintegrated subjects. These consequences of a subject's notional number do not constitute attraction (which is driven from the grammatical number of the local noun), but are a kind of semantically motivated agreement.

Evidence consistent with the notional-number account of semantic integration comes from research on the phenomena of collectivity and distributivity (see Figure 1), and with notional agreement to conjoined noun phrases. Collective nouns like *gang* refer to a group of people, containing more than one member. Humphreys and Bock (2005) showed that when a group of individuals is construed in a tight configuration, singular agreement increases in

probability; when construed in a loose configuration, plural agreement is more likely. For instance, when speakers used a sentence preamble like *the gang by the motorcycles,* the agreeing verb was more likely to be singular than with preambles like *the gang on the motorcycles.* A scene in which a gang stands next to some motorcycles tends to be imagined with the gang members grouped together (integrated); when the gang rides motorcycles, however, the group breaks up mentally into individuals (unintegrated). In this respect, conceptual integration promoted singular agreement.

Distributivity is another subtle property of situations that affects number agreement. Distributive and non-distributive construals of reference occur when features are attributed to multiple tokens of a single type or to a single token of a single type, respectively. Consider phrases like *the picture on the postcards* and *the key to the cabinets.* The number features and the syntactic structures of these phrases appear to be the same. By convention, both are grammatically singular. But they differ in how the referents relate to each other. The same picture can appear on many postcards, implicating multiple picture tokens behind *the picture on the postcards*. However, a single key can open many cabinets, so the common construal of *the key to the cabinets* is that just one key is intended. Non-distributive relationships seem to be more integrated than distributive relationships, so the notional hypothesis predicts more singular verbs with *the key to the cabinets* than with *the picture on the postcards.* This finding exactly has been reported often, beginning with Vigliocco, Butterworth, and Garrett (1996) and Eberhard (1999). Distributivity implies less integration among referents and promotes plural agreement; collectivity implies more integration among referents and promotes singular agreement.

Conjoined noun phrases show coalescence due to notional singularity—items become so tightly linked that they turn into one entity. Previous work has shown that abstract nouns are

more likely to coalesce into a notional singleton than concrete nouns (Lorimor, 2007), and that notional singletons are easier to mentally manipulate (Glenberg, Meyer & Lindem, 1987).

Thus, the prediction from the notional number hypothesis is the polar opposite of lexical interference. With increasing semantic integration, the notional number hypothesis predicts increasing rates of singular agreement whereas the lexical interference hypothesis predicts increasing rates of *plural* agreement. Correspondingly, with decreasing semantic integration, the notional number hypothesis predicts increasing rates of plural agreement whereas lexical interference predicts decreasing rates of *singular* agreement.

From the standpoint of the notional number account, then, Solomon and Pearlmutter's (2004) results are paradoxical. Solomon and Pearlmutter found more plural agreement when there was more semantic integration, the opposite of what the notional hypothesis predicts. The upshot is that the notional hypothesis appears to be wrong.

This conclusion may be premature, though. There are patterns in Solomon and Pearlmutter's data that open their results to an interpretation that aligns better with the notional account of semantic integration. In the five experiments taken together, 18% of trials yielded no responses at all. On an additional 13% of the trials, participants failed to produce the presented preamble accurately. Overall, problematic responses were more likely for unintegrated preambles (62% of all missing and distorted responses). For the critical subject noun phrases where attraction was most common (those with singular heads and plural local nouns), 34% of all responses were missing or distorted. These again predominantly affected the unintegrated condition (67% of all missing and distorted responses). These problematic responses obviously reduced the number of usable observations suitable for analysis, and they disproportionately

affected the unintegrated condition. In short, there were fewer responses available for analysis in the unintegrated condition, including fewer instances of attraction.

This correlation means that the incidence of attraction reflected the response rates in different conditions: Attraction may have varied in step with increased production of usable responses, rather than with changes in agreement implementation due to semantic integration. Because more responses occurred in the semantically integrated condition, it could be that greater semantic integration promotes easier conceptualization, easier message formation, and easier production of agreement. This ease is consistent with notional singletons and resulting singular agreement, considering singular as the default number in the absence of conceptual or lexical properties that bias plural agreement. In turn, this is what the notional hypothesis predicts about semantic integration and agreement: The computation of agreement is simpler because integrated mental models make sentence formulation simpler for speakers, rather than making it more complex in the way that lexical interference predicts.

In the first experiment, we tested the predictions of the notional number and lexical interference hypotheses about semantic integration and its effect on agreement. The experimental materials were the same as those of Solomon and Pearlmutter (2004). The experimental task was modified so as to increase the quantity of usable responses and thereby avoid covariations between the likelihood of responding and the likelihood of generating agreement attraction. Speakers produced sentence completions for subject noun-phrases using a predicate adjective from a small set provided. This procedure elicited verbs in a completely incidental fashion, verbs that were nearly always a singular or plural past-tense form of *to be*, either *was* or *were.*

The primary response measure was the latency to produce a number-agreeing verb. This measure directly targets the source of the agreement problem that is predicted to result from

semantic integration under the lexical interference account; that is, the parallel retrieval of nouns and resulting competition for control of number agreement. The lexical interference account predicts that interference and competition for control of verb number should slow the rate of verb production more for semantically integrated than for semantically unintegrated subjects. In contrast, the notional-number hypothesis predicts that verb production should be faster for integrated than for unintegrated subjects, without differences between conditions in the relative rates of attraction.

The second experiment used a task and measure that were identical to those of Solomon and Pearlmutter (2004). Speakers repeated and completed subject noun phrases as complete sentences, and the numbers of singular and plural verbs were counted. This made it possible to assess whether the results from Experiment 1 would generalize to the different measures.

In both these experiments, the experimental materials consisted of semantically integrated or unintegrated subject noun-phrases, or sentence preambles. All preambles contained a singular head noun followed by a prepositional phrase, relative clause, or subordinate clause. In the modifying phrase was another noun (the *local* noun), which immediately preceded the verb. These local nouns varied in grammatical number in both semantic integration conditions. Because plural local nouns create attraction, local plurality increases the probability of producing a plural verb.

The third experiment sought to turn the lexical and notional components of semantic integration into two variables, using conjoined noun phrases that varied on relatedness and abstraction. This made it possible to separately assess the contributions of both processes to agreement: Related words boost lexical interference, while abstract concepts boost notional

unity, and the two processes should operate independently. As in Experiments 1 and 2, lexical interference should increase response time.

The simple predictions for the first two experiments have to do with the effect of plural local nouns. With plural local nouns, the lexical interference hypothesis predicts slower completions for semantically integrated than for semantically unintegrated sentence subjects, and more attraction for semantically integrated than for unintegrated subjects. The notional number hypothesis predicts the opposite. For the third experiment, where the preambles are already proscriptively plural, the simple predictions have to do with the effects of relatedness and abstraction, and by extension the effects of lexical interference and notional number, respectively, on agreement.

CHAPTER 2

EXPERIMENT 1

2.1 INTRODUCTION

The first experiment was designed to examine the time course of verb agreement with semantically integrated and unintegrated subjects. The long-known association between increases in response time and increases in errors should lead to slower responses under the circumstances that typically yield errors in conventional number agreement (specifically errors of attraction), even on occasions when attraction does not actually surface in speakers' utterances. We measured the latency to produce number-agreeing verbs using the same sentence-eliciting preambles as Solomon and Pearlmutter. The preambles were subject noun-phrases like *The picture of/with the flowers* that differed 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 need for explicit instructions. Furthermore, to meet preconditions on the interpretation of reaction time, a wide majority of the elicited verbs were the same in form and grammatical number.

2.2 METHOD

2.2.1 Participants

In exchange for course credit or $7.00 compensation, 94 undergraduates at the University of Illinois participated in this experiment. Participants with fewer than 80% useable experimental

trials were excluded (N=21; see section 2.2.5 “Scoring” for criteria), and one participant was excluded due to recording failure, leaving 72 in total.

2.2.2 Equipment

Stimuli were presented using Psyscope X B53 (Cohen, MacWhinney, Flatt, & Provost, 1993) on a Macintosh Mini computer. They were displayed on a 17-inch LCD flat-screen monitor, and audio was recorded with a Senheiser directional microphone run through a USB button box and Tube MP preamp. The USB button box also recorded latency of vocal responses.

2.2.3 Materials

The 100 experimental items were taken from Solomon and Pearlmutter (2004; see Table 1 for examples). There were 24 representational items, 24 attribute/accompaniment items, 32 relative clause/subordinate clause items, and 20 functional items. Every item had four versions, half of them integrated and half unintegrated. The two integrated and two unintegrated versions of each item were equally divided between the two local noun numbers, with one version having a singular local noun and the other, a plural. For the 20 functional items, Solomon and Pearlmutter had alternative integrated forms that behaved similarly in their experiments, so we selected only one of the integrated forms for each item at random.

There were 152 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. Of these fillers, 46 contained simple noun phrases, 78 contained conjoined noun phrases, and 24 contained a noun phrase with a prepositional phrase. All of the fillers contained singular and/or plural nouns in different structural positions. Four additional fillers were randomly selected from among the original 36 relative clause/subordinate clause stimuli in Solomon and Pearlmutter (2004). Each

of these four fillers was presented in an invariant form divided equally among the four combinations of integration and plurality. Counting the fillers, 114 (45%) preambles had a singular subject noun phrase and 138 (54%) had a plural subject noun phrase.

Thirty of the fillers were placed at the beginning of every experimental list as a covert practice block, 11 singulars and 19 plurals.

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 item paired with the verb *was* and each of the adjectives (e.g., *The book with the red pen was good*). The resulting eight versions of each experimental item were divided equally among eight lists, so that participants viewed only one version of each item, and were counterbalanced so that each participant saw equal numbers of preambles with each adjective. All filler items were used in each list along with adjectives from the set. (See Appendix A for norming instructions.)

We calculated the mean plausibility rating across stimulus-adjective pairings, as well as the means of the most plausibly rated adjective in each stimulus set (the average rating of the best-fitting adjective for each preamble). Across pairs, the average rating was 4.65, and the average most-plausible rating was 6.01, indicating that the adjective set provided suitable completions and that at least one of the four adjectives provided a highly plausible completion for each preamble. The average plausibility rating of the stimuli without predicate adjectives, from Solomon and Pearlmutter was 5.02. Appendix B gives 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 integrated-local singular, integrated-local plural, unintegrated-local singular, and unintegrated-local plural items. Order within lists was quasi-random, with the filler trials appearing in randomized fixed positions across lists. No more than two experimental preambles and no semantically similar items appeared consecutively. Every list was also divided into two halves, with the order of the halves counterbalanced over participants. The same filler items occurred in the same positions in every list, relative to the experimental items.

2.2.4 Procedure

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

The second type of trial sequence is illustrated on the right in Figure 2. 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 milliseconds. When this happened, participants were to repeat back the preamble with an ending. Otherwise, the sequence was the same as for the standard trials. Catch trials were always filler items, and included 8 simple plural noun phrases, 16 conjunctions, and the four relative-clause/subordinate clause stimuli from the Solomon and Pearlmutter set. (Note that the catch-trial structure mimicked the task in Solomon and Pearlmutter's work, allowing these items to serve as a small-scale replication.) Catch trials

were pseudorandomly distributed so that they were distributed evenly across the first and second 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 128 pixels (10%) 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 adjective that made the most sense. Participants received two explicit practice trials and were queried about the set of adjectives and the procedure before beginning the experiment. The experimenter remained in the room for the entire session. Appendix C gives the complete instructions.

