

## The effect of the predicate on agreement error rates\*

*Christine M. Pittman*

*Ron Smyth*

University of Toronto

Subject-verb agreement errors exhibit patterns that allow researchers to test models of language production and comprehension. In this paper we address the question of whether semantic plausibility affects the implementation of agreement in an elicited production task. In particular, we investigate whether a nonsubject (local) noun that is also a plausible subject for the verb (1) will cause more agreement errors than an implausible local noun (2):

- (1) The boy near the dogs is/\*are running away.
- (2) The boy near the trees is/\*are running away.

If so, we may also ask whether these are always superficial concord errors, or whether they sometimes arise because the participant has mistaken the local noun for the head.

### 1. BACKGROUND

Agreement “attraction” appears to involve erroneous agreement with the local noun (Quirk, Greenbaum, Leech & Swartvik 1972, Francis 1986). In an agreement elicitation task, Bock and Miller (1991) gave participants sentence preambles in the form of a complex subject, and asked them to repeat each one and complete the sentence. They found a previously-unnoticed asymmetry whereby local plurals led to more errors than local singulars:

- (3) The key to the cabinets (SP)... more errors (e.g. "are rusty")
- (4) The keys to the cabinet (PS)... fewer errors (e.g. "is rusty")

They interpreted this as evidence that plurals, being morphosyntactically marked, bear a number feature which can interfere with the agreement implementation process, whereas singulars, being unmarked, have no such feature and are transparent to the agreement process.

Eberhard (1997) presents an activation model of agreement: singular agreement is the default, and plural agreement occurs when a plural feature in the subject phrase has enough activation to be “noticed” by the verb. Assuming that subject nouns have higher activation than others, a number feature associated with a subject noun will trigger normal plural agreement. Any other

---

\* **Acknowledgements:** This research was supported by a SIG grant from the Social Sciences and Humanities Research Council of Canada to Christine Pittman. We would like to thank Ana-Teresa Perez-Leroux and Elizabeth Cowper for comments on earlier drafts.

number features in the sentence, associated with nouns that have lower activation, will cause noise in the system. If the activation level on a non-subject plural noun is high for any reason, it can interfere and cause plural agreement.

Many other factors can affect agreement errors because of their effect on activation levels on nouns and/or their number features. For example, the *conceptual plurality* of a singular head noun (distributivity), as in (5), leads to more errors than for sentences like (3):

- (5). The label on the bottles is/\*are...

This has been confirmed for English (Eberhard 1999), Italian (Vigliocco, Butterworth & Semenza 1995), Spanish (Vigliocco, Butterworth & Garrett 1996), French (Vigliocco, Harsuiker, Jarema & Kolk 1996) and Dutch (Vigliocco, Harsuiker, Jarema & Kolk 1996). In terms of the activation model, the plural conceptual number of *label* increases plural noise and causes errors.

*Animacy* also affects agreement errors (Barker, Nicol and Garrett 2001). Animates are more "subjectworthy" and likely have higher activation than inanimate subjects. Barker et al. found that sentences with animate subjects are slightly less susceptible to agreement errors than those with inanimate subjects. Similarly, *Semantic Overlap* (Barker, Nicol and Garrett 2001) leads to more errors like (6) than (7). We interpret this to mean that the shared semantic features of *canoe* and *sailboat* induce semantic priming, leading to higher activation on the local noun:

- (6) The canoe by the sailboats \*are sturdy. more errors  
 (7) The canoe by the cabins \*are sturdy. fewer errors

Errors are also more frequent with (8) than (9), due to *Semantic Integration* (Soloman and Pearlmutter 2004):

- (8) The drawing of the flowers (direct link – NPs are integrated)  
 (9) The drawing with the flowers (indirect link – NPs not as integrated)

## 2. THE ROLE OF THE PREDICATE

Thus far the factors that increase activation levels all involve the subject phrase. Recently, researchers have asked whether other elements in the sentence can increase activation levels on the NPs, and thus affect error rates. Barker, Nicol and Garrett (2001) conducted a post-hoc analysis of their materials based on compatibility ratings, but found no such effect. However, this was a weak test of the plausibility hypothesis, since the division of materials into plausible vs. implausible predicates was not planned. On the other hand, Greenslit and Badecker (2000) did find the effect in a comprehension study. Reading times showed an effect of semantic compatibility between the predicate and distractor noun, and for the number of the distractor noun.

Thornton and MacDonald (2003) found predicate plausibility effects on error rates for both production and comprehension. They presented participants with items like (12), where the task began with a target verb, followed by the preamble, which were to be repeated and then completed with a form of *be* along with the predicate. They also varied the number of the head and local nouns (*album* and *composer*).

