Cognition 114 (2010) 151–164

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

# Experience and grammatical agreement: Statistical learning shapes number agreement production

Todd R. Haskell [a,*], Robert Thornton [b], Maryellen C. MacDonald [c]

[a] Western Washington University, Bellingham, WA, United States
[b] Pomona College, Claremont, CA, United States
[c] University of Wisconsin–Madison, Madison, WI, United States

ARTICLE INFO

*Article history:*
Received 2 October 2006
Revised 24 August 2009
Accepted 28 August 2009

*Keywords:*
Language production
Subject–verb agreement
Implicit learning

ABSTRACT

A robust result in research on the production of grammatical agreement is that speakers are more likely to produce an erroneous verb with phrases such as *the key to the cabinets*, with a singular noun followed by a plural one, than with phrases such as *the keys to the cabinet*, where a plural noun is followed by a singular. These asymmetries are thought to reflect core language production processes. Previous accounts have attributed error patterns to a syntactic number feature present on plurals but not singulars. An alternative approach is presented in which a process similar to structural priming contributes to the error asymmetry via speakers' past experiences with related agreement constructions. A corpus analysis and two agreement production studies test this account. The results suggest that agreement production is shaped by statistical learning from past language experience. Implications for accounts of agreement are discussed.



## 1. Introduction

Research in language production aims to explain how people can translate a thought into a spoken, signed, or written utterance. Much of this research has theorized about the architectural properties of the language production system that make possible this mapping from thought to language (e.g., Bock & Levelt, 1994). More recently, production researchers have begun to address the role of learning in the language production process. One such line of research involves the role of structural priming (or syntactic persistence), which refers to speakers' tendencies to use sentence structures that have been uttered or perceived in the recent past. The effects of structural priming are most clearly observable when a speaker has a choice between two alternative structures, e.g., a prepositional dative (*The author gave a book to the library*) or a double-object dative (*The author gave the library a book).* If the speaker has recently produced a prepositional dative sentence (*e.g., The man read a story to the boy*), that speaker is more likely to utter a prepositional dative than if a double-object dative (*e.g., The man read the boy a story*) has recently been produced (Bock, 1986). Although these effects are sometimes described in terms of short-term spreading activation (e.g., Pickering & Branigan, 1998), the effects can be relatively long-lasting, and it has been suggested that they are at least partly due to implicit learning of past utterances, possibly in addition to a transient activation-based mechanism (Bock & Griffin, 2000; Chang, Dell, Bock, & Griffin, 2000; Ferreira & Bock, 2006).

Another line of research to focus on learning mechanisms concerns the causes of speech errors in production. Dell, Reed, Adams, and Meyer (2000) noted that speech errors are typically constrained to follow the phonotactics of whatever language is being used. For example, an English speaker would be unlikely to produce a speech error such as [sil ŋo] for [siŋ lo] 'sing low,' because [ŋ] cannot occur in onset position in English. Dell et al. hypothesized that this phenomenon is the result of the implicit learning of sequential constraints. To test this hypothesis, they had participants recite lists of CVC syllables over the course

* Corresponding author.
*E-mail address:* todd.haskell@wwu.edu (T.R. Haskell).

0010-0277/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2009.08.017

of several days while varying the within-experiment phonotactics of the material that participants were to produce. They found that participants' speech errors not only obeyed the phonotactics of the language as a whole, but also obeyed the within-experiment constraints, demonstrating that utterance planning processes are affected by even very recent experience.

Findings such as these provide a potential avenue for linking theories of adult language production with theories of language acquisition. Indeed, Bock and Griffin (2000) raise the possibility that "structural priming is a dynamic vestige of the process of learning to perform language" (p. 189). On this view, the same learning mechanisms that help the child master speaking remain active in adults, serving the role of fine-tuning the production process.