2.2.5 Scoring

Response scoring for the reaction-time analyses excluded the catch trials. Responses on all other experimental trials were scored as valid or invalid, 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 voice-key triggering noises preceding the verb. Valid completions were scored as singular when the verb was *was* or *is* (6481 trials, 93%) and as plural when the verb was *were* or *are* (500 trials, 7%). Only singular responses were included in the statistical analyses of reaction time.

Scoring of singular and plural agreement in the spoken completions was carried out on transcriptions of the recorded responses. This scoring encompassed all completions where the first verb produced in the utterance was a number-specifying verb, regardless of whether

disfluencies were present. There were 7249 trials (99%) that met this criterion. Singular verbs occurred on 6726 trials (93%) and plural verbs on 523 trials (7%).

Catch trials were scored according to Solomon and Pearlmutter's scoring guidelines. Trials where the participant repeated the preamble back correctly and said "was," "is," or a third-person singular present-tense verb were marked as singular; trials where the participant repeated the preamble back correctly and said "were," "are," or a third-person plural present-tense verb were marked as plural. Trials in which the participant repeated the preamble back correctly and used a verb uninflected for number were marked as uninflected, and all other produced responses were marked as miscellaneous.

2.2.6 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 (integrated-singular, integrated-plural, unintegrated-singular, and unintegrated-plural). Every item was presented to 18 participants in each of the four 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 on each trial. Reaction times for plural verbs were not included in the statistical analyses because there were too few plural responses for meaningful interpretation. The dependent variable in the plural response analysis was the likelihood of a plural response on each trial.

In addition to treating integration as a dichotomized independent variable, as was done in Solomon and Pearlmutter (2004), we used a binned measure of integration derived from

Solomon and Pearlmutter's integration ratings for each version of each experimental item. Ratings were binned by rounding to the nearest integer. To roughly equate the number 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 (creating a bin for ratings greater than 5.5). All bins contained a mixture of stimuli from each of the two dichotomous categories, providing a more fine-grained distinction.

The binomial (dichotomized) measures of local plurality (singular or plural) and integration (integrated or unintegrated) were contrast coded with the values 0.5 and -0.5. The binned measure of integration was analyzed with Helmert contrasts, weighted to account for the number of responses in each bin. 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 (Bates, 2004; R Development Core Team, 2007).

2.3 RESULTS

2.3.1 Response time analyses

Response time to produce singular verbs was measured as a function of dichotomized integration and local plurality on correct-response trials. Integration was associated with faster response times, as was local singularity: Response times were faster with integrated preambles than their unintegrated counterparts (M= 1040.42 (Integrated); M= 1094.11 (Unintegrated)) and after singular-local-noun preambles compared to plural (M= 1060.01 (Singular); M= 1074.01 (Plural)).

Figure 3 shows response time as a function of binned integration and local plurality on correct-response trials. Overall, local plurality slowed responding, but response speed depended

on the level of integration. Higher levels of integration were also associated with faster response times regardless of local plurality, but at lower levels of integration, response times were slower when the local noun was plural than when it was singular.

The results of a multi-level linear regression are shown in Table 2. An interaction between binned integration and local plurality confirms that local-noun plurality mattered only at low levels of integration, and that at these lower levels, slower reaction times were associated with plural local nouns. That is, for preambles with plural local nouns, lower levels of integration consistently slowed singular sentence completions. At higher levels of integration, local noun plurality had no selective effect. A significant main effect of binned integration ratings confirms that reaction times were slower at lower levels of integration.

The model used in the analysis fit better than alternative models that were evaluated. Specifically, measures of goodness of fit (AIC and BIC values; measures which capture the tradeoff between capturing variance and adding needless complexity) were lower than those for an alternative model that allowed random slopes for integration type by items.

A similar multi-level linear regression was also run with the dichotomized measure of integration. This is reported in Table 2. The general data pattern is similar to the binned measure of integration, with slower reaction times associated with plural local nouns and with low levels of integration. However, no significant interaction was present between the two. The model used in this analysis had lower AIC and BIC values than an alternative model allowing random slopes for integration type.

2.3.2 Plural-response analyses

Table 3 shows descriptive statistics for all singular and plural verbs by dichotomized integration and local plurality. A greater proportion of plurals occurred with unintegrated than with integrated preambles, .09 to .03 respectively.

Figure 4 shows the percentage of plural responses by level of binned integration and local plurality. Plural response rates after singular local nouns were low and relatively consistent across integration level. The difference between singular and plural local nouns was largest at intermediate levels of integration, with more plurals after local plural nouns. As in the reaction time analysis, at high levels of integration there was little effect of singular and plural local nouns.

For the data broken down by the binned integration measure, a weighted empirical logit model of proportion of plural responses weighted by number of observations for each participant was the dependent measure, with items as a random factor. Results of this model can be seen in Table 2. There was a highly significant main effect of local plurality, with more plural responses after a plural local noun, but no clear patterns involving integration. The model employed for this analysis produced a better fit, with lower AIC and BIC values, than a model that included a random intercept for subjects.

A similar model was run for the data broken down by the dichotomous integration measure, a weighted empirical logit model of proportion of plural responses, with proportion of plural responses as the dependent measure and a random intercept for items. Table 2 shows the results. Again, there was a highly significant main effect of local plurality, with more plural responses after a plural local noun. There was also a main effect of integration, with more plural responses after unintegrated preambles. The model employed for this analysis produced a better fit, with lower AIC and BIC values, than a model including a random intercept for subjects.

The four catch trials that used the Solomon and Pearlmutter (2004) materials showed 7 instances (12%) of plural agreement in the integrated-local plural condition. This was more than in each of the other conditions, which were roughly equivalent: One plural occurred in each of the local singular conditions and two in the unintegrated-local plural condition.

2.4 DISCUSSION

The latency measure disclosed slower singular (i.e. grammatically correct) responses at lower levels of semantic integration. That is, agreement was faster with semantically integrated than with semantically unintegrated subjects. The penalty for reduced integration occurred for both singular and plural local nouns. Plural local-noun number slowed responding more than singular local nouns overall, displaying the reaction-time analog of an attraction effect (e.g., Pearlmutter, Garnsey, & Bock, 1999; Wagers, Lau, & Phillips, 2009, and Staub, 2009). This pattern implies greater difficulty with lower levels of integration and plural local-noun number. The outcome is the reverse of what the lexical interference hypothesis predicts about the effect of semantic integration, instead supporting the notional hypothesis. Unintegrated subjects slowed the production of singular verbs, behaving like notional plurals; integrated subjects speeded the production of singular verbs, behaving like notional singletons. Apparently, semantic integration affected agreement production in a way that fits the notional account.

These outcomes occurred with the same materials used by Solomon and Pearlmutter. However, our task differed from theirs, especially in the amount of constraint placed on the sentence completions. Perhaps the need to choose and produce one of a designated set of adjectives changed the implementation of agreement or reduced the plausibility of the completed sentences, especially at lower levels of integration. We evaluated this with a general index of the

plausibility of the possible responses, calculated as the average of the plausibility ratings 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 $r(198)= 0.29$. Though significant ($p < .001$), and in a direction that discounts the conjecture about plausibility, the correlation is weak enough to raise the worry that participants were forced to produce completions that made less sense to them than unconstrained completions would.

A second reservation stems from the agreement responses on the catch trials, which were carried out with the same repeat and complete procedure that Solomon and Pearlmutter used. The catch trials yielded an increase in plural agreement for integrated compared to unintegrated preambles, consistent with Solomon and Pearlmutter's findings. Although this again ameliorates concerns about the effect of differences due to the constraints on completion, it brings the apparent inconsistency between the two sets of results to the fore. With unconstrained completion, the findings were in the direction predicted by the lexical interference hypothesis; constrained completion supported the notional hypothesis. Experiment 2 was designed to address this inconsistency by changing the task from Experiment 1 to one that exactly replicated the Solomon and Pearlmutter procedure.

CHAPTER 3

EXPERIMENT 2

3.1 INTRODUCTION

The response times and errors from Experiment 1 hint at a substantial impact of notional number on agreement, with little support for lexical integration. Nonetheless, the bulk of the data came from a task that differed from Solomon and Pearlmutter's (2004) in several ways. This makes it important to assess whether the implications of Experiment 1 remain the same for unconstrained sentence completions. Experiment 2 used Solomon and Pearlmutter's task on all trials to explore whether and how agreement changes when speakers produce full sentences with completions they devise themselves.

3.2 METHOD

3.2.1 Participants

In exchange for course credit or $7.00 compensation, 50 undergraduates at the University of Illinois participated in this experiment. One participant was excluded due to recording failure, and one was excluded because she could not read the preambles in the time allotted per trial.

3.2.2 Equipment

All equipment was the same as Experiment 1.

3.2.3 Materials

The 100 experimental stimuli were the same as in Experiment 1. The implicit practice block at the beginning was shortened to six trials to reduce the length of the experiment. There

were 228 trials, 128 fillers and 100 experimental trials. Out of all trials in the experiment, including fillers, 53% were plural. Stimuli were presented in the same eight lists as Experiment 1, aside from the initial presentation of six rather than 30 filler trials.

3.2.4 Procedure

As in Experiment 1, participants were presented with a series of preambles on a computer screen, and were instructed to make these into complete sentences as quickly as possible. We asked them to repeat and complete the preamble using a termination that made sense to them. This was akin to the catch trials in Experiment 1, with no restrictions on predication. Participants were given one explicit practice trial, and were queried about the procedure before beginning the experiment. (See Appendix D for the instructions and explicit practice trials.) The time course of events within the trials was nearly identical to the catch trials from Experiment 1: 500 additional milliseconds of blank screen was presented instead of "Repeat", but all else was the same.

3.2.5 Scoring

The category of the response (singular/plural) was coded by hand according to the guidelines in Solomon and Pearlmutter (2004); which were the same as the procedure for scoring catch trials in Experiment 1. As in Experiment 1, singular and plural responses were coded. Singular and plural verbs occurred on 66% of trials, and of these responses, 92% were singular.

3.2.6 Design and data analysis

The design was identical to Experiment 1. The dependent variable was the plurality of the response, as outlined above.

Statistical analysis was carried out as in Experiment 1. Weighted empirical logistic regression was carried out using the lme4 package in R on the proportion of singular and plural responses by dichotomized integration, binned integration, and local plurality. The binomial

measures were contrast coded with the values of 0.5 and -0.5, and the binned measure of integration was analyzed with weighted Helmert contrasts.

3.3 RESULTS

Table 3 shows descriptive statistics for each scoring category by dichotomized integration and local plurality. More plural responses were associated with unintegrated preambles, as were more miscellaneous errors. The proportion of uninflected verbs is largest in the integrated condition, particularly with singular local nouns.

Figure 5 shows the proportion of plural responses by local plurality and dichotomous integration. More plural responses were present with a local plural noun, and with integrated stimuli. However, the difference in the number of plural responses between the local singular and local plural conditions was greater with integrated than unintegrated stimuli. This pattern replicates the attraction effect found in Solomon and Pearlmutter (2004).

