

## Constituent Structure and Linear Order in Language Production: Evidence From Subject–Verb Agreement

Todd R. Haskell  
University of Southern California

Maryellen C. MacDonald  
University of Wisconsin—Madison

A number of studies have shown that structural factors play a much larger role than the linear order of words during the production of grammatical agreement. These findings have been used as evidence for a stage in the production process at which hierarchical relations between constituents have been established (a necessary precursor to agreement), but before the final linear order of words is determined. The current article combines evidence from off-line ratings, online production studies, and a corpus analysis in support of the view that linear order effects do exist. These findings have implications both for theories of agreement production and language production more generally.

*Keywords:* subject–verb agreement, syntactic structure, word order, language production

Theories of language production differ on many details, but they generally agree that the production mechanism consists of a number of levels or stages, with one or more distinct processes occurring at each stage. For example, it is widely believed that the process of producing a word involves two distinct levels, one at which underlying conceptual features are used to select a particular word (the *lemma* level), and another at which the phonological form of that word is retrieved (the *lexeme* or *word form* level; for a recent incarnation of this view, see Levelt, Roelofs, & Meyer, 1999).

Although the notion of a series of levels or stages is typically applied to the production of complete sentences as well as individual words, in the sentence case the situation is more complex, and correspondingly less is known. Most theories of production assume that, in the course of producing a sentence, speakers generate a representation of its constituent structure, encoding the relationships between constituents as well as their order. There are two views on how this comes about. On one view (hereafter the *two-stage approach*), there is an initial stage during which the hierarchical relations between constituents are established, followed by a second stage during which the constituents are placed in their final order (Hartsuiker & Westenberg, 2000; Vigliocco & Nicol, 1998). The alternative (hereafter the *single-stage approach*) is that hierarchical relations and linear order are determined at the same time (Pickering, Branigan, & McLean, 2002).

To understand the distinction between the two approaches, consider how a speaker would produce a phrase such as *chips and a drink*. On the two-stage view, the first step is to construct a representation encoding the fact that there is some type of node that dominates both *chips* and *a drink*. It is important to note that this representation would be exactly the same whether the speaker eventually produces *chips and a drink* or *a drink and chips*. Only at the second stage is the linear order of the two constituents determined. The single-stage view, by contrast, posits a single production stage during which constituent structure and linear order are concurrently developed. On this view, there is never a point at which a speaker's representation of *chips and a drink* and *a drink and chips* would be the same; rather these two phrases have two distinct syntactic representations.

The choice between these alternatives has significant implications for theories of language production. For example, Vigliocco and Nicol (1998) suggest that having two stages helps resolve an important problem in production, namely, the fact that words may not be retrieved from the mental lexicon in the same order as they must be produced. Pickering et al. (2002) have argued that if there is only one stage, then information about different options for ordering the arguments of verbs must be explicitly represented in the lexical representations of those verbs.

Two phenomena that have been used in attempts to distinguish between the single-stage and two-stage approaches are structural priming and agreement errors. Structural priming phenomena are most commonly observed in production studies in which the speaker has a choice between two alternative syntactic structures in conveying a message. For example, a situation involving a giver, a gift, and a recipient may be described with a prepositional dative (*The author gave a book to the library*) or a double-object dative (*The author gave the library a book*). The key manipulation concerns the speaker's utterances in preceding trials of the experiment: If the speaker has recently produced a prepositional dative sentence (e.g., *The man read a story to the boy*), she is more likely to describe the giving situation using a prepositional dative than if she has just produced a double-object dative (e.g., *The man read the boy a story*; J. K. Bock, 1986).

---

Todd R. Haskell, Department of Linguistics, University of Southern California; Maryellen C. MacDonald, Department of Psychology, University of Wisconsin—Madison.

This work was partially supported by a University of Southern California, Los Angeles Dissertation Fellowship to Todd R. Haskell and by Grant MH64445 from the National Institutes of Health. Portions of this work were presented at the 15th Annual CUNY Conference on Human Sentence Processing. We thank Zenzi Griffin for helpful comments on an earlier version of this article.

Correspondence concerning this article should be addressed to Todd R. Haskell, Hedco Neurosciences Building, 3641 Watt Way, University of Southern California, Los Angeles, Los Angeles, CA 90089-2520. E-mail: thaskell@usc.edu

Note that, as with conjunctions, in some verb phrases it is possible to change the order of constituents without altering hierarchical relations. This can occur with so-called shifted prepositional datives such as *The author gave to the library a famous and very rare manuscript*. Shifted constructions and the alternative nonshifted version (*The author gave a famous and very rare manuscript to the library*) arguably share the same hierarchical relations. Pickering et al. (2002) asked whether there is a stage of sentence production at which this hierarchical information is represented, independent of information about the linear ordering of the constituents. If this is the case, then both shifted and unshifted prepositional datives should prime unshifted prepositional datives. In contrast, if shifted and unshifted prepositional datives are represented differently at all levels, then only unshifted prepositional datives should prime unshifted prepositional datives (priming of shifted constructions was not investigated, as they are very rare with short noun phrases of the sort used in the Pickering et al. experiments). The results followed the latter pattern. Pickering et al. concluded that the complete constituent structure, ordering included, is constructed in a single stage.

On the face of it, the findings of Pickering et al. (2002) would seem to conflict with data from the production of grammatical agreement, which seem to support the two-stage view. Speakers (and writers) make agreement errors following phrases like *the key to the cabinets* fairly frequently, producing utterances such as *the key to the cabinets were missing*. One potential cause of such an error is the proximity of the verb to the interfering plural noun *cabinets* (commonly termed the *local* noun). To investigate whether linear proximity was the cause of these agreement errors, Vigliocco and Nicol (1998) contrasted a condition in which speakers were asked to produce declarative sentences (*The helicopter for the flights was safe*) with a condition in which speakers were asked to produce questions (*Was the helicopter for the flights safe?*). In the declarative condition, the local noun immediately preceded the verb, while in the question condition, it was several words distant. If proximity of the verb to the local noun was causing agreement errors, then such errors should be more prevalent in the declarative condition than in the question condition. Vigliocco and Nicol found no evidence for a difference in the rate of agreement errors across the two conditions.

More recently, Franck, Vigliocco, and Nicol (2002) investigated the rate of agreement errors with three-noun phrases like *the threat(s) to the president(s) of the company(ies)*. If linear proximity to the verb is driving the occurrence of attraction errors, then one would expect more attraction errors with phrases like *the threat to the president of the companies* than with phrases like *the threat to the presidents of the company* because the mismatching noun is linearly closer to the verb in the first sentence. In contrast, if structural or hierarchical proximity to the verb is more important, one would expect the opposite pattern, because the mismatching noun is structurally closer to the verb in the second sentence. The results were consistent with the predictions of the structural account, supporting the claim that linear proximity does not affect agreement computation. Based on these findings, Vigliocco and Nicol (1998) and Franck et al. argued that agreement computation takes place at a stage when hierarchical relations have been established, but before the linear order of words has been determined. This is precisely the two-stage view that Pickering et al. (2002) argued against.

In short, two proposals have been put forward regarding the relative timing of determining hierarchical relations and determining linear order during the course of language production. The primary goal of this article is to help resolve the deadlock between these positions by taking a closer look at the agreement phenomena claimed to support the two-stage view. We present new agreement data which are problematic for the particular two-stage model that Vigliocco and colleagues have proposed but that are consistent with the single-stage model. Although our findings cannot rule out all possible two-stage models, to the extent that agreement data have been used to argue against a single-stage model and for a two-stage model, we believe our findings shift the balance of the evidence in favor of the single-stage model.

Before describing our experiments, we first turn to the existing agreement research and consider more closely the relationship between subject–verb agreement and accounts of production stages.

### Agreement and Linear Order

Computing agreement is a complex process; it is likely that different aspects of agreement are computed at different points in the production process (see K. Bock, Eberhard, Cutting, Meyer, & Schriefers, 2001, for an example of a two-step model). However, for present purposes we will focus on the more or less grammatical aspects of agreement that seem to be involved in the findings of Franck et al. (2002) and Vigliocco and Nicol (1998).

There is abundant evidence that agreement is sensitive to structural factors, the results of Franck et al. (2002) regarding *the threat to the president of the companies* versus *the threat to the presidents of the company* being a prime example (see also K. Bock & Cutting, 1992; K. Bock & Miller, 1991). Thus, it would seem that agreement must occur either at the same time or after hierarchical structure is determined.

At the same time, there is evidence suggesting that agreement cannot occur strictly after linearization. This evidence comes from speech errors such as the following (Garrett, 1980):

1. It probably *get outs* a little (for *gets out*).
2. It *deads end* into the. . . (for *dead ends*).
3. I'd *forgot abouten* that (for *forgotten about*).

These errors all involve the misplacement of a verbal affix within the sentence. Presumably such errors occur during the process of placing constituents in their proper order within the string, that is, during linearization. Note, however, that agreement must be computed before the error happens because agreement determines what the suffix will be. Vigliocco and Nicol (1998) make a similar argument based on slightly different errors. Thus, it would seem that agreement must occur either at the same time as or before linearization.