Despite this shift to incorporate a greater role for learning in accounting for production behavior, there is at least one prominent area of research in language production – grammatical agreement – in which there has been relatively little discussion of the role of experience. This omission is interesting because it is clear that significant aspects of agreement must be learned. Although agreement is widely attested across languages (Mallinson & Blake, 1981), there is substantial cross-linguistic variation in exactly what elements must agree with each other, in which properties they must agree, and in the handling of various sorts of special cases (Barlow, 1992; Corbett, 1983, 1986). Even within a specific language such as English, there are dialectical differences in certain instances (e.g., American English *the government is* versus British English *the government are;* see Bock et al., 2006). Nevertheless, contemporary theories of agreement remain largely neutral on the role of learning and experience in adult agreement processing. These theories, such as Marking and Morphing (Eberhard, Cutting, & Bock, 2005), do not explicitly reject a learning component; rather they focus on architectural properties of the production system and do not mention what, if any, role that experience may have in shaping subsequent productions.

Our goal in the present work is to directly investigate the role of learning and experience in the production of agreement. We believe that including a role for experience in a theory of agreement production would serve two important functions. First, it would help connect work on agreement production with work in other areas of production such as structural priming and the learning of phonotactic constraints, thus moving the field toward a more unified and comprehensive account of language production in general. Second, attention to learning helps to forge connections between work on adult performance and work on language acquisition.

Our hypothesis that agreement processes are shaped by experience is supported by several recent studies. For example, Haskell and MacDonald (2005) examined number agreement with disjunctive noun phrases such as *the shirt or the socks*. English speakers in production experiments demonstrated an overall tendency toward agreement with the nearer noun, i.e., *The shirt or the socks are dirty* but *Is the shirt or the socks dirty?* Using corpus analyses, Haskell and MacDonald also found that overall in the language, the closest noun to the verb is typically the subject noun, suggesting a distributional basis for a proximity bias. However, the strength of the proximity bias varied with the particular ordering of the nouns and the verb. Interestingly, the probability of the subject noun being the closest noun to the verb also varied in a parallel fashion in the corpus. This suggests that the particular pattern of behavior seen with the disjunctive phrases reflects a sensitivity to the distributional patterns in the language, presumably learned through experience with the language. A similar argument was made by Thornton and MacDonald (2003) who found that speakers make agreement errors more often when the resulting utterance has agreement between semantically plausible noun–verb pairs than between unrelated pairs. Thornton and MacDonald hypothesized that producers' errors tend to be shaped by their experience with the distributional patterns in the language, in which semantically related noun–verb pairs often agree. This view links agreement errors to more traditionally-studied speech errors, which are well-known to be shaped by the phonotactic patterns of the language (see Dell et al., 2000, for review).

Franck, Vigliocco, Anton-Mendez, Collina, and Frauenfelder (2008) also proposed a role for distributional patterns in agreement, but this time cross-linguistically. They conducted several experiments examining gender agreement in Spanish, French, and Italian. Participants were given a phrase (e.g., *El castillo del pueblo)* and an adjective (e.g., *VIEJO*), and were asked to combine the phrase and the adjective to form a complete sentence. The phrases always contained a noun preceded by a definite article, and Franck et al. manipulated the morphophonology of the article and the noun. In all three languages, gender is usually marked on the definite article, e.g., Spanish uses the article *el* for masculine nouns and the article *la* for feminine nouns. However, the transparency of gender marking on nouns varies; it is highly transparent in Italian, where most masculine nouns end in *-o* and most feminine nouns end in *-a*. It somewhat less transparent in Spanish, and even less transparent in French.

Franck et al. found that the gender morphophonology of articles affected agreement in Spanish and French, but not Italian. They hypothesized that in Italian, gender marking on the noun is sufficiently consistent that the production system can utilize that marking alone, without the need to also consider the article. In Spanish and French, however, the marking on the noun can frequently be ambiguous or misleading, and so the production system utilizes information about both the article and the noun when computing agreement. Again, this differential sensitivity to different sources of information must arise through learning and experience. Franck et al. also explicitly suggest that the sensitivity to phonology is a vestige of the language acquisition process, when children use phonology to help learn gender classes.

These previous studies offer specific hypotheses concerning how certain linguistic experiences may affect agreement production. However, they are correlational in nature – a particular pattern of agreement behavior is observed, and this pattern is seen to parallel distributional patterns in the language. Unlike the work elsewhere in production, such as structural priming and the learning

of phonotactic constraints, no agreement studies have actually manipulated the language experience of the participants. In the current studies, we seek to provide more direct evidence regarding the role of experience in agreement by directly manipulating speakers' experience and then examining the effect of that experience on agreement production.