(12) PRAISED/PLAYED The album(s) by the classical composer(s)...

There were more agreement errors when the passive verb was compatible with both nouns, but interestingly this effect was restricted to SP cases. In the comprehension study, reading times were slower at the passive verb under the same circumstances.

Thornton & Macdonald argue for a non-encapsulated, constraint-satisfaction processing model. The key question is whether the plausibility effect "penetrates" the agreement process online, or whether its effects come in later. We propose that during the initial stage of agreement processing (particularly in comprehension where the participant does not have a ready-made message plan, but arguably in production as well), plausibility increases the activation level of the local noun because of its potential semantic relationship with the predicate. If this relationship is not inhibited on the basis of syntactic information (i.e. that the local noun is not the head), then errors may arise. On at least some trials, this could even cause confusion between the local and head nouns, such that the local noun is treated not only as the agreement controller, but also as the subject, leading to an incorrect representation.

We therefore designed an elicited production experiment in which both plausibility and noun number were varied for the first part of the task, using head/local nouns with SP and PS number marking. After producing the initial preamble, verb, and predicate, the participant had to choose a follow-up predicate, conjoined with *and* to the initial response. This second part of the task allowed us to discern whether agreement errors are associated with subject mis-selection: if the local noun is mistakenly chosen as the agreement controller, was it also more likely to be mistakenly interpreted as the head? If so, participants should choose the wrong follow-up predicate. And if this depends on number, we should find more predicate mis-selections for SP preambles than for PS.

Finally, we wondered whether correct agreement in the first phase of the response would suppress the local noun's activation, as suggested above. If so, we expected fewer predicate mis-selections for SP items, where this would suppress the potentially interfering active plural marking, compared to PS items, where the local singular has its regular (unmarked) activation level.

### 3. METHOD

*Participants* Sixty University of Toronto students were paid \$7.50 for the 30 minute study.

*Materials* Forty-four sentence preambles (complex subjects) along with target predicates, were counterbalanced across 4 lists. There were 45 fillers. The items on each list were presented in the same random order.

SP-BTH	TALL	The boy by the trees...
PS-BTH	TALL	The boys by the tree...
SP-NP1	PLAYFUL	The boy by the trees...
PS-NP1	PLAYFUL	The boys by the tree...

*Procedure* Items were displayed on a computer screen using the DMDX software, as follows:

Target predicate (e.g. TALL): 250 msec  
 Blank screen: 50 msec  
 Sentence preamble (e.g. *The boy by the trees*): 1800 msec  
 Participant produces the subject phrase, verb, and predicate  
 Blank screen: 1000 msec  
 Two additional predicates: 2000msec (e.g. CHUBBY/GREEN)  
 "Press the SPACE BAR to continue"

Participants were instructed to remember the predicate, read the sentence preamble out loud, continue the sentence using *is*, *are*, *was* or *were*, then use the predicate (response: "The boy by the trees is/\*are tall"). Finally, they added "and" as well as their choice for the second predicate (see Part 2 below). They were instructed to speak quickly, and were recorded with a minidisk player. They first completed a practice session of 12 items.

*Scoring* Responses to the first part of the task were scored as Correct (C) if the participant said the preamble correctly, used a form of the verb *to be*, used the target predicate in the response and used the correct verb agreement. Responses were scored as Agreement Error (AG) if an incorrect form of the verb *to be* was the only error. Finally, Other (O) responses involved any error other than an agreement error. "Other" responses were discarded, and the analysis was based on the percentage of agreement errors per participant per condition out of the total number of scoreable responses (C + AG). For the second part of the task, we recorded the number of correct and incorrect predicate choices.

#### 4. RESULTS AND DISCUSSION

##### 4.1 Agreement errors

Participants made agreement errors on 12.39% of trials, although there were considerable individual differences.

**Table 1.** Means and standard deviations, verb agreement errors.

Type	<b>AG%</b>	<i>s.d.</i>
SP-NP1	<b>11.70</b>	<i>10.21</i>
PS-NP1	<b>8.58</b>	<i>10.77</i>
SP-BTH	<b>18.39</b>	<i>21.45</i>
PS-BTH	<b>10.90</b>	<i>14.96</i>

*Statistical note:* Agreement error data grossly violate both the normality assumption (all conditions have a sharp spike in the 0 and 1 error range) and the homogeneity of variance assumption (because the means are correlated with their standard deviations). Thus analysis of variance, while used by virtually all agreement researchers, is invalid. We present it for consistency with other research: a 2 (predicate plausibility: NP1 versus Both) x 2 (NP number: SP versus PS) repeated measures ANOVA was calculated. Because the materials were fully counterbalanced, only the analysis by participants is presented.