For the single-stage approach, this precisely situates where agreement occurs: during the (single) stage when hierarchical structure is built and linearization occurs. For the two-stage approach, there are three possibilities. First, agreement might occur concurrently with the building of hierarchical structure. Second, it might occur immediately after hierarchical structure but before linearization (this is what Franck et al., 2002, and Vigliocco &

Nicol, 1998, have suggested). With respect to effects of linear order, the first two possibilities make the same predictions and so will collectively be referred to as the *two-stage agreement-early* view. Finally, agreement might occur concurrently with linearization, hereafter the *two-stage agreement-late* view.

How do the extant data relate to these different possibilities? One might expect under the single-stage or two-stage agreement-late views that linear order would have at least a small effect on agreement. In that case, the failure of Franck et al. (2002) and Vigliocco and Nicol (1998) to find such effects would count against these views. However, in interpreting the findings of Franck et al. and Vigliocco and Nicol, it is important to distinguish between two different claims. The first claim is that structural relations seem far more important than linear order in determining agreement. This claim would seem to be well-supported by the results from Franck et al. and Vigliocco and Nicol discussed above (see also K. Bock & Cutting, 1992). The second, stronger claim is that linear proximity does not have any observable effect on agreement because linear order information is not yet available when agreement is being computed; this is the essence of the two-stage agreement-early view. It is logically possible for the first claim to be true while the second is false: The fact that linear order information is available at the time agreement is computed does not require that it be the dominant factor determining agreement.

That being said, if linear order has any influence on the agreement process, it is surprising that none of the previous studies have found order effects. However, none of the previous studies were designed to do so. For example, the experiments conducted by Franck et al. (2002) deliberately pit linear order and structural factors against each other. Thus, they demonstrated that structural proximity of the local noun to the verb is more important than linear proximity, but they do not tell us whether linear proximity plays any role at all.

Initially, it would seem that the experiment conducted by Vigliocco and Nicol (1998), which contrasted sentences such as *The helicopter for the flights was safe* and *Was the helicopter for the flights safe?*, was a more sensitive test for order effects. However, there are reasons to believe that any effects of linear order might have been masked by the much stronger structural factors at play. For example, early studies contrasting regular local nouns (*the trap for the rats*) versus irregular local nouns (*the trap for the mice*) did not observe any difference (K. Bock & Eberhard, 1993), but Haskell and MacDonald (2003) did observe such effects when they took steps to attenuate the strength of syntactic factors and allow the local noun to play a larger role. This is a particular concern with the Vigliocco and Nicol experiment because the overall rate of agreement errors was quite low relative to other studies (around 5%), suggesting that the local noun was having a minimal impact on agreement.

If strong structural factors tend to mask weak linear order effects, this raises the question of how to attenuate structural effects in an experiment setting in order to observe whether linear order effects then can be detected. The ideal situation would perhaps be to use a syntactic structure containing two nouns with the same structural proximity to the verb but with one linearly closer than the other. One such structure in English is disjunctive noun phrases such as *the shirt or the socks*.

There are some data on this structure from a survey conducted by Peterson (1986), who examined agreement with disjunctive

subjects. He collected verb preference ratings (singular vs. plural) for questions, in which the verb precedes the subject (*Have/has the president or the senators read the documents yet?*) and declarative sentences, in which the verb follows the subject (*Either this tree or those shrubs has/have to be cut down*). The results of Peterson's survey suggested that at least some raters were using a proximity strategy, whereby the verb agrees with the nearest noun. However, for a variety of reasons the data are not conclusive. Peterson's survey used only one or two items per condition; thus it is possible that the findings reflect idiosyncratic properties of his materials, rather than a general tendency in the language. In addition, his study involved off-line ratings, and it is possible that these contemplative, metalinguistic judgments may not accurately reflect speakers' behavior in the course of normal production.

To briefly summarize, the extant evidence indicates that linear proximity to the verb has little or no effect on the rate of agreement errors with noun phrases like *the key to the cabinets* or *the threat to the presidents of the company*. This evidence has been used to argue that hierarchical structure information but not linear order information is available at the time agreement is computed. However, there are several reasons to believe that this sort of phrase does not provide the maximally sensitive test for linear order effects and that a different type of phrase—disjunctions—would provide a more sensitive test. It is this more sensitive test that the current article aims to provide.

The remainder of the article is structured as follows. First, we present a series of three experiments exploring whether linear order influences the production of agreement with English disjunctive noun phrases, thereby providing some additional evidence relevant to the debate between the single-stage and two-stage views. The goal of Experiment 1 was to replicate Peterson's (1986) ratings data with a controlled set of stimuli. Experiments 2 and 3 involved an online production task to investigate agreement with disjunctions in sentences with subject-before-verb order (Experiment 2) and sentences with verb-before-subject order (Experiment 3). To anticipate the results, we found strong evidence that linear order plays an important role in agreement with disjunctions. Next, we explore the possible bases of this effect. Although our conclusions are more tentative here, we present a corpus analysis supporting the claim that this effect reflects the influence of distributional information on language production. In the general discussion, we consider the implications of our findings for theories of language production.

## Experiment 1

The survey data reported in Peterson (1986) are consistent with the claim that, given disjunctions like *the boy or the girls*, English speakers tend to prefer a verb form that agrees with the nearer of the two nouns. However, for the reasons discussed above, these data are far from conclusive. In order to more firmly establish the reality of the proximity effect, an off-line survey was conducted that replicated a portion of Peterson's survey with a more controlled set of stimuli.

## Method

*Participants.* Twenty native speakers of English from the University of Southern California community participated in the experiment for monetary compensation or credit in an introductory psychology course.

**Materials.** Forty items were constructed resembling the item presented in Table 1. Each item consisted of a question containing a main clause and a subordinate clause; the subordinate clause contained a disjunctive noun phrase with one singular and one plural noun, and a form of *be*. Questions were used because disjunctions are normally used in a request for information (i.e., “which one?”). Placing the disjunction in a subordinate clause preserved the normal subject-before-verb ordering of declarative sentences. There were two versions of each item. In the singular–plural (SP) version, the singular noun preceded the plural noun in the disjunction; in the plural–singular (PS) version the ordering of the two nouns was reversed. The complete set of items is presented in Appendix A. An additional 24 items were generated to serve as fillers. These items involved other types of noun phrases, for example, *the truck behind the warehouse, the blouse and the skirt*.

Two stimulus lists containing all 40 items were generated: For each item, one list contained the item in the SP version and the other contained the item in the PS version. Each list contained equal numbers of items in the SP and PS versions. Both lists contained all 24 fillers. The order of items was randomized separately for each list. Equal numbers of participants saw each list.

**Procedure.** The lists were presented to participants in a printed format, with each item followed by a 7-point scale ranging from  $-3$  (*left verb better*) to  $3$  (*right verb better*). Participants were instructed to read each item and to consider the relative acceptability of the two presented verb forms. If only the verb on the left was acceptable, they were to circle the  $-3$ . If only the verb on the right was acceptable, they were instructed to circle the  $3$ . If both (or neither) verb form was deemed acceptable, they were instructed to circle  $0$ . The remaining numbers were to be used to indicate graded degrees of acceptability, that is, both forms could be used, but one sounded somewhat better than the other. For half the participants, the singular verb was presented on the left and the plural on the right. For the other half of the participants, the order was reversed. The survey took approximately 5 min to complete.

## Results and Discussion

For purposes of the analyses, the raw data were recoded so that a rating of  $3$  represented a preference for the proximate noun, and a rating of  $-3$  represented a preference for the more distant noun. Overall, participants showed a preference to agree with the proximate noun ( $M = 0.89$ ,  $SE = 0.23$ ). When compared against a chance level of  $0$ , this proximity preference was significant,  $t_1(19) = 3.79$ ,  $p < .001$ ;  $t_2(39) = 17.8$ ,  $p < .001$ . However, the preference was much stronger in the SP condition ( $M = 1.96$ ,  $SE = 0.23$ ) than in the PS condition ( $M = -0.18$ ,  $SE = 0.39$ ), where there was even a slight numerical preference for agreement with the more distant noun. This resulted in a significant effect of noun order,  $t_1(19) = 4.95$ ,  $p < .001$ ;  $t_2(39) = 9.55$ ,  $p < .001$ .

Some insight into participants’ preferences in the PS condition can be gained by considering how the responses were distributed across the rating scale. Of all responses, 47% were 3s (either

positive or negative). An additional 29% were 2s (either positive or negative). Thus, some of the responses were consistent with proximity, and some were the exact opposite of what was expected. We return to this intriguing result in the context of the corpus analysis presented later in the article.

In summary, our results are broadly consistent with the findings of Peterson (1986), though they suggest that in some cases (i.e., the SP condition) English speakers almost always make use of proximity in determining appropriate verb number agreement for disjunctions, whereas in other cases (i.e., the PS condition) proximity plays a much smaller role. However, performance on an off-line task like verb preference ratings is not necessarily indicative of how speakers compute agreement online during language production. Thus, in Experiments 2 and 3 we sought to replicate the proximity effects observed in the current study with an online task.