This pattern was assessed statistically with a weighted empirical logit model. A weighted proportion of plural responses was the dependent measure, with subjects as a random factor. The results from this model are reported in Table 4. There was a significant interaction between local plurality and dichotomous integration. That is, there was a larger difference in the number of plural responses produced between the local plural and local singular conditions at high levels of integration than at low levels of integration. Also present was a main effect of local plurality, with more plural responses after a local plural noun, and a main effect of integration, with more plural responses to unintegrated preambles. This model produced a better fit, with lower AIC and BIC values, than a model including a random intercept for items.

A similar model was fitted using the binned measure of semantic integration. These patterns were assessed statistically with weighted empirical logistic regression. An empirical logit of proportion of plural responses weighted by number of observations was the dependent measure, with items as a random factor. The results from this model are reported in Table 4. As in the dichotomous integration model, there was a main effect of local plurality, with more plural responses after a local plural noun. Additionally, there were significant differences between the lowest levels of integration and the highest levels, with low levels of integration promoting more plural responses. The proportion of plurals is highest when following a local plural at high levels of integration. This interaction that is marginally significant at the midpoint of the scale, comparing bin 3 to the higher bins. This model produced a better fit, with a lower AIC and BIC values, than a model including a random intercept for subjects.

3.4 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. Put the other way around, less-integrated sentence subjects promoted plural agreement. Viewed as a product of attraction, these results are the reverse of what the lexical interference hypothesis predicts. Viewed instead as a consequence of variations in notional number, the findings align with a notional account of semantic integration. Perceptually and conceptually coherent states of affairs are more compatible with singular than to plural number, whereas perceptually and conceptually disparate states of affairs are more compatible with plural than singular number. The implication is that the effect of semantic integration on agreement has less to do with attraction and more to do with a speaker's evaluation of notional number.

The immediate question is why our results reverse those of the initial studies of semantic integration. As described in the introduction, the present experiments had a strikingly lower number of unusable responses. Table 5 displays the incidence of such responses in Experiment 2 against the combined incidence in Solomon and Pearlmutter's experiments. There were patently fewer failures to respond in Experiment 2, with the greatest disparity arising in the larger number of responses in the unintegrated condition. Though plural production remained rare, the increase in plural agreement relative to the integrated condition is one byproduct of the generally higher rate of producing inflected verbs (both singular and plural) relative to the Solomon and Pearlmutter experiments.

One conjecture about this difference is that less-integrated sentence subjects make it harder to form suitable mental models for generating any kind of utterance. This difficulty may be more likely to arise with external pressure for fast responding, which could have been greater in Solomon and Pearlmutter's experiments than in ours. Although Experiment 1 found no evidence that faster responding led to less consistent agreement, any efforts at increased speed were made under the internally generated motivation of the participants. With continuously applied external pressure, as with dual-task performance, inconsistent performance could create a tradeoff between speed and accuracy (Pachella, 1974) that surfaces as increased attraction.

In one crucial respect, the results of Experiment 2 were consistent with Solomon and Pearlmutter's data. There was a significant interaction between integration and local noun number, such that the increase in plural responding was greater in the integrated than in the unintegrated condition. This increase is analogous to a common interaction seen between notional and grammatical factors in other work on agreement. To explore this interaction further, we examined plural responding as a function of the binned integration variable used with the

reaction times in Experiment 1. The measure of plural responding was the standard index of attraction, the difference between the singular and plural local noun conditions in the use of plural verbs (calculated as the proportion of plurals among all singular and plural responses in each condition). Figure 6 displays the results. At low levels of integration, there is no evidence for a difference in plural agreement between the conditions, with the only apparent trend a decrease in plural verbs predicted from notional-number effects. However, at higher levels of integration, an attraction effect emerges in a decreased slope for integrated sentence subjects.

On its face, this trend in attraction supports the lexical interference hypothesis. However, lexical interference has nothing to say about the strong agreement effects that are consistent with notional number variations. In contrast, the Marking and Morphing theory (Eberhard et al. 2005) predicts both patterns in a unified way that depends on interactions between notional and grammatical number, without positing lexical interference. In the General Discussion, we lay out how Marking and Morphing explains the data.

CHAPTER 4

EXPERIMENT 3

4.1 INTRODUCTION

In the previous two experiments, we observed that semantic integration promotes both notional agreement and attraction. This attraction effect could be explained through lexical interference, which does not speak to the notional effects also observed in the data, or it could be explained through Marking and Morphing terms, which speaks to both effects.

To determine the origin of the attraction effect in the first two experiments, the third experiment used conjoined noun pairs as preambles. Conjoined noun phrases are prescriptively plural in English, but this plurality needs to be computed from the sentence structure rather than from the plurality of any one noun in the phrase. This leaves considerable room for notional agreement—consider the example with *Bacon and eggs* laid out in the general introduction. Furthermore, since conjoined noun phrases are made up of just one clause, rather than two clauses, all parts should be processed nearly simultaneously, potentially boosting lexical interference (e.g. Allum & Wheeldon, 2007).

To differentiate between the notional component of semantic integration and the interference-producing component of semantic integration, we varied nouns in the conjoined noun phrases over two dimensions. The first dimension was abstractness. Abstractness reduces tendency toward notional plurality (e.g Lorimor, 2007). A conjoined pair of abstract nouns is less likely to promote plural notional agreement than a conjoined pair of concrete nouns. For this reason, concrete nouns should promote faster plural agreement: The more sources of plurality there are, the faster agreement should be. The second dimension was semantic relatedness.

Related words promote semantic interference (e.g. Wheeldon & Monsell, 1994). This semantic interference should boost lexical interference, resulting in slower agreement under the lexical interference hypothesis.

Local plural nouns here, as in the first two experiments, should increase plural agreement. Plural nouns contribute to the Morphing process as sources of plurality. Thus, the presence of plural nouns makes the likelihood of plural verb agreement increase.

4.2 METHOD

4.2.1 Participants

In exchange for course credit or $7.00 compensation, 113 undergraduates from the University of Illinois at Urbana-Champaign participated in the study. 13 participants were excluded from the study for having 15.6% (five or greater) unusable experimental trials, as described in "Scoring" below, and an additional four participants were excluded due to technical difficulties.

4.2.2 Materials

The experimental preambles were conjoined noun phrases varied on abstractness, relatedness, and local noun plurality, each with two levels. These preambles were created using a free association word database containing normed relatedness and abstraction ratings (Nelson, McEvoy, & Schreiber, 1998). A list of preambles is in Appendix E. Abstract preambles contained two nouns rated between one and three on a seven-point scale, while concrete preambles contained two nouns rated between five and seven on the same scale. The head noun was kept constant, and related and unrelated versions of preambles were also developed, with related preambles having an association rating between 30% and 50% for the pair and unrelated

preambles having an association rating between 0.01% and 10% for the pair. All words in the preambles but one (tooth/teeth) were selected as singular nouns from the database, and only nouns with a meaningful plural were selected, regardless of their status as the head or local noun.

Norming was carried out to establish the fit of the set of nouns with the set of adjectives from Experiment 1 (*good, bad, ready* and *true).* The singular form of each noun was presented in the phrase *How likely is it that X is Y* with each adjective (e.g. *How likely is it that a vegetable is bad?)*. These phrases were divided into 4 lists, each containing all nouns once, and an equal number of instances of each adjective. Participants were asked to rate the likeliness of the pairing on a five-point scale (Not likely = 1). Twenty participants contributed to the norming, with each noun-adjective pair presented to 5 participants. The average rating for nouns across adjectives was 2.9 (range 2.2 to 3.8), and the average of the best-fitting adjective for all nouns was 4.9 (range 3.5 to 5), suggesting that as a whole, the set of adjectives was somewhat plausible, with at least one well-fitting adjective for each noun. An average absolute value of head-local differences across adjectives was calculated to provide a measure of predicability ease for the word pair; that is to say, the facility with which the set of adjectives matches the preamble. This average adjective difference score was 0.87 for concrete noun pairs and 0.59 for abstract noun pairs, suggesting that the abstract noun pairs may be easier, and therefore faster, to complete with this set of adjectives. To control for this, an adjective fit measure was entered into response time analyses that comprised of the average rating of all adjectives with the head and local nouns for each stimulus.

For the experimental stimuli, the four different versions of the preambles were divided into four lists, and as in the previous experiments, they were counterbalanced such that all lists contained only one version of each item, with an equal number of items of each type. Order

across the lists was quasi-random, with filler trials appearing in randomized fixed positions across lists, with no more than two experimental preambles and no semantically similar items appearing consecutively. The same filler items occurred in the same positions in every list, relative to the experimental items. Every list was also divided into two halves, with the order of the halves counterbalanced over participants, for a total of eight lists.

4.2.3 Equipment

All equipment was the same as Experiment 1 and Experiment 2.

4.2.4 Procedure

The procedure was nearly identical to Experiment 1, differing only in the critical trials and the proportion of fillers. Participants read preambles and completed them with an adjective from the set *good, bad, ready,* and *true.* However, there were 32 experimental trials, and 192 filler trials, for a total of 224 trials. 34 of the filler trials were catch trials. The implicit practice block was 12 trials long. Out of all the trials in the experiment, including fillers, 104 were plural. Instructions and explicit practice trials were identical to Experiment 1 (Appendix C).

4.2.5 Scoring

Scoring was as in Experiment 1 and Experiment 2. Experimental trials were excluded from response time analyses as in Experiment 1. Plural responses occurred on 86% of trials, and singular responses occurred on 14% of trials.

4.2.6 Design and data analysis

Every participant received exactly one version of each of the 32 experimental preambles, four preambles in each of the eight combinations of relatedness, abstraction and local–noun number. Every item was presented to 24 participants. The fixed effects in the statistical analyses were relatedness (dichotomized: related-unrelated, linear: normed rating from 0.01 to 0.50),

abstraction (dichotomized: abstract-concrete, linear: normed rating from 1 to 7), local-noun number (singular-plural), adjective fit, and their interactions.

As in Experiment 1, the dependent variable in the reaction time analysis of the plural verbs was the amount of time taken to initiate the verb from the onset of cue presentation to the onset of speech on each trial. Reaction times for singular verbs were not included in the statistical analyses because there were too few singular responses for meaningful interpretation. As in Experiment 1 and Experiment 2, the dependent variable in the response type analysis was the likelihood of a plural response on each trial.

Statistical analyses were carried out as in the previous experiments. Multi-level linear and weighted empirical logistic regression were carried out using the lme4 package in R. All binomial predictor variables were contrast coded with the values of 0.5 and -0.5.

4.3 RESULTS

4.3.1 Response time analyses

Response time to produce plural verbs was measured as a function of dichotomized relatedness, abstraction, and local plurality on correct-response trials. Abstraction was associated with faster response times: Response times were faster with abstract preambles than concrete ones (M= 1130 (Abstract); M= 1158 (Concrete)). Local plurality was also associated with faster response times: Response times were faster with local plural nouns than local singulars (M= 1159 (Plural); M= 1169 (Singular). Relatedness was associated with slightly slower response times, with response times slower with related preambles than their unrelated counterparts (M= 1146 (Related); M=1143 (Unrelated)), but this difference was extremely small.