The typical paradigm for studying agreement production in the laboratory, originating with Bock and Miller (1991), involves presenting speakers with the beginning of a sentence, such as *The key to the cabinets...*, and asking them to generate an ending for this sentence. Some portion of the time, speakers will make an agreement error, e.g., *The key to the cabinets were missing*; the rate of such errors serves as the dependent variable. One frequently replicated finding is that participants produce more agreement errors when a number-mismatching noun intervenes between the head noun of a subject and its verb (e.g., *The key to the cabinets were...*), as opposed to the case where both nouns have the same number (*The key to the cabinet were...*) (Bock & Cutting, 1992; Bock & Miller, 1991; Bock, Nicol, & Cutting, 1999), presumably because the mismatching number of this intervening noun (conventionally referred to as the *local* noun) interferes with the agreement process. This interference effect varies in an interesting way that may shed light on the nature of agreement processes: In a number of studies, plural local nouns appear to cause many more agreement errors than do singular local nouns. Thus, *the key to the cabinets are...* is a much more frequent error than *the keys to the cabinet is...* (Bock & Cutting, 1992; Bock & Miller, 1991; Eberhard, 1997; Thornton & MacDonald, 2003; Vigliocco & Nicol, 1998). We will refer to this finding as *the agreement error asymmetry*. This asymmetry has been found in several different languages (Fayol, Largy, & Lemaire, 1994; Hartsuiker, Schriefers, Bock, & Kikstra, 2003; Vigliocco, Butterworth, & Garrett, 1996; Vigliocco, Butterworth, & Semenza, 1995).

Because the agreement error asymmetry is such a robust effect, it is a useful test case for examining the role of experience in agreement. In particular, we investigated whether the asymmetrical pattern of errors might result at least in part from an asymmetrical distributional pattern in the language, analogous to the case of agreement with disjunctions considered by Haskell and MacDonald (2005).

Other accounts of the asymmetry do not posit a role for experience. One explanation, offered by Bock and Eberhard (1993), concerns feature markedness. Their hypothesis was that, in the abstract syntactic representations that form the basis for computing agreement, plural nouns carry an explicit plural feature, but singular nouns are the unmarked default and usually carry no number feature. The consequence is that singular and plural nouns affect the agreement production process somewhat differently. In a phrase with a singular head and a plural local noun, such as *the key to the cabinets* (hereafter, singular–plural or SP phrases), both the head noun and the noun phrase (NP) as a whole would be unmarked. However, the local noun has explicit plural marking. Thus an agreement error can result if this explicit plural feature is erroneously copied or percolates up to the top level of the NP, overwriting the unmarked status. In NPs with a plural head and singular local noun, such as *the keys to the cabinet* (hereafter, plural–singular or PS phrases), the explicit plural feature on the head noun is passed up to the NP. The unmarked local noun has no explicit feature to be passed up, so there is nothing to overwrite the explicit marking. The ultimate result is that more errors are observed in the SP than PS condition.

Eberhard (1997) tested this account by providing explicit marking on singular nouns through the use of number-marked determiners such as *one*, as in *one key to the cabinets* and *the keys to one cabinet*. In these cases, both the singular and plural NPs have explicit number features. The markedness account predicts that because all of the NPs have explicit features available to percolate up the syntactic tree, the error asymmetry should be reduced considerably. In support of this prediction, Eberhard found that explicitly marking singular heads did reduce error rates in the SP condition, whereas explicitly marking singular local nouns increased error rates in the PS condition. These findings indicate that explicit number cues affect agreement error rates.

However, as Eberhard noted, the singular–plural error asymmetry still remained quite robust even with the explicitly marked determiners. In the unmarked singular conditions, consistent with previous results, over 35% of the responses in the SP condition were agreement errors, whereas only about 1% were in the PS condition-a 35:1 asymmetry. In the conditions in which the singulars were marked with *one* as a determiner, there was a reduced but still substantial asymmetry: About 24% in the SP condition versus 3% in the PS condition, an 8:1 asymmetry. Thus while determiner marking clearly has an effect on agreement production, a full account of the asymmetry must appeal to other factors.