## Experiment 2

The initial pilot testing for this study used stimuli like those used in the previous study, combined with the fragment completion task of K. Bock and Miller (1991), in which speakers are given a sentence fragment and must use it as the beginning of a complete sentence. This pilot work revealed that speakers found it extremely difficult to correctly reproduce the fragment that was provided to them. In particular, participants frequently changed the number of one of the nouns. This is not surprising, because the fragment completion task requires participants to comprehend and memorize a complex noun phrase, including number on multiple nouns, with minimal semantic context. The nouns in disjunctions (e.g., *the boy or the girls*) likely have a less structured semantic relationship between them than the type of materials used in most previous fragment completion experiments (e.g., *the key to the cabinets*), making the memory burden relatively higher for disjunctions. We therefore attempted to develop a task that provided number information in a more meaningful way, thus yielding a richer message-level representation. In this task, participants were presented with two pictures. Each picture depicted either a single object, or a pair of identical objects. The task required the participant to identify which of the two pictures had a certain property by asking the experimenter. This provided a context in which it was natural to use a disjunction while eliminating the memory requirements of the fragment completion task.

## Method

**Participants.** Twenty students from the University of Southern California participated in exchange for credit in an introductory psychology class. All participants were native speakers of English.

**Materials.** The stimuli consisted of 40 public-domain photographs of everyday objects and animals obtained from various sites on the Internet. Each photograph was converted to a grayscale image and scaled so that all images were approximately the same size. Twenty of the photographs were randomly selected for use in the experimental trials; the remaining photographs were used in the filler trials.

**Lists and procedure.** The intent of the task was to elicit utterances containing disjunctive noun phrases from the participants. To achieve this, a card sorting task was designed in which the participant had to ask the experimenter a question in order to arrange cards in correct locations on the table. First, five sheets of paper were placed on the table in front of the participant. On each sheet of paper was a color name. This is illustrated in Figure 1.

Table 1  
Example Stimulus Item for Experiment 1

Condition	Item
SP	Can you ask Brenda if the boy or the girls is/are going to go first?
PS	Can you ask Brenda if the girls or the boy is/are going to go first?

Note. SP = singular–plural version; PS = plural–singular version.

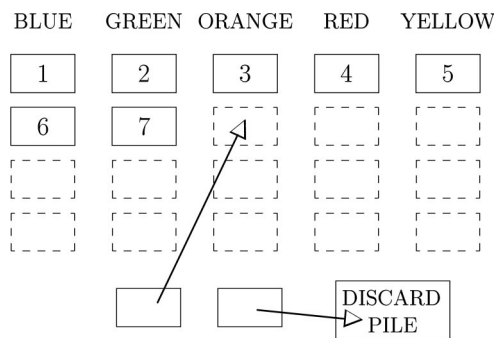


Figure 1. Schematic of the card sorting task used for Experiment 2, illustrating the sequence in which target cards were placed on the board.

Over the course of the experiment, participants were presented with 20 pairs of cards placed side by side. Each card had either a single grayscale image or two identical grayscale images on it. An example pair is shown in Figure 2. For each pair of cards, participants were instructed to determine which one was the *target card* (defined below). The target card was then placed below the appropriate sheet of paper, and the other card was discarded. Cards were placed below the sheets in sequence, moving from left to right across the five colors and then returning to the first sheet (see Figure 1).

The target card was defined as follows. For each pair of cards, there was a target color, determined by the color name at the top of the sheet where the target card was to be placed. After handing a pair of cards to the participant, the experimenter took a pair of cards for himself from a separate pile of cards, hidden from the participant. These cards were the same as those given to the participant, except they were colored, and one of the experimenter's cards was the target color. The task of the participant was to determine which card this was; to do this, participants were instructed to ask the experimenter a question beginning with "Can you tell me whether," followed by a description of what was on each card, followed by the target color. In the course of producing this question, the participant would have to produce either a singular or plural verb, for example, "Can you tell me whether the horses or the clock is/are red?" The experimenter would then refer to his own cards and answer by using a simple noun phrase (e.g., "the horses"). Upon receiving the answer, the participant would then place the target card in the appropriate spot on the table and discard the other card. The next trial would then commence.

The factors of interest were manipulated by independently varying the number of objects or animals portrayed on the left and right cards. Half of the trials were filler trials. For these trials, two picture cards both depicting either one or two entities were used, with each type occurring equally often. The other trials were experimental trials. For these trials, half the time the card on the left would portray one entity, and the card on the right portrayed two entities, whereas the remainder of the time the reverse was true. Although participants were not explicitly instructed to describe the

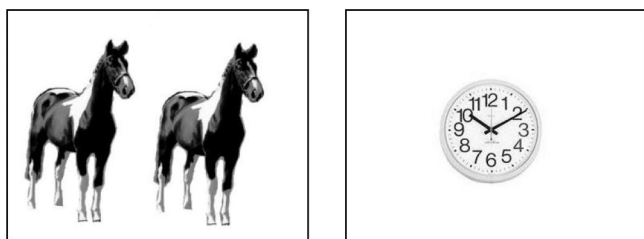


Figure 2. Example card pair for Experiment 2.

cards from left to right, pilot testing revealed that there was a strong natural tendency to do so.

To ensure that responding did not reflect characteristics of particular images or card pairs, five lists of card pairs were generated. The pairings of the images and the location of the target card (left or right) were randomized separately. Four participants were presented with each list.

**Scoring.** Responses were considered "scorable" if the participant (a) produced a question beginning with "Can you tell me whether" or any similar phrase, (b) described the entities on both the cards in a left-to-right order by using a singular noun for one entity and a plural noun for two entities, and (c) produced a single main verb.

Scorable responses were categorized according to the number of the first and second nouns and the number of the verb. Noun number was coded as SP or PS. Verb number was scored as singular or plural (no unmarked verbs were produced). If the participant produced one verb and then self-corrected ("is—I mean are"), the first verb produced was scored.

### Results and Discussion

Nonscorable responses (10%) were discarded from all analyses. In general, the nonscorable responses occurred when only one card was described ("Can you tell me whether the clock is red?"), when a grammatically singular noun was used to describe two entities ("Can you tell me if the clock or the pair of sofas is red?"), or when separate verbs were used for each card ("Can you please tell me if the clock is or the sofas are red?").

The dependent variable in the analyses was the rate of plural-marked verbs, that is, plural verbs/(singular verbs + plural verbs). To determine whether noun order (SP vs. PS) had a significant effect on the rate of plural verbs, parallel *t* tests were conducted that used participants and items as the random variables. For the items analysis, one item was discarded because it did not yield any scorable responses. Order was treated as a within-groups variable in the analysis by participants, and a between-groups variable in the analysis by items. For the items analysis, each pairing of cards was treated as a separate item.

The proportion of plural-marked verbs in each condition is presented in Figure 3. If verb number is influenced by proximity,

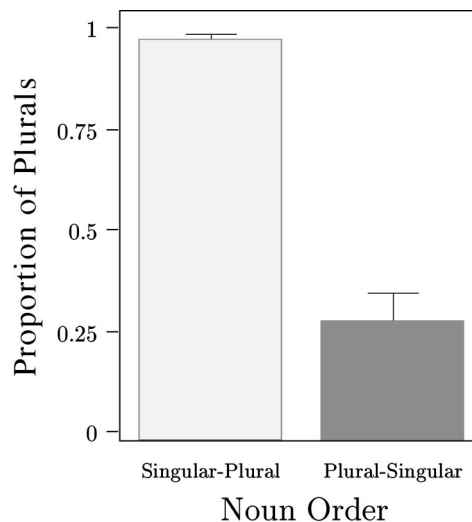


Figure 3. Mean proportion of plural verbs by condition in Experiment 2.

there should be more plural verbs in the SP condition than in the PS condition. Indeed, plural verbs were overwhelmingly more common in the SP condition, as reflected in a significant main effect of order,  $t_1(19) = 9.7, p < .001$ ;  $t_2(47) = 12.1, p < .001$ . For purposes of comparison, the rate of plural verbs was 10% in the SS condition and 100% in the PP condition.

It is interesting to note that, as in Experiment 1, the tendency toward proximity was stronger in the SP condition (98% agreement with the proximate verb) than in the PS condition (72% agreement with the proximate verb),  $t_1(19) = 3.72, p < .001$ ;  $t_2(47) = 4.48, p < .001$ . Thus, whatever causes this asymmetry, it is apparently not an artifact of the off-line ratings task. This finding will be addressed more fully in the context of the corpus analysis later in the article.

The occasional use of plural verbs in the SS condition is somewhat surprising. We cannot offer a definitive explanation for these responses, but we can offer a speculative possibility. Note that a question like "Can you tell me whether the horses or the clock is/are red?" can be interpreted in two different ways. One reading can be paraphrased as, "Is it the case that either the horses or the clock is/are red?" (i.e., the *yes-no* reading). The other reading roughly corresponds to, "Either the horses or the clock is/are red, and I want to know which one" (the *which-alternative* reading). Peterson (1986) found that plural verbs were often accepted for SS items with a *yes-no* reading. Thus, there appears to be a certain amount of semantically conditioned variability with these items, and although the pragmatics of our task encouraged a *which-alternative* reading, this variability may have occasionally led to the production of a plural verb.