Figure 7 shows response time as a function of dichotomized relatedness, dichotomized abstraction, and adjective fit on correct-response trials. Overall, abstract preambles were associated with slightly faster response times than concrete ones, but most response times were fairly constant across all levels of abstraction, relatedness, and adjective fit. The one outlying point is in the concrete-unrelated trials at high levels of adjective fit—these response times were significantly slower than in any other condition.

The results of a multi-level linear regression are shown in Table 6. A three-way interaction between linear relatedness, linear abstraction, and adjective fit confirms this response time penalty at highly concrete, highly related, and high levels of adjective fit. The model used in the analysis fit better than alternative models that were evaluated, though all models showed similar patterns. AIC and BIC values were lower in this model than in models which contained the dichotomized relatedness and abstraction variables, or models with any of the variables eliminated.

4.3.2 Plural-response analyses

Table 7 shows descriptive statistics for all singular and plural verbs by dichotomized relatedness, dichotomized abstraction, and local plurality. A greater proportion of plurals occurred with local plural nouns. Abstract preambles elicited fewer plural responses, particularly in the local singular noun condition.

A weighted empirical logit model of proportion of plural responses weighted by number of observations for each participant was the dependent measure, with items as a random factor, and dichotomized abstraction, dichotomized relatedness and local plurality as predictors. Results of this model can be seen in Table 6. There was a highly significant main effect of local plurality, with more plural responses after a plural local noun, and a trend towards a main effect of

abstraction, with fewer plural responses to abstract preambles. The model employed for this analysis produced a better fit, with lower AIC and BIC values, than a model that included a random intercept for subjects, or models including the linear relatedness and abstraction variables.

4.4 DISCUSSION

The responses produced in Experiment 3 parallel the previous two experiments. In this experiment, as in the previous two, local plural nouns were associated with plural agreement, while local singular nouns showed an increased tendency to produce singular agreement. This can be directly attributed to the increased plural value that is added up in the Morphing process during the formation of agreement.

No significant notional effects were shown here, in contrast to Experiments 1 and 2. Abstract preambles were weakly associated with singular agreement, though this was not statistically significant.

Importantly, Experiment 3 showed no lexical interference effect. In fact, preambles that are highly concrete, highly unrelated, and have a high adjective fit show the largest response times of all. This can be attributed to predication difficulty, the difficulty of choosing an appropriate verb and adjective to describe two distinct concepts. This effect of predication difficulty mirrors the large number of missing responses in the unintegrated conditions in Solomon and Pearlmutter's experiments.

In addition to adjective fit, a second variable was entered into the response time analyses in attempt to determine the influence of the predicate adjectives on agreement. This measure was a ranked-signed-difference score of the best-fitting adjective for the two nouns in each pair, and

was created by taking the absolute value of the difference between the best-fitting adjective ratings for the head and the local noun, adding a negative sign if the best-fitting adjective was different for the two nouns, and ranking the scores in order from lowest (-29) to highest (28). This measure produced similar results to the average adjective fit measure, with low (widely differing) ranked-signed-difference scores interacting with high relatedness to produce slower responses.

Abstraction slightly speeded responding, contrary to the predictions from Marking and Morphing. This may be a result of the notional unity of abstract preambles, and the fact that notional singletons are easier to mentally manipulate (Glenberg, Meyer, & Lindem, 1987).

Interestingly, even with a local plural noun, there was a non-negligible rate of singular agreement in this experiment, suggesting that conjunctions, no matter what their content, tend to be less strongly plural than regular plural noun phrases: That is, *cat and dog* is less plural than *dogs*. This is perhaps due to the contribution of Morphing to plural agreement, and the conceptual difficulty of computing plurality from structures rather than words.

Taken together, these results suggest that the Marking and Morphing model neatly describes agreement in complex referential situations, and that conceptual conflict slows response times.

CHAPTER 5

GENERAL DISCUSSION

The results of the first two experiments show that participants respond more slowly, using more plural responses given a low level of semantic integration, regardless of local noun number. This penalty of semantic *dis*-integration reflects notional agreement. The results of the notional portion of the third experiment, though not significant, are numerically consistent with the first two, showing that participants use more plural responses with (notionally plural) concrete noun pairs.

Experiment 1 demonstrated the relationship between difficulty and semantic integration. Slower responses, suggesting greater difficulty, were associated with local plural noun-containing and unintegrated preambles, by both dichotomized and continuous integration measures. Plural agreement, a measure of difficulty and notional agreement, was also associated with unintegrated preambles. Experiment 2 demonstrated the relationship between agreement and sematic integration, replicating the procedure and measures used in previous work (Solomon & Pearlmutter, 2004, Gillespie & Pearlmutter, 2011). Plural responses, indicative of notional agreement, were associated with low levels of integration. Experiment 3 demonstrated the relationship between agreement, a different source of notional plurality, and lexical interference. Plural responses were associated with the conditions with the most sources of plurality (from Marking and Morphing), and response times were slow when predication was particularly difficult for the stimulus.

Semantic integration is concerned with the relationship between items in a referential context, and is linked to unitization and the resulting notional singularity of a mental message.

The results of the present experiments suggest that differences in semantic integration are in fact associated with differences in notional number, with corresponding effects on number agreement.

Semantic integration seems to impact notional agreement by dictating what types of responses (singular, plural, or uninflected) are plausible. This is similar to some studies on notional effects on pragmatic interpretation in comprehension. Ferriera and McClure (1997) and Patson and Ferriera (2009) found no garden path effect on a verb that would be implausible with a conjoined noun phrase. That is, in *After Jose and the bride kissed the party began in earnest*, participants were able to determine that *kissed* would not take an additional argument, making the sentence unambiguous. Kaup, Kelter, and Habel (2002) find that notional information derived from pronouns *sie (they* in German) and *biede (both* in German) can be used to determine the number of people doing an action. This number information relates to the way that concepts are licensed to fit together.

The speeding of responses following semantically integrated stimuli follows from classic work on mental models which shows that comprehension is facilitated when referents are spatially associated, or notionally singular, as compared to spatially dissociated, or notionally plural (Glenberg, Meyer, & Lindem, 1987). Thus, the pragmatic information given by notional connectedness and the corresponding numerosity of referents is used to inform processing. Though our results are unlike other notional effects on agreement errors (e.g., Bock & Eberhard, 1993), the observed notional effects on agreement in our experiments seem to reflect conceptual processing.

How can the observed notional effects be reconciled with previous work on semantic integration? As stated earlier, one possibility is that notional effects were present in the data

collected by Solomon and Pearlmutter, but were overshadowed by the task demands. In their data set, there were a large number of mistakes that speakers made in their task which did not enter into the analysis of agreement. Specifically, uninflected and miscellaneous responses were most common at low levels of integration, inversely co-occurring with the observations of attraction (i.e., plural agreement) which constitute agreement errors. This suggests that if difficulty is measured in uninflected and miscellaneous responses, unintegrated preambles are harder than integrated preambles, as we saw in our data. Experiment 3 speaks to this issue: The condition with the most predication difficulty had the slowest response times.

Another related possibility is that our participants had an easier time with the task: Comparing the percentages of uncategorizable responses in our Experiment 2 (10%) to the average of Solomon and Pearmutter's five experiments (31%), it is obvious that the participants in our experiments had less difficulty repeating and completing the experimental preambles. This in turn suggests that our participants were better able to conceptualize the situations referred to by the subject noun phrase: They had a clearer conception of what *the book with the red pens* entailed. This not only increased the overall number of responses, but also supported the effects of notional number that we observed.

One difference between our first two experiments and those of Solomon and Pearlmutter is that we took stimuli from several of their experiments, collapsing across a variety of sentence structures. We found differences between the different types of integration that came from separate experiments in the Solomon and Pearlmutter studies. Particularly, there was a high number of uninflected responses with the functional stimuli in our Experiment 2. These stimuli represent relationships between people (e.g., *the chauffeur for the actors).* Because specific, lexical verbs are more likely to describe people, due to the higher levels of predicability

associated with animates (see Bock, Loebell, & Morey, 1992, for review)**,** the increase in uninflected responses with highly integrated stimuli in this set suggests an interaction between this bias and conceptual features of the mental model of the message, supporting a notional, conceptually-driven account of agreement. For two sets of our stimuli, we saw less evidence for notional effects. Representational stimuli (e.g. *The drawing of/with the flowers)* and relative clause-subordinate clause stimuli (e.g. *The report that (Megan) described the accident*) showed increased plural agreement for integrated sentences, in both of our experiments 1 and 2, opposite from the rest of our data. This also fits with a conceptually-driven account of agreement: These stimuli feel harder to visualize than many of the other sentences, causing difficulty with conceptualization. This conceptualization difficulty may have caused notional number to be ignored, leaving room for a pure lexical interference effect.

Furthermore, there was some evidence for a lexical interference effect in the first two experiments. In Experiment 2, we observed an interaction between the dichotomous measure of semantic integration and local plurality, with high levels of integration promoting a larger difference between the two local noun conditions in the proportion of plural responses. This is consistent with a lexical interference effect. In Experiment 1, there was no significant interaction, but both response time and the proportion of plural responses followed the same numerical pattern consistent with lexical interference, with high levels of integration promoting a larger difference between the two local noun conditions

Though this interaction could be explained through lexical interference, as outlined previously, it could also be explained parsimoniously through the processes of Marking and Morphing, accounting for both notional and attraction effects simultaneously under the same model. Marking and Morphing posits that singular notional number is unmarked, making it the

default state for the system. Noun phrases with singular notional number are therefore susceptible to attraction from plurals within the phrase, creating erroneous plural agreement. In contrast, plural notional number is marked. This marking means that the notional number for a noun phrase is specified as plural and the noun phrase is not susceptible to attraction effects from morphing. This predicts an interaction effect between integration and local noun number such that attraction is only licensed given notional singularity and a local plural noun, which is the pattern observed in the current studies. Marking and Morphing can therefore account for the effects of semantic integration without appealing to lexical interference, parsimoniously explaining the full pattern of the data in both experiments.

Plural responses in agreement are associated with two different kinds of effects. First, they reflect plural notional agreement, correct agreement based upon the semantic properties of the referent. Second, they reflect attraction, a type of grammatical error. Comparing the proportion of plural responses in Experiment 2 with different local nouns allows us to separate the two kinds of difficulty. Plural responses after a singular local noun cannot come from attraction, as there is no attractor. These responses are purely influenced by notional agreement, and they occur only at low levels of integration. In contrast, plural responses after a plural local noun could be influenced by attraction from the plural local noun, or they could be influenced by notional agreement. Plural responses after a plural local noun occurred over the entire integration range, from low to high, but increased with low integration. This increase can be attributed to the same notional properties that increase plural responses after a singular local noun, leaving the plural responses at high levels of integration to be based on attraction. These results across the entire range of integration are more consistent with a notional agreement mechanism, as in the Marking and Morphing model than the lexical interference mechanism proposed by Solomon

and Pearlmutter.