There was a significant main effect of NP number:  $F(1,59) = 4.65$ ,  $p = .035$  (SP mean = 15.05, PS mean = 9.74). There was also a main effect of predicate plausibility:  $F(1,59) = 9.48$ ,  $p = .003$  (NP1 mean = 10.14, BTH mean = 14.65). The interaction of predicate plausibility and NP number was marginal:  $F(1,59) = .277$ ,  $p = .10$ .

*Non-parametric tests:* The only essential question is whether there is a plausibility effect for SP or for PS. A significant interaction is not necessary for these planned orthogonal comparisons. Thus we used the Wilcoxon signed ranks test. For SP items, the comparison of NP1 vs. Both was significant ( $Z = 3.09$ ,  $p = .002$ ), but for PS items there was no significant difference ( $Z = 0.80$ ,  $p = .422$ ). Our agreement error data provide firm evidence for a plausibility effect on agreement errors, but only for SP items. This is compatible with Thornton & Macdonald (2003).

These findings are consistent with the view that a local plural NP (SP items) has greater activation than a local singular, and that when the local NP is also semantically compatible with the predicate, its activation is further enhanced, causing an even greater number of errors. For PS items, the local singular has lower activation (no number marking; not a subject) and predicate plausibility does not increase the number of errors significantly. We acknowledge that in a more sensitive experiment, one might find a plausibility effect for PS, but it is clear from our results that this would be a weaker effect than for SP. Thornton and Macdonald attribute the lack of a plausibility effect for PS to the fact that

"...error rates in the PS condition are typically on a par with the PP match condition, suggesting a relative insensitivity to variables that increase agreement error rates. Thus, the influence of plausibility might not be observed in the PS condition because of that insensitivity".

While this is true of both their data and ours, their explanation amounts to the questionable claim that because PS error rates are low in some circumstances, they must always be low. To add explanatory value to this finding, we suggest that in the BOTH condition, competition between an overtly marked and highly active subject NP, along with its plausibility, should result in a larger effect than that of the local singular. However, nothing in our theoretical approach rules out the possibility that, perhaps with much larger sample sizes, there could be a smaller plausibility effect for PS.

#### 4.2 SECOND PREDICATE SELECTION ERRORS

The agreement error data show that the predicate's compatibility with a distractor noun affects agreement error rates. But how and why does this happen? One possibility is that the system misanalyses the distractor noun as the subject. However, in the production model used by most agreement error researchers (Levelt 1989; Bock and Levelt's 1994) agreement is computed before structure is built. Thus if the distractor noun is misanalyzed it should be placed in subject position in the participants' responses. This clearly did not happen in our study.

Our task simulates normal production in that participants had the predicate in mind while listening to the preamble. Thus they could have made overt subject selection errors, but they did not. Moreover, Thornton & Macdonald replicated their results with items for which the participants heard the preamble first and saw the predicate only when it was time to respond.

To explain the findings, we propose that the system has strongly considered the distractor noun as a potential subject and this has increased its activation level. This increase in activation can come into play when the agreement process is underway, leading to either a simple concord error or a semantic error (treating the local noun as though it were the subject). That is, the local noun can control either the superficial form of agreement, or its semantic consequences, or both.

There is a precedent for the claim that participants' early confusion leads to errors downstream. Hollingworth, Halliwell and Ferreira (2001) found that the original structure given to a garden path sentence persists even after it has been reanalyzed. After reading a sentence like (13):

(13) While the man hunted the deer paced in the zoo.

participants often answered *yes* to the question "Did the man hunt the deer?", and this reflects the rejected (garden path) interpretation of *deer* as the object of *hunted*. It is possible that if the production system has strongly considered assigning the role of subject to a distractor noun, this interpretation might persist even after it is rejected, at least on some trials.

Recall that during each item of Experiment 1, participants completed the trial by choosing one of the adjectives, conjoined with *and*, to end the sentence. The filler predicates were chosen to be opposites, synonymous or

unrelated to the first predicate, to prevent participants from guessing the aims of the experiment.

SP-BTH TALL	The boy by the trees [...] <i>and</i>	<b>chubby/green</b>
PS-BTH LARGE	The lawyers beside the car [...] <i>and</i>	<b>confident/shiny</b>
SP-NP1 YOUNG	The baby under the blankets [...] <i>and</i>	<b>lovable/fluffy</b>
PS-NP1 POLITE	The girls by the piano [...] <i>and</i>	<b>kind/ancient</b>

In this part of the task, the correct and incorrect predicates appeared on the screen, one above the other, with their order randomly determined. They were displayed for 2000 msec. A "deadline" beep was then played to encourage participants to choose quickly. Participants were told to select the predicate that went best with the sentence they had already constructed, and add it to complete the response.