During debriefing, participants were asked whether they had any ideas what the experiment was investigating. Responses from 5 of the participants made reference to choosing between *is* and *are*. When informed of the purpose of the experiment, a few additional participants reported encountering some difficulty computing agreement. Thus, it is possible that at least some participants used conscious, experiment-specific strategies in their choice of verb number. If this were true, the results of the current study might not reflect the normal mechanisms used in computing agreement. To test for this possibility, a separate analysis of variance (ANOVA) was conducted to compare the pattern of results for those participants who thought the experiment was related to the choice between *is* and *are* (the *aware* group) and those who did not suspect that the experiment was investigating agreement (the *naive* group). A proximity effect was evident for both groups; the rate of plural verbs was 92% (SP) versus 65% (PS) for the aware group, and 100% (SP) versus 15% (PS) in the naive group. It is interesting to note that it was actually significantly weaker in the aware group, as reflected by a Group  $\times$  Noun Order interaction (SP vs. PS),  $F_1(1, 18) = 30.7, p < .001$ .<sup>1</sup> Thus, the proximity effect observed in the current experiment is probably not the result of conscious attention to agreement processing.

The results of Experiment 2 provide strong evidence for a proximity effect in the production of subject-verb agreement. In producing agreement with disjunctive noun phrases, speakers showed an overwhelming tendency to inflect the verb to agree with the nearer noun, whether that noun was singular or plural. However, in Experiment 2, the nearer noun was always the second noun. To provide converging evidence that speakers are inflecting the verb to agree with the nearer noun (rather than always the

second noun), in Experiment 3 participants were asked to form simple questions, for example, "Is/are the horses or the clock red?" Thus, in this task speakers utter sentences in which the normal subject-before-verb ordering is reversed, to place the verb before the subject. If the effect observed in Experiment 2 was truly a proximity effect, then reversing the order of subject and verb should reverse the effect, such that agreement is predominantly with the first noun in the disjunction.

### Experiment 3

#### Method

*Participants.* The participants were 20 students from the University of Southern California, who participated in exchange for credit in an introductory psychology class. All participants were native speakers of English. Data from 3 additional participants were discarded because they failed to follow the instructions or failed to produce scorable responses in all four cells of the design.

*Materials and procedure.* The five stimulus lists from Experiment 2 were also used in the current experiment; again, 4 participants were presented with each list. The procedure was the same as in Experiment 2, with the exception that rather than being requested to produce questions of the form "Can you tell me whether the horses or the clock is/are red?," participants were asked to produce simple questions (e.g., "Is/are the horses or the clock red?").

*Scoring.* Scoring was the same as in Experiment 2.

#### Results and Discussion

Nonscorable responses (22%) were discarded from all analyses. This percentage was much higher than in Experiment 2. The increase was almost entirely due to a sharp rise in circumlocutions (e.g., "Are the horses or is the clock red?"), which constituted 8% of all responses.

Analyses were conducted in the same fashion as for Experiment 2. The proportion of plural-marked verbs in each condition is presented in Figure 4. If verb number is influenced by proximity, there should be more plural verbs in the PS condition than in the SP condition. Plural verbs were in fact far more common in the PS condition, as reflected in a significant effect of order,  $t_1(19) = 33.2, p < .001$ ;  $t_2(48) = 17.8, p < .001$ .

For purposes of comparison, the rate of plural verbs was 1% in the SS condition and 98% in the PP condition. Thus, unlike in Experiment 2, speakers rarely used plural verbs with SS items. Thus, if the occasional plural verb in Experiment 2 reflects the influence of the *yes-no* reading, speakers in the current experiment apparently did not consider this reading.

As was done for Experiment 2, during debriefing participants were asked whether they had any ideas about what the experiment was investigating. Responses from 8 of the participants (40%) were related to the choice between *is* and *are*. To determine whether awareness had any effect on agreement behavior, an additional ANOVA was conducted with group (aware vs. naive) and order (SP vs. PS) as variables. The pattern of data in the two groups was almost identical, and a proximity effect was evident for both groups. The rate of plural verbs was 0% (SP) versus 97% (PS) for

<sup>1</sup> An analysis by items could not be conducted because dividing the participants into two groups resulted in too many empty cells.

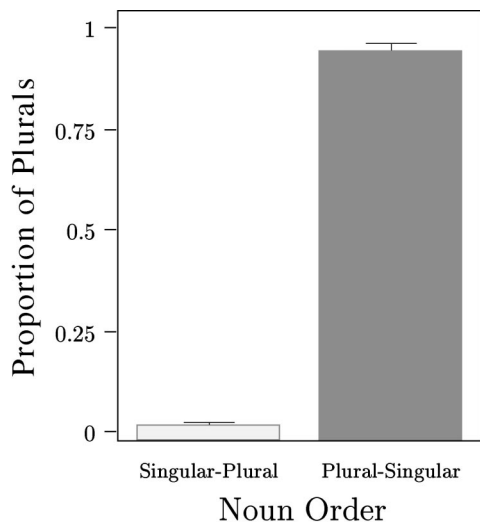


Figure 4. Mean proportion of plural verbs by condition in Experiment 3.

the aware group and 3% (SP) versus 95% (PS) for the naive group. Neither the main effect of group nor the interaction with order approached significance (all  $F$ s < 1). Thus, as with Experiment 2, the proximity effect observed in the current experiment appears not to be caused by conscious awareness of agreement processing. Unlike Experiment 2, however, the interaction of group and order did not approach significance.

The results of the current study indicate that when speakers produce sentences in which the verb precedes a disjunctive subject, agreement is predominantly with the first noun in the disjunction. This contrasts with the findings of Experiment 2, in which the verb followed the subject and agreement was almost always with the second noun. Thus, the pattern observed in Experiment 2 cannot be

due to a general bias toward agreeing with the second noun in the disjunction; rather, the results strongly argue that speakers inflected verbs to agree with the nearer of the two conjuncts, thus exhibiting a proximity effect.

#### Comparison of Experiments 2 and 3

Although there was a clear proximity effect in both Experiment 2 and Experiment 3, in some conditions this effect was numerically larger than in others. An additional analysis was conducted to compare the magnitude of the proximity effect across conditions and experiments. Each response from both experiments was recoded according to whether it was consistent with the predictions of proximity (scored as 1) or inconsistent with the predictions (scored as 0). Order was recoded as proximate noun singular or proximate noun plural; this was done so it would be possible to assess whether the likelihood of agreeing with the proximate noun depended on whether it was singular or plural. The recoded data are presented in Figure 5.

An ANOVA was then conducted on the recoded data with study and proximate noun number as factors. For the items analysis, the Unique (Type III SS; SPSS Inc., 2003) method was used to adjust for the nonequal cell sizes that resulted from the discarded item from Experiment 2. There was a main effect of Experiment, such that a higher proportion of responses were consistent with proximity in Experiment 3 than were in Experiment 2,  $F_1(1, 38) = 9.0$ ,  $p < .01$ ;  $F_2(1, 97) = 6.5$ ,  $p = .01$ . However, this main effect was qualified by a significant interaction,  $F_1(1, 38) = 15.2$ ,  $p < .001$ ;  $F_2(1, 97) = 11.8$ ,  $p < .01$ . The interaction reflected the fact that for the subject-before-verb, proximate noun singular condition (e.g., "Can you tell me if the horses or the clock is/are red?"), speakers showed some tendency to produce a plural verb (the left-most bar in Figure 5), rather than a singular verb to agree with the nearer noun, whereas in all other cases agreement was almost always with

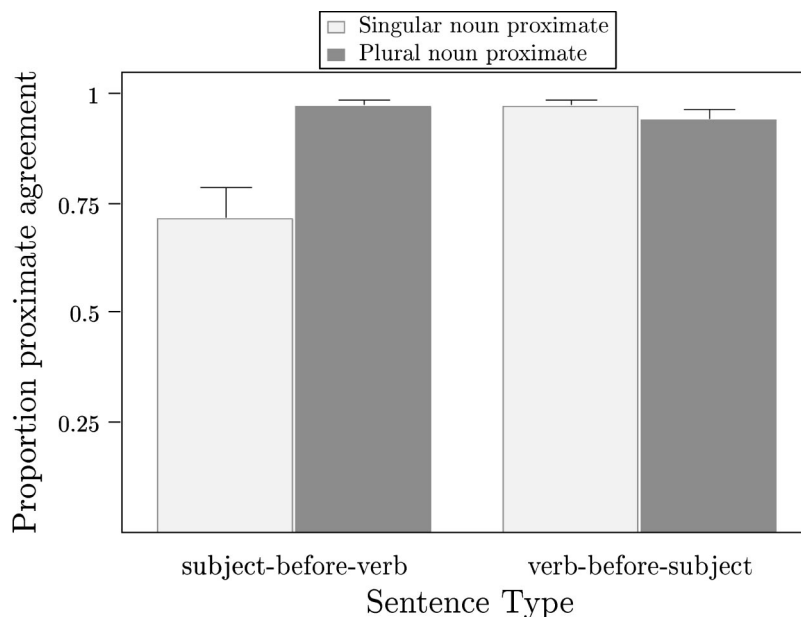


Figure 5. Mean proportion of agreement with proximate noun by condition in Experiments 2 and 3.

the nearer noun. The crucial point of these analyses is that the proximity effect showed a somewhat different pattern in Experiment 2 than in Experiment 3. Below we consider whether these differences might be related to distributional patterns in the English language.