In conclusion, the evidence suggests that semantic integration supports differences in notional agreement rather than differences in attraction. This is the opposite of Solomon and Pearlmutter's conclusion. Semantic integration made sentence completion easier, as measured by reaction time and sentence completion responses. Semantic integration likewise seems to have increased notional singularity, making intuitively simpler situations, while the absence of integration did the opposite, increasing notional plurality and yielding plural agreement. Attraction was observed as well, but when the contribution of dis-integration to notional agreement is factored in, attraction was the product of grammatical plurality of local nouns. Though we cannot conclude that semantic integration is irrelevant to the processes of lexicalization, we can say with some confidence that semantic integration *is* relevant to the construal of notional number.

TABLES AND FIGURES

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

Stimulus Type by experiment	Integrated		Unintegrated	
	Local Singular	Local Plural	Local Singular	Local Plural
	Representational			
Experiment 1	The drawing of the flower	The drawing of the flowers	The drawing with the flower	The drawing with the flowers
	Functional			
Experiment 2	The translator of the ambassador	The translator of the ambassadors	The translator with the ambassador	The translator with the ambassadors
Experiment 3	The translator for the ambassador	The translator for the ambassadors	The translator with the ambassador	The translator with the ambassadors
	Attribute/ Accompaniment			
Experiment 4	The apple with the brown spot	The apple with the brown spots	The apple with the fresh peach	The apple with the fresh peaches
	Relative clause/ Subordinate clause			
Experiment 5	The report that described the traffic accident	The report that described the traffic accidents	The report that Megan described the accident	The report that Megan described the accidents

Table 2. Parameters for experiment 1 reaction time and plural response models.

Reaction time model (binned integration)	Estimate	95% CI Lower	95% CI Upper	p value (MCMC)
Intercept	1072.28	1031.19	1112.54	< 0.001
Local plurality	-13.36	-29.03	3.00	0.1
Integration				
Level 5 versus 6	10.02	-22.73	39.22	0.58
Level 4 versus 5 & 6	27.01	0.34	55.84	< 0.05
Level 3 versus 4, 5, & 6	5.51	-28.51	45.30	0.67
Level 2 versus 3, 4, 5, & 6	46.34	14.85	79.41	< 0.001
Local plurality x Integration				
Plurality x Int level 5 vs 6	29.16	-28.92	81.93	0.32
Pl x Int level 4 vs 5 & 6	40.33	-8.18	88.14	< 0.1
Pl x Int level 3 vs 4, 5 & 6	-78.01	-141.85	-16.79	< 0.05
Pl x Int level 2 vs 3, 4, 5, & 6	-42.67	-96.54	4.16	< 0.1

Reaction time model (dichotomous integration)	Estimate	95% CI Lower	95% CI Upper	p value (MCMC)
Intercept	1072.43	1032.69	1114.17	< 0.0001
Local plurality	-15.17	-30.54	1.19	0.06
Integration	-55.21	-84.36	-26.62	< 0.0001
Local plurality x Integration	-8.58	-38.17	23.16	0.58

Table 2 (cont.)

Plural response model (binned integration)	Estimate	95% CI		p value (MCMC)
		Lower	Upper	
Intercept	-2.26	-2.14	-1.95	< 0.001
Local plurality	0.75	0.61	0.95	< 0.001
Integration				
Level 5 versus 6	0.07	-0.22	0.32	0.68
Level 4 versus 5 & 6	0.06	-0.04	0.42	0.13
Level 3 versus 4, 5, & 6	0.06	-0.2	0.38	0.58
Level 2 versus 3, 4, 5, & 6	-0.19	-0.52	0.01	0.06
Local plurality x Integration				
Plurality x Int level 5 vs 6	-0.16	-0.68	0.39	0.55
Pl x Int level 4 vs 5 & 6	0.4	0.01	0.92	< 0.05
Pl x Int level 3 vs 4, 5 & 6	-0.02	-0.47	0.67	0.74
Pl x Int level 2 vs 3, 4, 5, & 6	-0.16	-0.79	0.28	0.37

Plural response model (dichotomous integration)	Estimate	95% CI		p value (MCMC)
		Lower	Upper	
Intercept	-2.11	-1.97	-1.74	< 0.0001
Local plurality	0.70	0.48	0.93	< 0.0001
Integration	-0.32	-0.61	-0.25	< 0.0001
Local plurality x Integration	0.18	-0.06	0.64	0.12

Table 3. Responses by scoring category and experiment.

Experiment	Integration Level	Local Noun	Singular	Plural	Uninflected	Misc	Proportion Plural
Experiment 1	Integrated	Singular	1764	42	0	13	.02
		Plural	1650	162	0	16	.09
	Unintegrated	Singular	1728	98	0	7	.05
		Plural	1584	221	0	13	.12
Experiment 2	Integrated	Singular	785	2	304	90	0.00
		Plural	756	77	257	86	0.07
	Unintegrated	Singular	701	74	302	98	0.06
		Plural	687	107	266	116	0.09

Table 4. Parameters from Experiment 2 plural response models.

Plural response model (binned integration)	Estimate	95% CI		p value (MCMC)
		Lower	Upper	
Intercept	-1.98	-1.82	-1.62	< 0.001
Local plurality	0.53	0.34	0.73	< 0.001
Integration				
Level 5 versus 6	-0.11	-0.33	0.27	0.86
Level 4 versus 5 & 6	-0.08	-0.26	0.26	0.97
Level 3 versus 4, 5, & 6	0.27	0.05	0.67	< 0.05
Level 2 versus 3, 4, 5, & 6	-0.01	-0.25	0.29	0.81
Local plurality x Integration				
Plurality x Int level 5 vs 6	-0.16	-0.68	0.52	0.76
Pl x Int level 4 vs 5 & 6	0.13	-0.44	0.59	0.72
Pl x Int level 3 vs 4, 5 & 6	-0.60	-1.09	0.16	0.15
Pl x Int level 2 vs 3, 4, 5, & 6	0.07	-0.61	0.47	0.82

Plural response model (dichotomous integration)	Estimate	95% CI		p value (MCMC)
		Lower	Upper	
Intercept	-2.22	-2.28	-1.92	< 0.0001
Local plurality	0.87	0.65	1.23	< 0.0001
Integration	-0.93	-1.26	-0.68	< 0.0001
Local plurality x Integration	1.30	0.77	1.93	< 0.0001

Table 5. Proportion of response types in Solomon and Pearlmutter (2004); Experiment 2.

Experiment	Integration Level	Singular	Plural	Miscelaneous	Missing
Solomon & Pearlmutter	Integrated	0.39	0.04	0.11	0.13
	Unintegrated	0.33	0.02	0.15	0.24
Experiment 2	Integrated	0.64	0.03	0.07	0.02
	Unintegrated	0.58	0.08	0.09	0.02

Table 6. Parameters from Experiment 3 reaction time and plural response models.

Response time model	Estimate	95% CI		p value (MCMC)
Linear abstraction and relatedness		Lower	Upper	
Intercept	1166.93	1120.68	1212.22	< 0.001
Local plurality	-2.85	-29.98	23.78	0.87
Abstraction	7.85	-6.95	21.67	0.27
Relatedness	-120.19	-321.76	77.08	0.24
Adjective fit	-53.55	-115.17	6.94	0.09
Local pl x Abstraction	-11.01	-24.75	3.65	0.14
Local pl x Relatedness	-326.28	-543.15	-86.75	< 0.01
Local pl x Adjective fit	87.98	-34.31	194.97	0.15
Abstraction x Relatedness	-12.70	-131.34	98.75	0.83
Abstraction x Adjective fit	-28.45	-57.83	3.91	0.08
Relatedness x Adjective fit	-526.39	-948.58	-85.20	< 0.05
Local pl x Abstraction x Relatedness	-110.62	-237.07	22.51	0.12
Local pl x Abstraction x Adjective fit	41.90	-21.60	92.59	0.18
Local pl x Relatedness x Adjective fit	-81.80	-986.16	912.91	0.95
Abstraction x Relatedness x Adjective fit	-486.31	-696.87	-255.72	< 0.001
Local pl x Abstraction x Relatedness x Adjective fit	42.84	-429.30	602.18	0.78

Table 6 (cont.)

Plural response model	Estimate	95% CI		p value (MCMC)
Dichotomized abstraction and relatedness		Lower	Upper	
Intercept	-3.57	-3.63	-3.45	< 0.001
Local plurality	0.19	0.02	0.37	< 0.05
Abstraction	-0.13	-0.31	0.05	0.18
Relatedness	-0.11	-0.29	0.06	0.20
Local pl x Abstraction	-0.21	-0.53	0.15	0.29
Local pl x Relatedness	0.18	-0.15	0.54	0.28
Abstraction x Relatedness	0.04	-0.35	0.36	0.92
Local pl x Abstaction x Relatedness	-1.01	-3.55	1.90	0.50

Table 7. Experiment 3 responses by scoring category.

Abstraction	Relatedness	Local Noun	Singular	Plural	Uninflected	Misc	Proportion Plural
Abstract	Related	Singular	125	303	0	0	0.71
		Plural	20	406	1	0	0.95
	Unrelated	Singular	102	324	0	1	0.76
		Plural	30	398	0	0	0.93
Concrete	Related	Singular	64	360	2	1	0.84
		Plural	22	402	1	0	0.95
	Unrelated	Singular	75	351	1	0	0.82
		Plural	24	402	1	1	0.94

Figure 1. Notional plurality in semantic integration, distributivity and collectivity.