*Scoring:* The dependent measure was the percentage of incorrect predicate selections, based on the number of useable responses. We created two data sets, one for errors made after a correct initial agreement, and one for errors made after an agreement error.

*Analysis of variance:* As with the first part of the task, we present ANOVA results in order to allow for comparisons with other studies. The correct hypothesis tests are nonparametric and are given below.

A 2 (NP Number: SP versus PS) x 2 (predicate plausibility: NP1 versus BTH) x 2 (Agreement Error: AG versus No AG) repeated measures analysis of variance was conducted on the predicate selection data. The main effect of Plausibility was significant,  $F(1,59) = 23.21$ ,  $p < .0001$ . There were far more errors in the BOTH condition (12.75% vs. 4.15%). The interaction of NP Number by Agreement Error (Table 2) was also significant,  $F(1,59) = 5.03$ ,  $p = .029$ . No other effects were significant, although the 3-way interaction was marginal ( $F(1,59) = 3.01$ ,  $P = .09$ ). Post hoc tests show a Number effect following agreement errors only ( $P < .05$ ); all other comparisons were nonsignificant.

**Table 2.** Mean percentage of incorrect predicate selections (IP) for the Agreement Error x NP Number interaction

AG versus C	NP Number	IP%
After Agr Error	SP	<b>11.45</b>
	PS	<b>6.38</b>
After Correct Agr	SP	<b>6.45</b>
	PS	<b>8.78</b>

*Nonparametric tests:* We conducted four independent planned comparisons on SP vs. PS, one for each combination of Agreement Error and Plausibility, using the Wilcoxon test. These were conducted on the full set of results, since there is no way to calculate the interaction term with this nonparametric test. Neither comparison was significant for the NP1 condition (see Tables 3 and 4). Only the comparison between SP and PS BOTH, after an

agreement error (17.40 vs. 9.44), was significant ( $Z = 2.23$ ,  $p < .05$ ). The same comparison when no error was made not only did not show this pattern, but was marginally significant in the opposite direction (8.95 vs. 13.71,  $Z = 1.72$ ,  $p = .09$ ).

**Table 3.** Mean percentage of incorrect predicate selections (IP) after an agreement error

	<b>IP%</b>	<i>s.d.</i>
SP-NP1	<b>5.49</b>	<i>20.71</i>
PS-NP1	<b>3.33</b>	<i>18.10</i>
SP-BTH	<b>17.40</b>	<i>32.15</i>
PS-BTH	<b>9.44</b>	<i>26.64</i>

**Table 4.** Mean percentage of incorrect predicate selections (IP) after successful agreement

	<b>IP%</b>	<i>s.d.</i>
SP-NP1	<b>3.94</b>	<i>8.46</i>
PS-NP1	<b>3.85</b>	<i>8.34</i>
SP-BTH	<b>8.95</b>	<i>13.33</i>
PS-BTH	<b>13.71</b>	<i>15.65</i>

It is intriguing that SP caused large numbers of predicate errors, which are diagnostic of subject mis-selection, only after an agreement error. Why would participants not make such errors if they had correctly produced the agreement in the first part of the item? The fact that the difference not only disappeared, but was marginally significant in the opposite direction, indicates that inhibition may be involved. According to our model, when an SP item is successfully processed, the active plural feature on the nonsubject noun is inhibited, leading to a large decrease (from 17.40% to 8.95%) in the number of predicate selection errors compared to responses after an agreement error. On the other hand, the PS Both condition is not affected in the same way; correct agreement does not result in fewer errors when agreement is implemented correctly; in fact, there are more errors. But this too is readily explained by our model: after correct agreement, the local singular does not undergo inhibition, because there is no active number feature to inhibit.

We have already argued that the difficult SP conditions can be explained by the interaction of plural number and plausibility, both of which can lead to at least a superficial concord error. If an error is produced, we assume that the local noun gets further activation because it now controls agreement, and this convergence of factors causes a reanalysis whereby the local noun is thereafter treated as the subject, at least on some trials. Conversely, in the easier PS conditions, the local noun receives little activation because it is not a subject and has no number marking.

Another way to test our hypotheses -- one that does not suffer from the lack of multifactor nonparametric repeated measures analysis of variance procedures -- is to examine the correlations between the percentage of

agreement errors a participant made in the first part of the item and the percentage of predicate selection errors both after correct agreement and after an agreement error. We ran four separate correlations, one for each combination of Plausibility and Agreement type.