### *Why Should Proximity Matter?*

Together, the findings from Experiments 1–3 clearly indicate that in at least one construction of English (disjunctions), linear order of elements has a robust effect on the production of agreement. This finding presents somewhat of a puzzle. The findings of K. Bock and Cutting (1992), Franck et al. (2002), and Vigliocco and Nicol (1998) firmly establish that structural relationships are much more important than are linear relationships in the process of computing agreement. In fact, within generative linguistics agreement is typically defined in structural terms (e.g., Chomsky, 1981). Why should linear order play any role at all?

One possibility is related to a contrast drawn by Vigliocco and colleagues (Vigliocco & Franck, 1999; Vigliocco & Hartsuiker, 2002; Vigliocco & Zilli, 1999) between *minimal-input* and *maximal-input* approaches to language production. In the minimal-input approach, production is efficient and accurate because it is carried out by several special-purpose modules that operate on limited types of information. In the maximal-input approach, which Vigliocco and colleagues favor, production is efficient and accurate because the production system makes use of all available and relevant information. The use of multiple types of information protects against errors that might occur if more reliable cues are lost or corrupted during the production process, and it can increase efficiency by allowing the production mechanisms to start working on certain processes as soon as any source of relevant information is available, rather than waiting for one particular information source. Perhaps linear order serves as one of several types of information that help to ensure accuracy in agreement production.

Some terminology introduced by MacWhinney and Bates (1989) provides a convenient way to talk about the usefulness of different cues in agreement processing. The “validity” of a cue is a function of two factors: How available it is and how reliable it is (i.e., how often it points to the correct conclusion). For example, one potential cue is the plausibility of the relationship between a given noun and the verb. For example, in *The door to the room slammed shut* it is more plausible for a door to slam than for a room, and the door is also the subject of the verb. Of course, sometimes plausibility points in the wrong direction (e.g., *The bystander near the soccer player kicked the ball*), and sometimes both nouns are equally plausible, in which case plausibility is not available as a cue. However, it seems that plausibility is usually available and fairly reliable and so likely has moderate cue validity. In accord with this, Thornton and MacDonald (2003) found that the extent to which the local noun is plausibly related to the agreeing verb affects the rate of agreement errors (but see Barker, Nicol, & Garrett, 2000).

In the case of linear order, whenever two nouns occur in a sentence, one must come before the other. Thus, linear order is always available. It is not clear, however, whether linear order information provides a reliable cue for identifying which noun should control the number of the verb. To evaluate the potential influence of a linear order cue, we carried out a corpus analysis

examining the typical linear relationships of nouns controlling agreement and the verbs that agree with them. If linear order of nouns relative to the verb is a reliable cue for subject–verb agreement, then we should find a strong correlation between linear order of the various nouns in a sentence and which of those nouns happens to be the agreement controller.

### Corpus Analysis

The corpus analysis had two goals. The first was to determine the extent to which linear order is related to the identity of the agreement controller in the experience of English speakers or, in the terms of Bates and MacWhinney (1982), the reliability of linear order as a cue to agreement. The second goal was to determine the extent to which particular distributional patterns in the English language might be linked to the particular pattern of proximity effects observed in Experiments 2 and 3. In this context, we use the term *agreement controller* to refer to the particular noun within the subject–noun phrase that appears to dictate the number of the verb and *subject* to refer to the entire phrase (this distinction becomes critical later on). To investigate this issue, we used a corpus of English text as an approximation of speakers’ experience and applied a heuristic for identifying the agreement controller based on linear order information to this corpus. If the heuristic were to yield significantly better performance than a chance baseline, then it could be concluded that linear order is a potentially reliable cue to identifying the agreement controller.

### *Method*

*Choice of corpus.* The sample of English sentences used in the current study was taken from a version of the Brown corpus (Francis, 1964; Marcus, Santorini, & Marcinkiewicz, 1993) that was tagged for part of speech. Because subject–verb pairs are not explicitly marked in this corpus, coding was done by hand.

*Selection and coding of sentences.* In order to obtain a representative sample from the corpus, sentences were randomly chosen from the corpus until 500 “codable” sentences were obtained. A sentence was considered codable if it was possible to identify a unique nominal as the agreement controller along with a corresponding agreeing verb and if this nominal could be identified as either singular or plural. For purposes of this analysis, “nominal” was used fairly loosely to include common nouns, proper nouns, cardinal numbers, personal and demonstrative pronouns, and some quantifiers (when used like nouns). To carry out the analysis, it was necessary to identify the agreement controller, the verb, and all noncontroller nominals in each codable sentence. After dropping all other elements from the sentences, this yielded a representation of the serial order of all nominals and the verb.

### *Results and Discussion*

The goal of the analysis was to determine whether the serial order of nouns and verbs in a sentence provides any information which could be useful in identifying the noun controlling agreement. If order information is useful, then a principle based on it would result in the correct identification of the agreement controller more often than the “chance” value obtained when the same principle was applied after the order of the nouns had been scrambled (thus eliminating any information inherent in their linear order). The chance value was estimated by selecting one nominal at random from each coded sentence; the chance value was taken



to be the proportion of times this procedure identified the true agreement controller. Averaging over 10 rounds of random noun selection, the true agreement controller was identified 37.5% of the time.

This chance value was compared with the value resulting from a procedure that selected, out of all nouns that preceded the verb, the noun that was closest, or, if no nouns preceded the verb, the closest following noun. This procedure was intended as a simple heuristic, and it is not claimed that this is what speakers actually do. However, it provides a straightforward metric of the potential usefulness of linear order information. That being said, the heuristic procedure identified the true agreement controller 82.3% of the time. This percentage was significantly higher than the chance baseline,  $\chi^2(1, N = 1,000) = 209, p < .001$ .

It should be noted that our procedure actually used both proximity and order information (i.e., whether a noun appears before or after the verb). The results of Experiments 2 and 3 suggest that speakers are using proximity specifically, above and beyond relative order of nouns and verbs. Thus, it is important to demonstrate that proximity on its own is a useful cue to identifying the agreement controller and that the success of our heuristic is not based solely on the use of relative order.

To address this issue, we selected from the original sample just those sentences in which two nouns preceded the verb or two nouns followed the verb, corresponding to the conditions in Experiments 2 and 3. This resulted in a set of 118 sentences. We then recomputed the success rate of the linear order procedure on just these sentences. Note that in this case, the linear order heuristic reduces to a proximity heuristic, as none of the sentences had nouns both before and after the verb. With this set of sentences, the heuristic was successful 67.8% (80/118) of the time. This success rate was significantly different than the chance rate of 50% ( $p < .001$ , by binomial test).

The results of this analysis demonstrate that linear order in general, and proximity in particular, is a good cue to the identity of the agreement controller. Thus, it is plausible that distributional cues might have contributed to the proximity effects seen in Experiments 2 and 3. However, although overall effects of proximity were quite strong in those experiments, there was one case in which speakers frequently appeared to produce agreement with the farther noun, namely, the PS condition in Experiment 2, in which the subject preceded the verb. A stronger test of the ability of distributional patterns to account for our data would be to examine whether proximity is a somewhat less reliable cue in this particular case. Thus, we conducted a further analysis in which we compared the success rate of the linear order heuristic for subject-before-verb SP structures, in which speakers showed a strong tendency to agree with the nearer noun, and subject-before-verb PS structures, in which this tendency was much weaker.

For this analysis, we began by selecting all sentences from our original sample in which exactly two nominals preceded the verb and in which the verb agreed with one of these two nouns (i.e., subject-before-verb order). This set of sentences was further limited to those in which the first nominal was singular and the second plural (SP), or vice versa (PS). This procedure yielded only 27 sentences. In order to have sufficient data on which to base our conclusions, an additional 73 sentences of this type were sampled from the Brown corpus and coded as before to yield a total sample of 100 sentences (58 SP and 42 PS). We calculated how often the

verb agreed with the proximate (second) noun separately for the SP and PS cases. In the SP case, the proximity procedure was accurate 75.9% (44/58) of the time. In the PS case, the proximity procedure was accurate only 38.1% (16/42) of the time. This difference was statistically significant,  $\chi^2(1, N = 100) = 12.9, p < .001$ . Thus, focusing on the case of subject-before-verb order, in the Brown corpus, a proximity procedure is much less reliable in PS constructions than SP constructions, and in Experiment 2, deviations from proximity occurred frequently with PS phrases but almost never with SP phrases. This close parallel between the distributional patterns in the corpus and the behavior of speakers supports the claim that speakers' behavior might have a distributional basis.

As a final point, it should be noted that the corpus used in this study primarily consists of text drawn from written sources. It is possible that the distribution of nouns and verbs in spoken language would be somewhat different. As a preliminary step toward addressing this concern, we conducted a small-scale replication of the corpus analysis with two randomly chosen spoken conversations from the SWITCHBOARD corpus (Godfrey, Holliman, & McDaniel, 1992). A total of 93 sentences were coded, of which only 6 had no nouns before the verb. We therefore analyzed only overall proximity results and did not compare proximity effects across word order. The overall success rate of the proximity heuristic was 95.6% (89 out of 93 cases). This suggests that a tendency toward proximity is present in spoken as well as written language. An examination of the conversations suggests that this is largely because in speech subjects tend to be simple noun phrases with no noncontroller noun intervening between the controller and the verb.