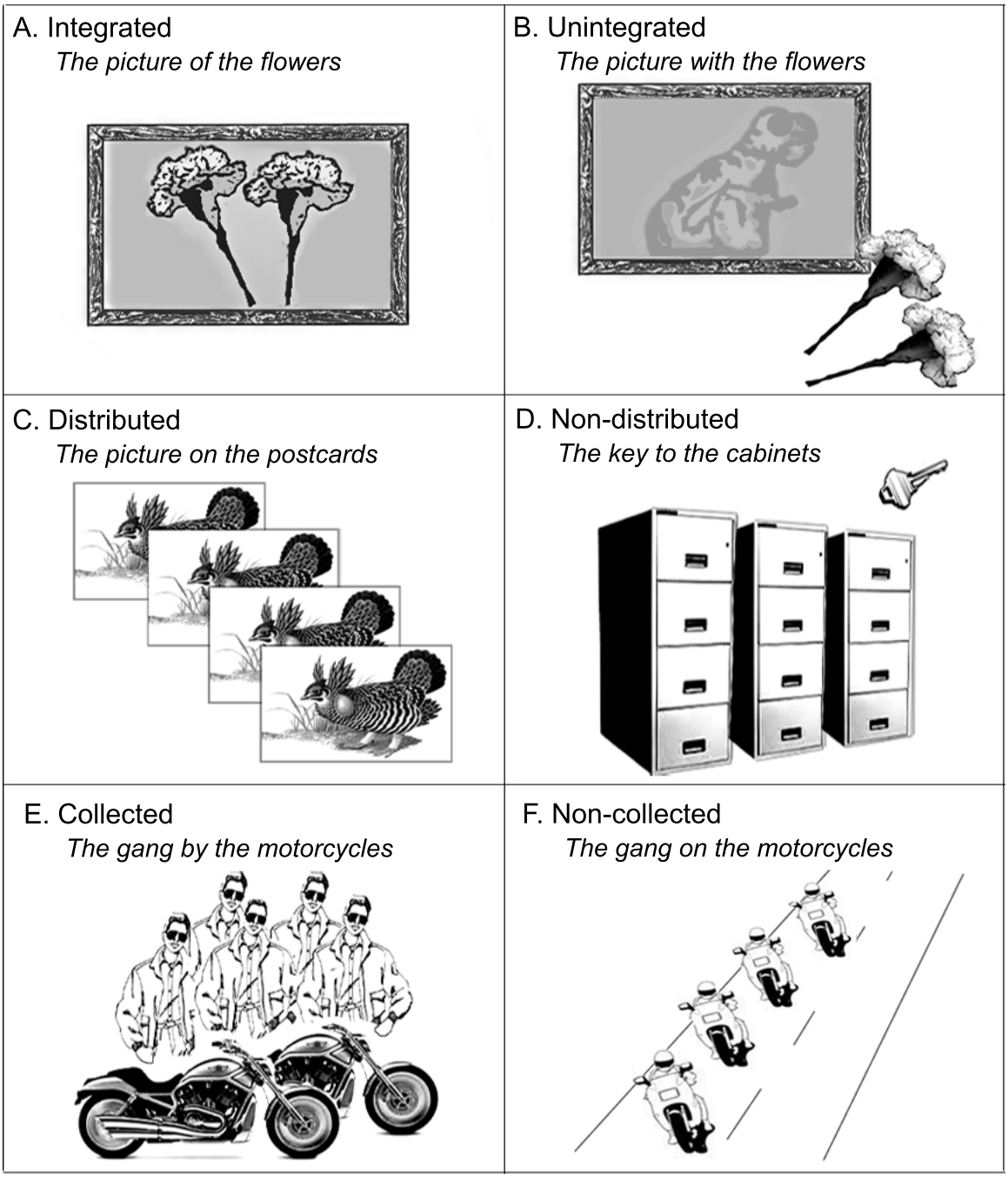


Figure 2. Trial time course (Standard trial on left; catch trial on right).

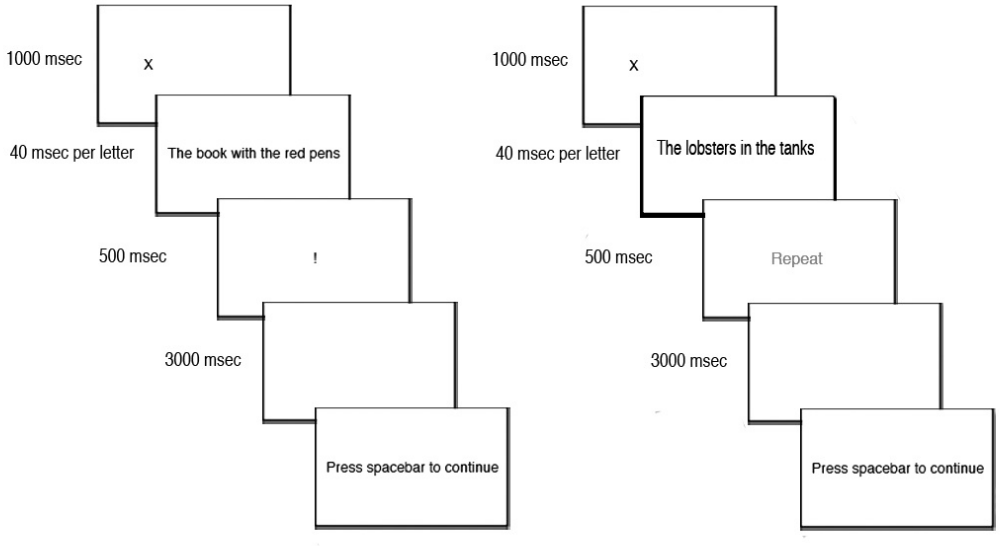


Figure 3. Experiment 1 mean voice onset time (in msec) by binned integration and local plurality. Error bars are 95% confidence interval on the difference between plural and singular local nouns, calculated from the interaction term in an ANOVA.

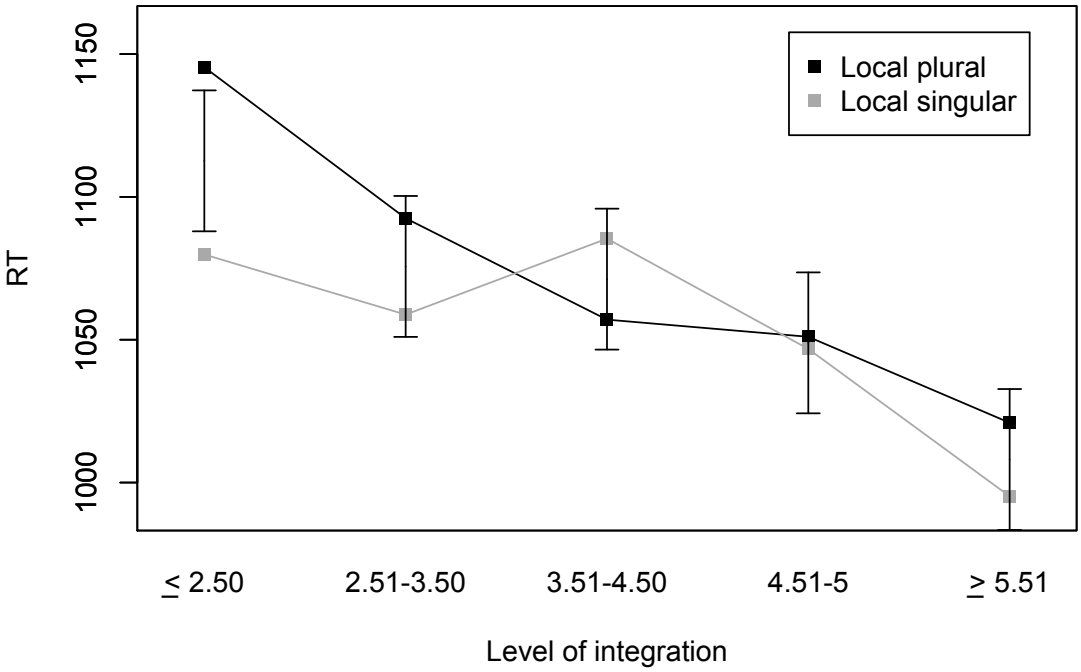


Figure 4. Experiment 1 proportion of plural responses by integration and local plurality. Error bars are 95% confidence intervals around differences between means, calculated from the interaction term in the empirical logit model.

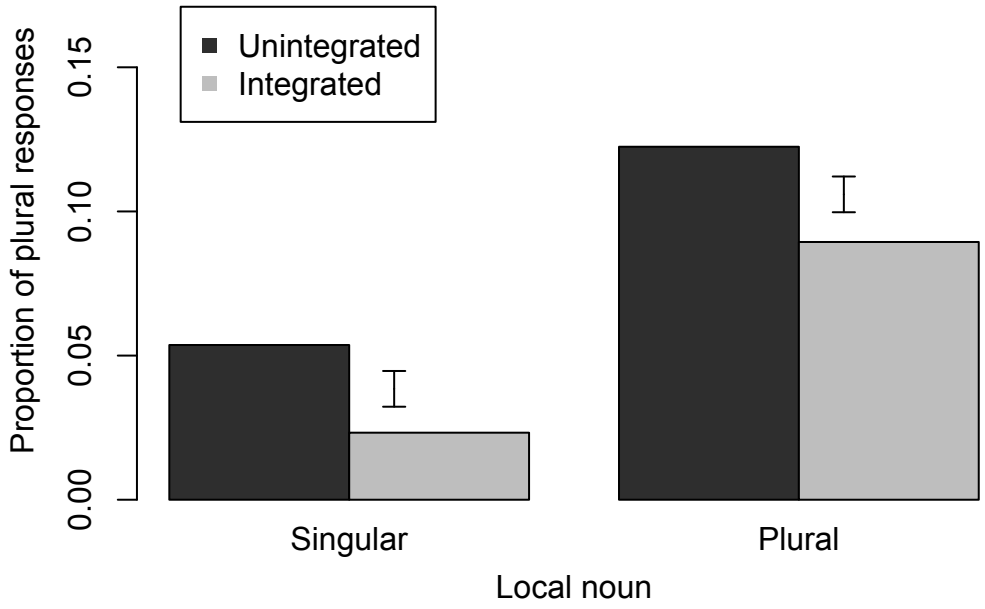


Figure 5. Experiment 2 proportions of plural response after unintegrated and integrated preambles with singular and plural local nouns. Error bars are 95% confidence intervals around differences between means, calculated from the interaction term in the empirical logit model.

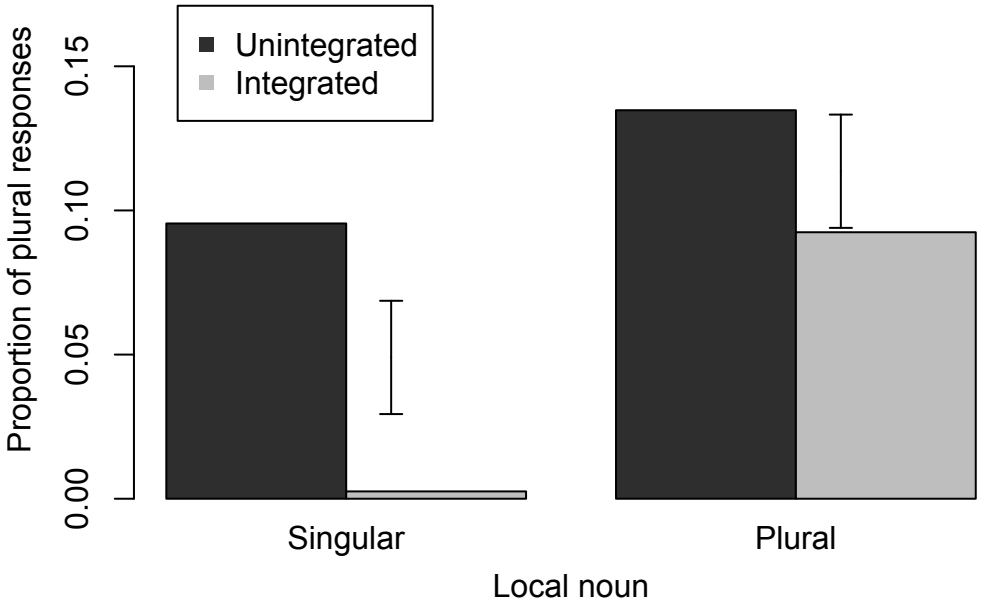


Figure 6. Experiment 2 plural responses by binned integration and local plurality. Error bars are 95% confidence intervals around differences between means, calculated from the interaction term in the empirical logit model.