One way to extend Eberhard's (1997) account to accommodate these facts is evident in the Marking and Morphing account (Eberhard et al., 2005). On this view, determiners and nouns make separate contributions to the overall number marking of the phrase, with the contribution of nouns being weighted much more heavily than that of determiners. Thus, the singular number marking achieved with a phrase such as *one key* is still much weaker than the plural number marking for *keys*. This approach can, in principle, explain why manipulations involving determiners attenuate but do not eliminate the error asymmetry. However, the implemented model presented by Eberhard et al. does not fully account for Eberhard's findings. For example, Eberhard found more agreement errors with phrases like *the keys to one cabinet* than with phrases like *the keys to the cabinet*, while the model presented by Eberhard et al. predicts no difference between these conditions (see their Fig. 12). In short, in making the effect of determiners weak enough to explain the relatively small effect of using quantifiers with singular head nouns, it becomes too weak to explain the effect of using quantifiers with singular local nouns.

Our own view is that there is clearly an effect of explicit marking of number on agreement processes, but that there is at least one important additional force in shaping the agreement error asymmetry. Given that an independently motivated implicit learning mechanism has been success-

fully applied to structural priming and the learning of phonotactic constraints, we hypothesize that it underlies aspects of agreement production as well, such that agreement production is also shaped by distributional patterns in speakers' experience. For the language-learning child, this shaping aids in the acquisition of "correct" agreement, as defined by the speech community in which the child lives. For example, in American English, the noun *government* is typically used with a singular verb, e.g., *the government is weak,* while in British English it is commonly used with a plural verb, e.g., *the government are weak* (see, e.g., Bock et al., 2006). An English-learning child must learn which verb form is appropriate to use in the immediate speech community. On our view, this learning takes place by means of adjustments to the language production system that increase the likelihood of producing structures that are well-attested in the child's experience, and decrease the likelihood of producing rare or non-existent structures.

Of course, such mechanisms would be of limited usefulness if the behavior of each lexical item had to be learned separately. In the case of the contrast between British and American English, the observations about *government* apply to many other collective nouns (e.g., *family, management)* as well as grammatically singular names of sports teams and musical groups. Ideally, the learning mechanism should be capable of generalizing the choice of verb form to other words and constructions that have similar properties, in the same way that structural priming effects can be found even between items that do not have identical syntactic properties (Bock & Loebell, 1990).

In this paper, we pursue the view that the SP error asymmetry is due at least in part to an asymmetry in speakers' experience with a similar construction, one in which singular head nouns are often grammatically paired with plural verbs. Consider collective phrases such as *a small coalition of special-interest groups have been pushing for change, a number of other analysts are recommending them, and a majority of people-oriented companies are likely to adopt it* (all grammatical examples taken from the Penn Treebank corpora, Marcus, Santorini, & Marcinkiewicz, 1993). Notice that each of these examples contains a grammatically singular head noun (*coalition, number*, and *majority*) followed by a prepositional phrase containing a plural local noun. In each case, the verb is plural rather than singular, thus seemingly agreeing with the plural local noun rather than the grammatically singular head noun. These cases are broadly similar to true agreement errors e.g. *the key to the cabinets were…*, with a singular head followed by a plural local noun and a plural verb. Although the underlying syntax may well be different in the collective examples and true agreement errors, the distributional pattern on the surface is very similar. In contrast, we are unaware of any constructions in which a plural noun routinely co-occurs with a singular verb in English. Our basic claim, essentially, is that experience with these collective constructions influences production of the closely related SP phrases and increases the likelihood that a plural verb will be used with SP phrases. We suggest that this process is analogous to the structural priming cases in which locative constructions prime passive sentences, even though the two share only a superficial syntactic relationship (Bock & Loebell, 1990). Such learning effects from similar but not identical constructions emerge in Chang's (2009) computational model of sentence production and form an important part of his account of different production preferences in English and Japanese.

If this implicit learning account of the agreement error asymmetry is correct, then it should be possible to manipulate the asymmetry by varying the participant's experience with agreement constructions, in the same way that other studies of priming or implicit learning have changed the patterns of participants' productions. We tested this hypothesis in several steps. First, we present corpus data on the rate of occurrence of singular and plural verbs with noun phrases of the type commonly used in agreement production studies. This analysis establishes that there is indeed an asymmetrical distributional pattern in the language. Next, we describe two experiments in which participants' short-term experience with collective phrases was directly manipulated, and the effect on participants' subsequent agreement production was measured.