## General Discussion

This article began by contrasting two views of the process of constituent assembly during language production. In the two-stage approach, the hierarchical relations among constituents are computed in one stage, followed by a second stage in which the constituents are placed in their final linear order. In the one-stage approach, both of these tasks are accomplished in a single stage. In the case of constructions such as *the helicopter for the flights*, there is abundant evidence that mere linear proximity between the noncontroller noun and the verb does not measurably affect the likelihood that the verb will erroneously agree with that noun. This finding has been used to argue that agreement must follow the determination of hierarchical relations but precede linearization, which suggests that these two processes occur during distinct stages. We investigated whether proximity effects might emerge with a different construction, namely, disjunctions.

Experiment 1 demonstrated that, for agreement with disjunctions, people find a verb that agrees with the nearer noun more acceptable than a verb that agrees with the more distant noun. Experiments 2 and 3 extended this finding to an online production task. In producing sentences with disjunctions, speakers exhibited a strong tendency toward agreement with the noun nearest the verb, whether the verb followed (Experiment 2) or preceded (Experiment 3) that noun.

It was argued that speakers of English might use linear proximity information during agreement because it is a fairly reliable cue as to the noun the verb should agree with, and therefore the use

of proximity can help protect agreement processing from errors. A corpus analysis was conducted to show that proximity is, in fact, a reliable cue to agreement. In addition, in the one case in which we observed weaker proximity effects in our online experiments, proximity was also a less reliable cue in the corpus. These findings support a close link between speakers' experience with the language and their agreement behavior.

In the remainder of the general discussion, we discuss four main issues. First, we consider possible alternative accounts and limitations of our findings. Second, we outline a potential mechanism for proximity effects. Third, we consider the implications of our findings for theories of agreement. Finally, we return to the main point of the article—the debate between the single-stage and two-stage views—and discuss how our results bear on this debate.

### *Alternative Accounts of Our Findings*

In this section, we consider three basic alternative accounts of our findings: (a) that our results stem from particular characteristics of our task, (b) that the results reflect monitoring or feedback from later levels of processing, and (c) that the behavior of disjunctions is not representative of agreement as a whole.

*Is our task special?* One possible concern is that our task encouraged particular strategies or heuristics that yielded apparent proximity effects. Specifically, in Experiment 3, speakers may have looked at the first picture, planned and articulated the first part of the sentence (e.g., “Are the horses. . .”), and only then looked at the second picture. The reverse pattern may have occurred to some extent with Experiment 2.

It seems unlikely that this is the only cause of the proximity effects, as similar effects were found using a different task in Experiment 1 and also by Peterson (1986). This is not to say that our task is perfect—all experimental tasks are only approximations to the natural phenomenon being studied. For example, in the fragment completion task, participants must first comprehend and remember the preamble before producing their utterance, potentially providing an opportunity for processes outside of production to influence the results (this issue is discussed further below). On the other hand, because the card sorting task requires the use of pictures, it is of limited usefulness for studying several categories of nouns, including mass nouns like *sand* and collective nouns like *committee*. Ultimately, the most reliable conclusions about agreement processing are those based on converging evidence from multiple paradigms. Thus, research on agreement production could benefit from further development of tasks to complement fragment completion and, where possible, the direct comparison of results obtained through the use of different tasks. In the meantime, we believe that when the same degree of critical scrutiny is applied to all available paradigms, the card sorting task has fewer shortcomings than most and therefore represents a methodological step forward.

*Monitoring and feedback.* A second alternative explanation for why verbs tend to agree with the nearer noun appeals to a monitoring mechanism or to feedback within the production system. It is widely accepted that speakers monitor their own speech for errors and that speech is available for monitoring after being formulated but before it is actually articulated (see, e.g., Hartsuiker & Kolk, 2001; Postma, 2000). Typically, it is assumed that this monitoring is carried out by the comprehension system, that is, that

it is external to language production. Assuming the monitor is sensitive to “local coherence” (see Tabor, Galantucci, & Richardson, 2004), it might classify a sentence such as *The boy or the girls is getting it* as an error, because by itself *girls is* sounds incongruous. In contrast, a sentence such as *The boy or the girls are getting it* would not be identified as a potential speech error. If the proximity effect were due to monitoring in this way, it would have little bearing on the “one stage” versus “two stage” issue.

However, it is not clear how a monitoring account would explain the different pattern of effects for subject-before-verb and verb-before-subject orders. If anything, it seems that it would predict a stronger proximity effect in the subject-before-verb than the verb-before-subject case. This is because *girls is* should be classified as an error more readily than *is girls*, because nouns following verbs are often objects which do not have to agree with the verb in number. In contrast, the only significant deviation from proximity in our data comes in one of the subject-before-verb conditions. For these reasons, monitoring does not appear to provide an adequate explanation of the current findings.

Closely related to monitoring is the possibility of feedback within the production system—essentially an internal monitor, rather than an external one. Specifically, the stage at which linearization occurs might come after the stage at which agreement computation occurs but because of feedback linearization could exert some influence on agreement. Some authors have argued that feedback is the best explanation of certain effects on agreement, in particular effects of morphophonology (see Vigliocco & Hartsuiker, 2002, for discussion). One attractive property of a feedback-based account is that it provides a natural explanation of why effects of structural factors are large and robust, whereas effects of linear order are much more elusive. Unfortunately, as with monitoring, it is not clear how this approach could explain the different pattern of proximity effects seen with subject-before-verb and verb-before-subject order.

*Are disjunctions special?* A third possible concern with our experiments is that the findings with disjunctions might reflect processes or circumstances specific to this construction rather than mechanisms of agreement production more generally. There are several possible shapes such an alternative might take. We consider each in turn.

*Asymmetrical disjunctions.* In a number of non-English languages, verbs may sometimes agree with only one member of a conjunction. For example, Aoun, Benmamoun, and Sportiche (1994) discussed dialects of Arabic in which a verb that appears before a conjoined subject may optionally agree with only the first (nearer) conjunct. This kind of agreement is often termed *partial agreement*. One explanation that has been proposed for these facts is that conjunctions can have an asymmetrical structure, such that one of the conjuncts is in a privileged position for purposes of agreement (Johannessen, 1996; Munn, 1999). For example, Johannessen posited the existence of a “conjunction phrase,” in which one conjunct fills the specifier position, while the other fills an argument position. It is plausible that this account could be extended to disjunctions as well.

Such an approach can explain why agreement with only one noun in a conjunction is sometimes observed. However, because the order of the specifier and argument in a conjunction phrase is fixed within a given language, it predicts that this kind of agreement will always occur with the same conjunct (e.g., the first one).

The contrast between the results of Experiments 2 and 3 demonstrates that this is not the case with English disjunctions; agreement truly seems to be dominated by the nearer noun, regardless of whether it occurs first or second. Thus, this approach does not provide a satisfying account of our data.

*Disjunctions of clauses.* Another explanation that has been put forward for partial agreement with conjunctions appeals to the fact that speakers may in some cases omit redundant portions of conjoined sentences, resulting in what appears to be agreement with just one of the conjuncts. Aoun et al. (1994) applied this type of analysis to partial agreement in Arabic.

Extending this analysis to disjunctions provides a potential explanation for our results. That is, participants might have initially planned to produce separate predicates for each noun (e.g., "Can you tell me whether the horses are red or the clock is red?"), then omitted the first predicate. Arguably, if the second predicate rather than the first is omitted, a similar mechanism could account for the inverted verb-before-subject order as well. However, this sort of account would have to explain why the first predicate is omitted in one case and the second in the other and would thus likely have to invoke proximity in some way. This account also does not explain why speakers showed a much reduced tendency to agree with the nearer noun in the subject-before-verb PS condition.

*Agreement with disjunctions is different than other kinds of agreement.* Even if the above alternatives are not viable, it still might be the case that agreement with disjunctions is a special case that is handled differently than cases of "ordinary" agreement. As mentioned earlier, there are abundant suggestions in the linguistics literature that linear order effects also occur with conjunctions and in a wide range of languages. It has also been argued that linear proximity effects occur with so-called equative sentences, as in the following examples adapted from Allerton (1992):

- 4a. The poor qualifications of the applicants has/have become the main obstacle.
- 4b. The applicants' poor qualifications has/have in recent weeks become the most serious obstacle.
- 5a. Among linguists, what seems to matter most is/are the arguments.
- 5b. For linguists, the most important factor is/are the coherence and logical soundness of the arguments.

For some speakers, agreement with the postverbal noun phrase is acceptable in the (a) sentences but not the (b) sentences. The crucial difference is whether the preverbal or postverbal noun phrase is linearly closer to the verb.

Naturally, these observations suffer from the same shortcomings as the survey conducted by Peterson (1986), which provided the motivation for the experiments presented here. Ultimately, informal observations should be supplemented by controlled psycholinguistic experiments, in particular experiments involving a production task rather than grammaticality judgments. However, even if it turns out that proximity effects are limited to a few specific cases, this would still have important implications for the debate between single-stage and two-stage models. In the two-stage

agreement-early account, linear order is prevented from affecting the computation of agreement by the basic architecture of the production system. Even demonstrating a single case in which proximity effects do occur would seem to require revision of this claim.