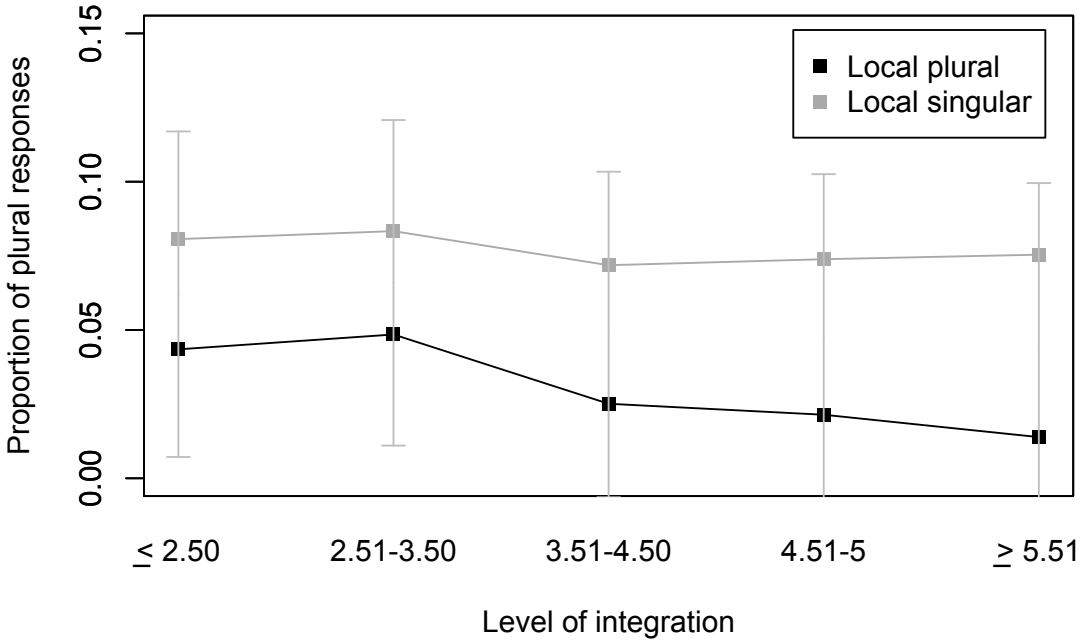
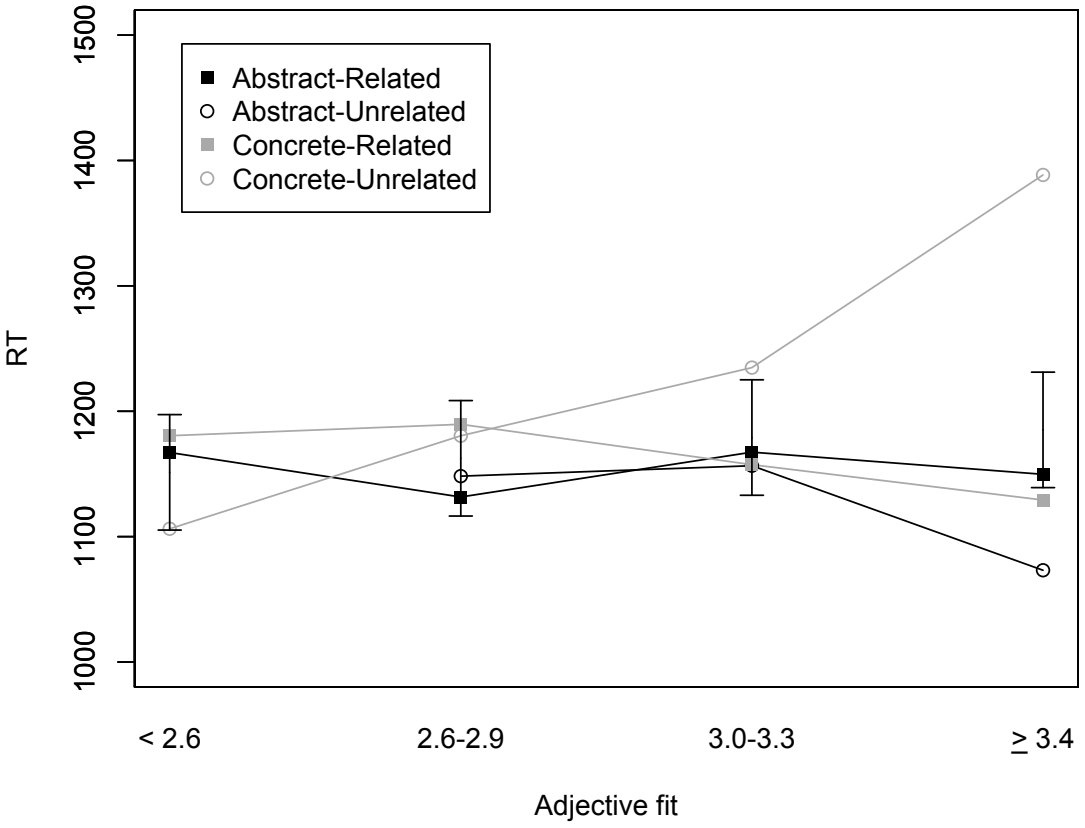


Figure 7. Experiment 3 response time by abstraction, relatedness, and adjective fit (lower is poorer fit). Error bars are 95% confidence intervals around differences between means, calculated from the interaction term in the multi-level model.



REFERENCES

Allum, P.H. & Wheeldon, L.R. (2007). Planning scope in spoken sentence production: The role of grammatical units. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 791-810.

Bates, D. M. (2005). Fitting linear mixed models in R: Using the lme4 package. R News: The Newsletter of the R Project, 5(1), 27–30.

Bock, K. & Eberhard, K. M. (1993). Meaning, sound and syntax in English number agreement. *Language and Cognitive Processes*, *8*(1), 57-99.

Bock, K., & Miller, C. A. (1991). Broken agreement, *Cognitive Psychology*, 23(1), 45-93.

Bock, K., Eberhard, K. M., & Cutting, J. C. (2004). Producing number agreement: How pronouns equal verbs, *Journal of Memory and Language*, 51(2), 251-278

Bock, K., Loebell, H., & Morey, R., (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review,* 99(1), 150-171

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(2), 257-271.

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, K. (2005). Making sense of syntax: Number agreement in sentence production. *Psychological Review*, 112, 531–559.

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

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.

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

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

Lorimor, H. (2001). Conjunctions and Grammatical Agreement. (Unpublished doctoral dissertation). University of Illinois, Urbana-Champaign.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. URL http://www.usf.edu/FreeAssociation/.

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

Pachella, R. G. (1974). The interpretation of reaction time in information processing research. In B. H. Kantowitz (Ed.), Human information processing: Tutorials in performance and cognition. (pp. 41-62). Hillsdale, NJ: Erlbaum.

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

Pearlmutter, N. J., Garnsey, S. M., & Bock, 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, URL http://www.R-project.org.

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.

Vigliocco, G., Butterworth, B., & Garrett, M. F. (1996). Subject–verb agreement in Spanish and English: Differences in the role of conceptual constraints. *Cognition,* 61, 261–298.

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

Wheeldon, L. R., & Monsell, S. (1994). Inhibition of Spoken Word Production by Priming a Semantic Competitor. *Journal of Memory and Language, 33*, 332-356.

APPENDIX A: 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?

APPENDIX B: RATINGS FOR PREAMBLE INTEGRATION (SOLOMON & PEARLMUTTER, 2004) AND PREAMBLE PLAUSIBILITY (PRESENT EXPERIMENTS AND SOLOMON & PEARLMUTTER).

Preamble	Integration rating	Plausibility ratings		
		Mean	Median	Solomon & Pearlmutter
The drawing of/with the flower	5.0/3.4	5.6/5.4	7/7	6.6/6.1
The picture of/with the gem	4.3/4.8	5.1/4.8	6/7	5.5/3.8
The sculpture of/with the key	4.0/3.8	5.4/4.7	7/5	3.5/2.5
The sketch of/with the bookcase	4.2/4.1	5.4/4.9	7/6	5.2/3.0
The painting of/with the costume	3.3/3.0	5.5/4.7	7/6	4.9/2.6
The statue of/with the bird	3.3/4.6	5.7/4.2	7/4	5.3/4.9
The tape of/with the record	5.3/5.6	5.4/5.7	7/7	4.1/3.1
The photo of/with the document	4.3/4.9	6.3/5.6	7/6	4.6/4.7
The xerox of/with the memo	4.6/4.6	5.2/5.5	6/6	5.9/4.4
The video of/with the puppet	3.2/3.6	5.0/6.4	6/7	4.0/3.7
The illustration of/with the map	5.3/4.4	5.7/5.5	7/6	4.9/5.1
The photocopy of/with the article	4.4/3.9	6.5/5.3	7/6	6.5/4.7
The reproduction of/with the antique	4.2/3.6	4.8/4.1	6/4	5.4/3.1
The fax of/with the blueprint	3.5/5.6	5.8/5.2	7/6	4.8/4.6
The telecast of/with the movie	5.2/3.7	5.8/5.8	7/7	5.0/3.8
The snapshot of/with the letter	3.2/4.1	5.5/5.3	6/6	3.9/4.5
The description of/with the CD	3.7/2.8	5.8/5.5	7/7	5.9/4.5
The slide of/with the magazine	3.2/3.8	4.6/4.8	5/6	3.4/2.9
The broadcast of/with the show	5.8/5.5	6.1/4.9	7/5	6.2/3.3
The draft of/with the report	5.2/4.8	6.7/5.6	7/7	6.4/5.2
The sonogram of/with the infant	5.3/5.3	6.2/4.7	7/5	5.8/5.0
The polaroid of/with the stamp	2.9/2.4	5.3/4.0	6/4	3.2/2.2
The postcard of/with the shoe	2.4/2.8	4.6/3.7	6/4	2.6/2.7
The portrait of/with the crown	3.5/4.8	6.1/5.3	7/6	4.7/4.5
The assistant for/with the inspector	4.3/3.6	5.3/4.9	7/6	5.4/4.5
The chauffeur for/with the actor	4.2/2.8	4.8/4.7	7/6	5.4/4.9
The apprentice for/with the tailor	5.1/4.5	5.4/4.9	7/6	5.0/5.5
The supporter of/with the evangelist	3.8/3.6	4.7/4.7	6/5	5.0/4.2
The translator of/with the ambassador	5.2/4.1	5.5/5.3	7/7	5.8/5.7
The secretary of/with the supervisor	5.4/3.9	4.9/5.1	7/7	6.0/6.1
The accountant for/with the millionaire	4.3/4.4	5.2/4.3	7/4	6.6/6.0
The nurse for/with the surgeon	4.8/5.0	5.5/5.5	7/6	5.2/6.2
The consultant of/with the producer	4.3/4.3	5.2/4.7	6/6	5.1/5.4
The advisor of/with the attorney	5.3/4.7	5.4/5.2	7/6	4.8/5.1
The servant for/with the diplomat	3.9/3.7	5.7/4.8	7/6	5.6/5.4
The manager of/with the band	4.7/5.1	5.5/5.4	7/6	5.9/5.9
The agent for/with the artist	4.0/3.5	5.2/4.4	6/5	5.6/5.0
The trainer for/with the athlete	5.5/5.3	5.2/5.8	6/7	6.7/6.2
The doctor of/with the patient	5.8/5.9	5.1/4.9	6/6	6.9/6.8