When only the first noun was plausible, there was no correlation between the number of original agreement errors and the number of adjective errors downstream. For both SP and PS, it seems that the clarity of the representation of the subject, because of its unique plausibility, meant that subsequent errors were limited to superficial concord errors.

However, when both nouns were plausible, we found that participants who made a large number of initial agreement errors also made a large number of adjective errors downstream. This correlation was significant for the percentage of adjective errors following an agreement error in both the SP and PS conditions (SP:  $R(60) = 0.47$ ,  $P < .0001$ ; PS:  $R(60) = 0.47$ ,  $P < .0001$ )

## 5. GENERAL DISCUSSION

In conjunction with the findings of Thornton and Macdonald (2003), our research has shown that agreement errors are influenced by semantic factors beyond the subject phrase. In addition, we have shown that when a local noun is also a plausible subject, people make more agreement errors on SP only, and also exhibit confusion about the subject in their downstream adjective choices. This led to productions such as the following:

(14) The boy beside the trees \*were tall ... and leafy.

These results conform to our prediction that activation of a local noun by both its plural number and its plausibility as a subject can lead the speaker to erroneously accept it as the subject for the duration of the sentence, even though it is not in subject position. In these cases we can say that the local plural both controls agreement and acts as a (revised) subject for the purposes of adjective modification. This means that the initial representation of a complex subject NP is not necessarily the final one. Processing breakdowns caused by the interaction of overt plural marking and compatibility with the first adjective can boost the activation of the local noun to the point where it controls both superficial agreement and semantic agreement.

Finally, what can our experiment tell us about everyday agreement processes? It is unlikely that in a non-experimental setting anyone would produce a sentence like (15), because the speaker controls the message plan, and does not intend to discuss a *varnished waiter*:

(15) The waiter by the tables \*are old and varnished.

In our study, the second predicate was introduced after the participant had already completed the first part of the sentence, so that the addition to the message plan was not internally generated. Although it could be argued that this

is not a natural production situation, there are in fact similar situations involving semantic inclusion relationships such as that between *occurrence* and *fires* in (16), where even speakers might shift subjecthood to the local noun in mid-sentence:

(16) The occurrence of the fires \*were unfortunate.

In such cases, both the fires and their occurrence are not only plausibly interpreted as unfortunate; they are co-extensive.

**NOTE:** Experimental items are available from the authors upon request.

### References

- Barker, J., Nicol, J., & Garrett, M. (2001). Semantic factors in the production of number agreement. *Journal of Psycholinguistic Research*, 30, 91-114.
- Bock, K. & Levelt, W. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 945-984). San Diego, California: Academic Press, Inc.
- Bock, K. & Miller, C. (1991). Broken agreement. *Cognitive Psychology*, 23, 45-93.
- Christianson, K., Hollingworth, A., Halliwell, J. F. & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology*, 42, 368-407.
- Eberhard, K. (1997). The marked effect of number on subject-verb agreement. *Journal of Memory and Language*, 36, 147-164.
- Francis, W. (1986). Proximity concord in English. *Journal of English Linguistics*, 19, 309-317.
- Greenslit, N. & Badecker, W. (2000). The role of the lexical item in the comprehension of subject-verb agreement. Poster presented at the 13<sup>th</sup> Annual CUNY Conference on Human Sentence Processing, San Diego, California.
- Levelt, W. (1989). *Speaking: From Intention to Articulation*. Cambridge, Massachusetts: MIT Press.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1972). *A Grammar of Contemporary English*. New York: Seminar press.
- Solomon, E. & Pearlmutter, N. (2004). Semantic integration and syntactic planning in language production. *Cognitive Psychology*, 49, 1-46.
- Thornton, R. & MacDonald, M. (2003). Plausibility and grammatical agreement. *Journal of memory and language* 48, 740-759.
- Vigliocco, G., Butterworth, B., & Garrett, M. (1996). Subject-verb agreement in Spanish and English: Differences in the role of conceptual constraints. *Cognition*, 61, 261-298.
- Vigliocco, G., Butterworth, B., & Semenza, C. (1995). Construction subject-verb agreement in speech: The role of semantic and morphological factors. *Journal of Memory and Language*, 34, 186-215.
- Vigliocco, G., Hartsuiker, R., Jarema, G. & Kolk, H. (1996). One or more labels on the bottles? Notional concord in Dutch and French. *Language and Cognitive Processes*, 11, 407-442.
- Vigliocco, G. & Nicol, J. (1998). Separating hierarchical relations and word order in language production: Is proximity concord syntactic or linear? *Cognition*, 68, 13-29.