*The role of prescriptive pressures.* As with many grammatical phenomena, the matter of agreement with disjunctive noun phrases has received some attention in style manuals. *The American Heritage Book of English Usage* (American Heritage Dictionaries, 1996) has this to say about agreement with either/or constructions:

When the construction mixes singular and plural elements, however, there is some confusion as to which form the verb should take. Some people argue that the verb should agree with whichever noun phrase is closest to it. The Usage Panel has much sympathy for this view. Fifty-five percent prefer the plural verb for the sentence *Either the owner or the players is going/are going to have to give in*. Another 12% find either verb acceptable. . . (p. 16)

This quote, which is representative of grammarians' preferences, is generally consistent with our findings. The judgments described in this quote may therefore reflect proximity-based processes in the grammarians' production systems; alternatively, our data may reflect experiment participants' allegiance to explicitly taught prescriptive rules of agreement in disjunctions.

It is in principle impossible to rule out any role for prescriptive pressures in our results. To the extent that dictates on proper usage become incorporated into the grammar, whether of an individual speaker or a speech community, they can no longer be differentiated from any other idiosyncratic aspect of English grammar. The crucial question is, if speakers were following prescriptive rules, were they doing so consciously? The debriefing interviews for Experiments 2 and 3 suggest that any application of prescriptive rules was not conscious. Although many participants became consciously aware of the choice between *is* and *are*, not a single participant articulated anything resembling the usage guidelines above. Similarly, the fact that the aware and unaware groups showed the same pattern of results militates against a prescriptive explanation.

Moreover, disjunctions are hardly the only construction investigated in psycholinguistic experiments to be subject to prescriptive recommendations in style books. For example, agreement with collective terms such as *family* is extensively discussed in grammar texts, yet collectives have played a central role in a number of agreement studies (e.g., K. Bock, Nicol, & Cutting, 1999; Haskell & MacDonald, 2003). Thus disjunctions are in good company in this regard. On all these counts, it seems that disjunctions, along with other constructions for which explicit prescriptive teaching may have occurred, can offer insight into agreement processes.

#### *An Activation-Based Mechanism for Proximity Effects*

If the proximity effects we have observed are not naturally explained by one of the alternatives above, can they be explained through the distributional patterns found in the corpus analysis? One possible explanation builds on the idea that at the time agreement is computed, one of the nouns in the disjunction is more highly activated than the other.

Dell, Schwartz, Martin, Saffran, and Gagnon (1997) have described production as requiring the coordination of three simulta-

neous tasks: preparing to produce what comes next, producing what comes now, and suppressing production of what came just before. In the subject-before-verb case, if the verb form is chosen after the first noun has already been produced, then activation of this noun may be suppressed to ensure that it does not get produced again. This would result in higher activation for the second (and nearer) noun. In the verb-before-subject case, the reverse situation obtains; if, at the time the verb form is being chosen, the first noun is being prepared for production but little or no planning of the second noun has taken place, then there will be much higher activation for the first (and nearer) noun.

Thus, there is a correlation between a noun and a verb having an agreement relation and a noun and a verb being coactivated. These patterns of activation may come to be linked to agreement through a process of implicit or associative learning (cf. J. K. Bock & Griffin, 2000; Chang, Dell, Bock, & Griffin, 2000). In this view it is this coactivation, rather than linear proximity directly, that serves as the cue. In terms of syntax, there is nothing necessary about this relationship. Rather, it is a learned cue that helps the production system function rapidly and efficiently, similar to the maximal-input approach discussed earlier (e.g., Vigliocco & Hartsuiker, 2002).

One advantage of this approach is that it provides a principled basis for the use of proximity, as opposed to other distributional statistics that could be imagined. For example, in English a heuristic such as “the first noun in the sentence controls agreement” would likely do as well or better than a proximity-based heuristic. However, there is no obvious way in which “first noun” would be reflected in the processing dynamics of the system. An additional advantage is that coactivation, being based on linear order, is likely to be a language specific cue. In subject-object-verb languages such as Japanese, it is the object and the verb that are more likely to be coactivated. In an account based directly on activation dynamics, proximity confers an advantage in English but might cause interference in Japanese. In our approach, coactivation is only used as a cue if it is correlated with agreement, thus leading to implicit learning.

This possibility is clearly highly speculative, but it is broadly consistent with earlier claims that whether or not a noun and a verb are part of the same planning unit strongly affects how likely that noun is to influence agreement. For example, K. Bock and Cutting (1992) found that agreement errors were more likely with fragments such as *the editor of the history books* than *the editor that rejected the books*. This difference was attributed to the clause being an important planning unit in the production process, such that in the first case *editor* and *books* would be planned together, whereas in the second case they would be planned as part of separate clauses and therefore at separate times, presenting less possibility for interference. Presumably, a noun planned as part of the same clause as the agreement controller would be more active at the time agreement is being computed than a noun planned as part of a different clause.

An activation-based account of this sort could be related to several other proposals in the production and comprehension literature. First, it is reminiscent of accessibility accounts in production, in which the activation of nouns is thought to affect other syntactic-level production processes, such as choice of syntactic structure and word order (J. K. Bock, 1982). Second, this view is broadly consistent with Gibson and colleagues’ (e.g., Gibson,

1998; Grodner, Gibson, & Tunstall, 2002) locality proposals in comprehension, in which computational cost is incurred by increased distance between dependent elements in a sentence. On this view, having a nearby noun controlling agreement would exert a lower computational burden than having a more distant noun controlling agreement. To the extent that computational burden influences decisions speakers make, this could encourage agreement between proximate elements.

In this regard it is interesting to note that the fragment completion task, used in the majority of studies on agreement, may mask effects that are due to activation differences. In that task, the entire phrase is comprehended and held in memory before it must be produced. Thus, both nouns must be kept activated throughout the task. This may artificially attenuate the asymmetry in activation levels between them. More generally, relative activation levels of different nouns may be influenced by the particular task used. This issue remains to be investigated in future research.

### *Implications for Theories of Agreement Production*

Our findings clearly show that, at least in the case of disjunctions, linear order exerts an effect on agreement and therefore that linear order information must be available at the time agreement is computed. It is less clear precisely where in the agreement process proximity exerts its effects. In this regard, it is possible to distinguish between two different processes that must take place during production: function assignment and computing the number of the subject noun phrase.

The process of function assignment associates particular noun phrases with particular grammatical roles, such as the role of subject. It has been argued that some apparent agreement phenomena actually reflect function assignment (see, e.g., K. Bock & Miller, 1991; Hupet, Fayol, & Schelstraete, 1998). However, we do not believe this is occurring in the current experiments. A comparison between the results of Experiments 2 and 3 shows that not only is agreement affected by the order of the nouns within the disjunction but also by the position of the disjunction as a whole relative to the verb (i.e., subject-before-verb vs. verb-before-subject order). Thus, information about the linear position of the verb must have been available at whatever stage our effects arose. It is therefore difficult to see how our effects could have arisen any earlier than linearization, far downstream from function assignment.

In our view, the effects arise with a different process, computing the number of the subject. A similar line of reasoning was followed by Hartsuiker, Anton-Mendez, and van Zee (2001), who found that object noun phrases can occasionally disrupt agreement in Dutch. In explaining these findings, the authors proposed that the subject noun phrase was identified correctly but that constituents anywhere in the tree (including object noun phrases) can come to influence the number that is marked on this subject noun phrase. Computing the number of the subject is commonly believed to be a core aspect of agreement proper (e.g., K. Bock et al., 2001). Thus, if our proximity effects do reflect difficulty during this process, then they can properly be called agreement effects, and theories of agreement production should account for them.

Previous work examining the role of linear order in agreement (e.g., K. Bock & Cutting, 1992; Franck et al., 2002; Vigliocco & Nicol, 1998) has primarily focused on the relative contributions of

structural factors and linear order. However, in terms of the basic architecture of the system, it is a very different thing to say that order information has a negligible effect on agreement than to say that it cannot, in principle, have any effect on agreement. Although it is too early for any strong conclusions, our results suggest that the former claim is more accurate.

### *Implications for Single-Stage Versus Two-Stage Models*

The major question addressed by the experiments reported in this article is whether constituent structure is built in a single stage or two separate stages. Pickering et al. (2002) reported syntactic priming results argued to support a single-stage model and to be inconsistent with a two-stage model. Franck et al. (2002) and Vigliocco and Nicol (1998) presented agreement error data argued to support a two-stage model. With respect to the debate between single-stage and two-stage models, the crucial aspect of their data is that they failed to find any effects of linear order on agreement.

The current experiments suggest that, under certain conditions, linear proximity can have a measurable influence on agreement. Do these findings rule out the two-stage model? With respect to accounts in which agreement occurs early, strictly before linearization (Franck et al., 2002; Vigliocco & Nicol, 1998), the answer would seem to be *yes*. However, it is possible to devise a two-stage account in which agreement occurs later, concurrently with linearization. Such an approach could accommodate linear order effects such as the ones we observed.