The tutor for/with the student	5.5/5.4	5.5/4.9	7/6	5.9/6.9
The coach of/with the gymnast	5.3/5.8	5.5/4.8	7/6	6.7/6.5
The photographer of/with the supermodel	4.4/4.9	5.2/5.5	7/7	6.8/6.3
The promoter for/with the DJ	4.5/4.6	5.8/5.2	7/7	5.5/5.6
The groundskeeper of/with the landowner	4.4/3.6	5.1/4.5	5/5	5.1/5.6
The book with the torn page/ the red pen	6.0/6.1	4.2/4.6	4/5	5.9/4.3
The shirt with the crazy pattern/ the dirty towel	5.2/4.8	4.4/3.6	6/3	5.5/3.6
The ring with the fake diamond/ the gold bracelet	6.1/5.9	4.5/4.9	5/6	4.3/6.0
The apple with the brown spot/ the fresh peach	4.0/4.4	3.1/4.6	2/6	5.5/5.4
The tie with the hideous stripe/ the cotton blazer	4.1/5.4	3.4/4.2	3/4	5.0/5.8
The watch with the missing hand/ the black wallet	6.0/5.1	3.4/3.8	3/4	4.2/3.9
The jacket with the faulty zipper/ the wet umbrella	6.3/5.9	3.5/2.8	4/2	5.5/5.0
The razor with the rusty blade/ the empty can	6.2/6.6	3.9/2.7	3/2	5.7/3.6
The key with the jagged edge/ the shiny coin	5.0/4.8	3.8/3.7	3/4	6.6/3.4
The bed with the creaky spring/ the tall bookcase	5.8/5.3	4.0/3.9	5/5	5.9/4.5
The phone with the missing button/ the broken toaster	5.3/5.3	3.5/2.4	3/1	4.5/1.6
The pillow with the nasty stain/ the flannel sheet	3.5/4.8	2.7/4.2	2/4	5.1/5.6
The lamp with the florescent bulb/ the antique portrait	6.1/5.7	4.4/3.9	6/4	6.3/4.2
The magazine with the colorful ad/ the telephone book	6.1/5.8	5.5/3.3	7/3	6.4/2.9
The sweater with the tiny hole/ the linen suit	4.5/4.7	4.0/3.6	4/2	5.6/3.2
The receipt with the blurry price/ the sealed package	5.8/5.7	4.9/3.8	6/4	5.5/4.3
The tree with the dead branch/ the small shrub	6.4/6.4	3.2/3.9	2/4	6.4/3.7
The pizza with the yummy topping/ the tasty beverage	6.0/6.4	4.6/4.2	7/4	6.6/5.6
The milk with the extra vitamin/ the blueberry muffin	5.6/4.9	3.6/5.5	3/6	5.8/5.9
The guitar with the loose string/ the loud drum	6.0/6.0	4.1/4.9	4/6	6.2/4.8
The blanket with the soft fringe/ the clean skirt	5.4/5.8	4.4/3.5	5/4	6.2/3.4
The glass with the lengthy crack/ the crystal bowl	4.6/4.9	3.6/3.6	3/4	5.6/4.9
The bike with the bent spoke/ the surfboard	5.9/6.3	4.1/3.3	5/2	6.2/3.2
The chair with the wobbly leg/ the old table	5.0/5.9	3.6/5.0	3/6	6.6/5.7
The report that described the traffic accident/ Megan described the accident	4.8/4.8	6.1/5.9	6/7	6.3/4.8
The confirmation that excited the employee/ Phil fired the employee	3.6/3.4	4.6/4.9	5/6	5.5/4.6
The opinion that insulted the executive/ Claire should help the executive	3.3/3.3	4.6/3.9	6/4	5.6/5.3
The fear that consumed the prisoner/ Dan freed the prisoner	4.3/4.8	4.1/4.0	4/4	5.3/4.5
The idea that struck the crazy scientist/ the school sued the scientist	4.4/4.4	5.8/4.2	6/5	5.4/5.9
The proof that changed the original ruling/ the lawyer approved the ruling	4.1/4.2	5.7/4.3	7/4	5.6/5.0
The rumor that plagued the gorgeous dancer/ Ted dated the dancer	2.5/3.0	4.5/4.4	5/5	5.5/5.9
The discovery that revealed the precious gem/ Kim bought the precious gem	3.3/3.7	4.8/3.9	5/4	5.6/5.4
The request that pleased the energetic chef/ Troy please the talented chef	2.9/3.3	4.6/3.4	5/3	5.4/4.8
The speculation that disgusted the elite committee/ Julie disgusted the committee	3.7/3.1	4.7/4.6	6/5	5.0/4.8

The announcement that cancelled the scheduled meeting/ the senator moved the meeting	3.4/4.4	5.0/5.9	6/6	6.0/6.4
The dream that haunted the wealthy stockbroker/ Claire would assist the stockbroker	2.8/2.1	4.2/4.7	5/6	5.1/5.3
The point that calmed the anxious witness/ Kevin calmed the witness	2.4/2.4	4.3/3.2	5/2	4.6/5.3
The threat that mentioned the deadly bomb/ the thief planted the bomb	4.9/5.1	4.7/3.7	5/2	4.8/4.1
The observation that confused the brilliant scientist/ Sarah confused the scientist	4.4/4.8	4.2/3.6	4/4	4.6/4.4
The verification that consoled the busy worker/ David consoled the worker	2.6/2.3	5.0/4.3	6/4	5.5/5.6
The misconception that puzzled the boy/ Mark smacked the boy	2.8/2.8	4.0/4.5	5/6	5.4/4.9
The suggestion that disappointed the cranky governor/ the aide disappointed the governor	2.9/2.8	4.4/4.4	5/4	4.1/4.5
The possibility that interested the recent applicant/ Richard might interest the applicant	3.9/2.9	3.8/4.5	3/5	5.5/4.0
The feeling that troubled the little girl/ Curt troubled the small girl	3.9/3.3	3.7/4.5	4/5	5.9/4.4
The finding that showed the important result/ Beth lied about the result	5.8/5.0	5.2/5.0	5/6	5.3/3.5
The myth that amused the small town/ Ann amused the town	2.4/2.3	4.7/3.8	5/4	5.8/4.5
The argument that bothered the enraged striker/ the boss bothered the striker	3.3/4.6	4.9/4.0	6/4	5.0/4.1
The remark that insulted the good friend/ Seth insulted the friend	2.5/2.9	5.0/3.8	5/3	5.5/4.3
The recommendation that thrilled the avid reader/ Ms. Drew thrill the reader	2.8/3.4	4.7/3.7	6/3	4.8/4.1
The information that convinced the rich investor/ Mae convinced the investor	3.8/4.3	5.6/4.0	6/4	5.4/4.9
The evidence that convicted the criminal/ Dan hated the criminal	5.4/5.4	5.8/4.0	7/5	6.0/4.4
The hope that reassured the actor/ Bob would meet the actor	3.6/2.4	4.0/3.5	4/3	4.4/3.8
The insinuation that alluded to the affair/ Jane had the steamy affair	4.3/3.6	4.5/5.0	5/7	5.0/4.6
The judgment that annoyed the defending attorney/ Brianna annoyed the attorney	4.4/5.1	4.7/3.5	5/3	5.3/4.4
The news that encouraged the baseball team/ the coach encouraged the team	3.4/2.4	4.7/4.6	6/5	6.0/5.5
The notice that hurt the caring parent/ the son hurt the parent	3.4/1.8	3.5/3.6	3/3	4.3/4.6

APPENDIX C: EXPERIMENT 1 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.

APPENDIX D: EXPERIMENT 2 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 E: PREAMBLES FROM EXPERIMENT 3

Condition	Stimulus	Normed rating		
		Abstraction	Relatedness	Adjective Fit
Unrelated-Concete	The ambulance and the light(s)	6.20	0.04	2.95
	The tongue and the tooth (teeth)	6.13	0.07	3.08
	His cabbage and salad(s)	6.45	0.02	3.23
	The tail and the animal(s)	5.87	0.03	2.90
	The cap and the gun(s)	6.43	0.02	2.93
	Their missile and plane(s)	6.06	0.01	3.54
	The dish and the cat(s)	6.49	0.01	2.63
	His donkey and the monkey(s)	6.07	0.01	2.98
	The hill and the valley(s)	6.34	0.05	2.31
	Her leg and her muscle(s)	6.02	0.01	3.00
	Her flask and his jar(s)	6.39	0.03	2.50
	His sword and her weapon(s)	5.82	0.02	2.96
	Her pouch and pipe(s)	6.70	0.01	2.70
	Their raft and beach(es)	6.03	0.01	3.00
	The rail and fence(s)	5.53	0.02	2.48
	The spine and the neck(s)	6.03	0.01	2.85
Unrelated-Abstract	The cause and the force(s)	2.91	0.01	2.95
	His chance and her option(s)	2.65	0.01	3.18
	The chaos and the headache(s)	2.38	0.01	2.78
	His compulsion and his tendency (tendencies)	2.16	0.01	2.70
	Her concept and his belief(s)	1.76	0.02	3.20
	The condition and the issue(s)	2.36	0.01	3.48
	Her curiosity and her interest(s)	2.09	0.06	2.96
	Their destiny and outcome(s)	2.34	0.01	3.20
	Their domain and rule(s)	2.95	0.01	3.08
	The fantasy and the wish(es)	2.35	0.01	2.80
	The honesty and the loyalty (loyalties)	2.13	0.02	3.34
	The hypothesis and the thought(s)	1.77	0.03	3.50
	The mischief and the mistake(s)	2.80	0.01	2.79
	His mood and his attitude(s)	1.68	0.02	3.03
	The norm and the standard(s)	2.39	0.04	2.85
	Her weakness and her fault(s)	2.58	0.02	2.73
Related-Concrete	The ambulance and the hospital(s)	6.90	0.17	3.38
	The tongue and the mouth(s)	5.80	0.38	2.65
	His cabbage and vegetable(s)	6.42	0.14	3.45
	The tail and the dog(s)	5.92	0.25	2.98
	The cap and the head(s)	6.38	0.10	2.46
	Their missile and the bomb(s)	6.36	0.07	2.95

	The dish and the plate(s)	6.26	0.32	2.43
	His donkey and the horse(s)	6.35	0.14	2.93
	The hill and the mountain(s)	6.14	0.43	2.39
	Her leg and her arm(s)	5.79	0.50	2.68
	Her flask and his bottle(s)	6.97	0.10	2.73
	His sword and her knife (knives)	5.67	0.30	2.88
	Her pouch and bag(s)	6.40	0.09	2.35
	Their raft and boat(s)	6.16	0.17	2.95
	The rail and train(s)	5.50	0.25	3.05
	The spine and the bone(s)	5.99	0.08	2.90
Related-Abstract	The cause and the effect(s)	2.56	0.36	3.13
	His chance and her risk(s)	2.71	0.10	3.05
	The chaos and the confusion(s)	2.52	0.15	2.78
	His compulsion and his obsession(s)	2.10	0.21	2.86
	Her concept and his idea(s)	2.42	0.42	3.48
	The condition and the situation(s)	2.74	0.06	3.23
	Her curiosity and her wondering(s)	1.96	0.06	2.94
	Their destiny and fate(s)	2.32	0.21	3.05
	Their domain and range(s)	2.97	0.17	2.74
	The fantasy and the dream(s)	2.53	0.50	2.75
	The honesty and the truth(s)	2.71	0.45	3.36
	The hypothesis and the theory (theories)	2.08	0.27	3.50
	The mischief and the trouble(s)	2.58	0.30	2.95
	His mood and his emotion(s)	1.96	0.07	2.88
	The norm and the average(s)	2.27	0.17	2.89
	Her weakness and her strength(s)	2.60	0.25	3.15