Given that either a single-stage or a two-stage agreement-late model can account for our findings, which is to be preferred? At the beginning of this article, we noted that the structural priming data of Pickering et al. (2002) seemed to support the single-stage model, whereas the failure to find effects of linear order on agreement by Franck et al. (2002) and Vigliocco and Nicol (1998) supported a two-stage model, resulting in something of a deadlock. Given our findings, the agreement data can no longer be taken to support the two-stage model. Thus, although we cannot offer a definitive answer, it would seem that the balance of the evidence now favors the single-stage approach.

### References

- Allerton, D. (1992). Problems of modern English grammar II: Disagreement about agreement. *English Studies*, 73, 458–470.
- American Heritage Dictionaries (Ed.). (1996). *The American Heritage book of English usage: A practical and authoritative guide to contemporary English*. Boston: Houghton Mifflin.
- Aoun, J., Benmamoun, E., & Sportiche, D. (1994). Agreement, word order, and conjunction in some varieties of Arabic. *Linguistic Inquiry*, 25, 195–220.
- Barker, J., Nicol, J., & Garrett, M. (2000, March). *Semantic and probabilistic factors in the processing of number agreement*. Paper presented at the 13th Annual CUNY Conference on Human Sentence Processing, La Jolla, California.
- Bates, E., & MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner & L. R. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 173–217). New York: Cambridge University Press.
- Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review*, 89, 1–47.
- Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, 18, 355–387.
- Bock, J. K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General*, 129, 177–192.
- Bock, K., & Cutting, J. C. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language*, 31, 99–127.
- Bock, K., & Eberhard, K. M. (1993). Meaning, sound, and syntax in English number agreement. *Language and Cognitive Processes*, 8, 57–99.
- Bock, K., Eberhard, K. M., Cutting, J. C., Meyer, A. S., & Schriefers, H. (2001). Some attractions of verb agreement. *Cognitive Psychology*, 43, 83–128.
- Bock, K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology*, 23, 45–93.
- Bock, K., Nicol, J., & Cutting, J. C. (1999). The ties that bind: Creating number agreement in speech. *Journal of Memory and Language*, 40, 330–346.
- Chang, F., Dell, G. S., Bock, K., & Griffin, Z. M. (2000). Structural priming as implicit learning: A comparison of models of sentence production. *Journal of Psycholinguistic Research*, 29, 217–229.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, the Netherlands: Foris Publications.
- Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, 104, 801–838.
- Francis, W. N. (1964). *A standard sample of present-day English for use with digital computers* (Report to the U.S. Office of Education on Cooperative Research No. E-007). Providence, RI: Brown University.
- Franck, J., Vigliocco, G., & Nicol, J. (2002). Subject-verb agreement errors in French and English: The role of syntactic hierarchy. *Language and Cognitive Processes*, 17, 371–404.
- Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production* (Vol. 1, pp. 170–220). London: Academic Press.
- Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68, 1–76.
- Godfrey, J. J., Holliman, E. C., & McDaniel, J. (1992). Switchboard: Telephone speech corpus for research and development. *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1, 517–520.
- Grodner, D., Gibson, E., & Tunstall, S. (2002). Syntactic complexity in ambiguity resolution. *Journal of Memory and Language*, 46, 267–295.
- Hartsuiker, R. J., Anton-Mendez, I., & van Zee, M. (2001). Object attraction in subject-verb agreement construction. *Journal of Memory and Language*, 45, 546–572.
- Hartsuiker, R. J., & Kolk, H. H. J. (2001). Error monitoring in speech production: A computational test of the perceptual loop theory. *Cognitive Psychology*, 42, 113–157.
- Hartsuiker, R. J., & Westenberg, C. (2000). Word order priming in written and spoken sentence production. *Cognition*, 75, B27–B39.
- Haskell, T. R., & MacDonald, M. C. (2003). Conflicting cues and competition in subject-verb agreement. *Journal of Memory and Language*, 48, 760–778.
- Hupet, M., Fayol, M., & Schelstraete, M.-A. (1998). Effects of semantic variables on the subject-verb agreement processes in writing. *British Journal of Psychology*, 89, 59–75.
- Johannessen, J. B. (1996). Partial agreement and coordination. *Linguistic Inquiry*, 27, 661–676.
- Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1–75.
- MacWhinney, B., & Bates, E. (Eds.). (1989). *The crosslinguistic study of sentence processing*. New York: Cambridge University Press.
- Marcus, M. P., Santorini, B., & Marcinkiewicz, M. A. (1993). Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics*, 19, 313–330.

- Munn, A. (1999). First conjunct agreement: Against a clausal analysis. *Linguistic Inquiry*, 30, 643–668.
- Peterson, P. G. (1986). Establishing verb agreement with disjunctively conjoined subjects: Strategies vs. principles. *Australian Journal of Linguistics*, 6, 231–249.
- Pickering, M. J., Branigan, H. P., & McLean, J. F. (2002). Constituent structure is formulated in one stage. *Journal of Memory and Language*, 46, 586–605.
- Postma, A. (2000). Detection of errors during speech production: A review of speech monitoring models. *Cognition*, 77, 97–131.
- SPSS Inc. (2003). *SPSS Base 12.0 user's guide*. Chicago: Author.
- Tabor, W., Galantucci, B., & Richardson, D. (2004). Effects of merely local syntactic coherence on sentence processing. *Journal of Memory and Language*, 50, 355–370.
- Thornton, R., & MacDonald, M. C. (2003). Plausibility and grammatical agreement. *Journal of Memory and Language*, 48, 740–759.
- Vigliocco, G., & Franck, J. (1999). When sex and syntax go hand in hand: Gender agreement in language production. *Journal of Memory and Language*, 40, 455–478.
- Vigliocco, G., & Hartsuiker, R. J. (2002). The interplay of meaning, sound, and syntax in language production. *Psychological Bulletin*, 128, 442–472.
- Vigliocco, G., & Nicol, J. (1998). Separating hierarchical relations and word order in language production: Is proximity concord syntactic or linear? *Cognition*, 68, B13–B29.
- Vigliocco, G., & Zilli, T. (1999). Syntactic accuracy in sentence production: The case of gender disagreement in Italian language-impaired and unimpaired speakers. *Journal of Psycholinguistic Research*, 28, 623–648.

## Appendix A

### Stimulus Items for Experiment 1 (SP versions)

- Can you ask Brenda if the girl or the boys is/are going to go first?  
Do you know if the tiger or the lions was/were roaring louder?  
Do you think the pickle or the olives is/are going to get eaten first?  
Have you heard whether the swimmer or the runners is/are raising more money for charity?  
How do I know if the stove or the refrigerators is/are arriving last?  
I can't tell whether the nurse or the doctors is/are more annoyed.  
Do the directions say if the store or the restaurants is/are further away?  
We need to know if the necklace or the rings was/were more damaged.  
I want to know if the sheet or the blankets is/are longer.  
I need to know if the table or the benches is/are heavier.  
Maria probably knows if the photograph or the paintings is/are more expensive.  
It didn't matter to me if the magazine or the books was/were placed on top.  
It is hard to tell whether the miner or the truckers is/are more drunk.  
Ask Ronnie if the ruby or the diamonds is/are going in the nicer box.  
I wonder if the bee or the flies is/are bothering Vijay more?  
It doesn't really matter whether the contractor or the banks is/are losing the most money.  
Can you tell me whether the swing or the slides is/are more popular?  
Do you think the window or the walls is/are more likely to be damaged in an earthquake?  
Do you remember if the dresser or the beds was/were bought first?  
Did Naomi say whether the bookshelf or the desks is/are going to be moved first?  
Can you ask the guide if the sword or the guns is/are more valuable?  
Did Samantha say whether the lamp or the plants was/were harmed more in the flood?  
Can you tell me if the TV or the radios was/were sold more quickly?  
Can you tell me whether the mug or the cups is/are dirtier?  
The guidebook must say whether the river or the mountains is/are more picturesque.  
Would you say the copier or the printers is/are breaking down more often?
- Ask the doctor whether the passenger or the drivers was/were injured worse in the accident.  
Marcus will tell you whether the pitcher or the teapots is/are going to be needed first.  
Do you remember if the martini or the beers is/are for table 10?  
Ask the boss if the crate or the boxes is/are supposed to be loaded first.  
I'm confused about whether the statue or the vases was/were insured for more money.  
Do you think the light or the signs is/are more distracting to drivers?  
Find out whether the shovel or the rakes is/are more useful in the garden.  
Did you think the orange or the apples was/were tastier?  
Can you find out if the jar or the bottles is/are more fragile?  
Do you know whether the watch or the cameras was/were more fragile?  
The board wants to know if the dolphin or the sharks was/were attracting more visitors at the aquarium.  
Vivian must know whether the bookstore or the galleries was/were closer to School Street.  
Can you tell me whether the racket or the balls was/were newer?  
Tell me whether the outlet or the switches is/are broken first.

## Appendix B

### Objects and Animals Depicted in the Pictures for Experiments 2 and 3

airplane	bed	bike	bookcase
bowl	bow tie	camera	car
cat	chair	clock	coathanger
dog	fork	glass	globe
handmirror	hat	horse	knife
lamp	map	pan	pillow
phone	piano	pitcher	plate
pot	rabbit	radio	sofa
spoon	stapler	suitcase	table
television	tennis racket	toaster	vase

Received June 12, 2003

Revision received March 5, 2005

Accepted March 15, 2005 ■