## 2. Choice of verb number with complex NPs

If collective phrases are a source of implicit learning that results in the agreement error asymmetry, they must be frequent enough to plausibly affect the production of alternative forms. To evaluate this claim, we conducted an analysis on the Treebank version of the Brown corpus (Marcus et al., 1993). This corpus contains approximately 50,000 sentences and 1 million words of written American English text, tagged and parsed to indicate the structural relations between words in a sentence. We used the *tgrep* program to extract all sentences from this corpus in which an SS, SP, PS, or PP phrase was used with a number-marked verb, i.e., a present-tense verb or the past-tense forms *was* and *were*. For determining noun number, we used the tagging present in the corpus, in which collective nouns are generally coded as singular.

For each type of phrase, we then computed the proportion of all number-marked verbs that exhibited strict agreement with the head. For SS and SP phrases, this meant a singular verb; for PS and PP phrases, this meant a plural verb. These proportions, along with the raw counts, are presented in Table 1.

As can be seen in the table, there was a strong adherence to strict agreement for all phrase types except SP, where more than 20% of the phrases were used with a plural verb. Grouping together the other three phrase types for purpose of analysis, the proportion of strict agreement in the SP case was significantly different from the propor-

**Table 1**
Strict agreement by phrase type in the brown corpus.

| Phrase type | Proportion strict agreement |
|---|---|
| SS | 0.994 (1091/1098) |
| SP | 0.791 (321/406) |
| PS | 0.975 (354/363) |
| PP | 1.000 (191/191) |

tion of agreement with the head in the other three cases ($\chi^2$ (1, $N = 2058$) = 274, $p < 0.001$). It is these same SP phrases which consistently exhibit the highest rate of errors (that is, deviations from strict agreement) in agreement production studies. Inspection of the particular phrases used with plural verbs revealed that they were almost all of the collective type; examples include *a number of considerations suggest, a series of consequences begin,* and *a group of public-spirited Barbour County citizens have.*

These results suggest that English affords extensive experience with collective phrases containing plural verbs. Of course, the existence of this pattern does not guarantee that the pattern has a role shaping the agreement error asymmetry. It could be that some factor (such as markedness) is responsible for both the asymmetry in the corpus and the asymmetry in the agreement errors. To address this possibility, we directly manipulated speakers' experiences with collective phrases in Experiment 1. If the error asymmetry arises exclusively from principles of markedness, then such manipulations should have little effect. If, on the other had, the error asymmetry is in part caused by the distributional asymmetry, then manipulating the distributional information should have a measurable impact on agreement production in subsequent sentences.

## 3. Experiment 1

### 3.1. Method

#### 3.1.1. Participants

Sixty undergraduate students at the University of Wisconsin–Madison participated in the experiment. Participants received course credit in an introductory psychology class or were paid for their participation. All participants reported that they were native speakers of English.

#### 3.1.2. Materials and procedure

The experiment was administered as a web-based survey. Participants first read a short story about two teachers planning a field trip. Following the story were 33 sentence fragments that participants were instructed to complete in a way that was consistent with the story. Participants were led to believe that the experiment concerned how people interpret stories.

For simplicity of exposition, we will describe the structure of the experiment in terms of primes and targets, although we believe the mechanisms involved bear more resemblance to implicit learning than the transient changes in activation often associated with the term *priming*. Using this terminology, there were three prime conditions, one that was intended to promote plural agreement with collective phrases, one that was intended to promote singular agreement, and a neutral condition that served as a baseline. Prime condition was manipulated between participants.

Eight sentences in the story contained a collective phrase with a collective head noun as the subject. These sentences constituted the *story prime* sentences. In the singular-prime condition, these contained a singular verb, in the plural-prime condition, they contained a plural verb, and in the neutral condition, they contained a verb without number marking. To ensure that any observed effect was specific to agreement with collective phrases, rather than priming of a particular verb form or number, an additional eight sentences contained a non-collective subject followed by a verb opposite in number from the prime verbs (the *story counterbalance* sentences). Specifically, in the singular-prime condition these verbs were plural, and in the plural-prime condition these verbs were singular. In the neutral condition, equal numbers of singular and plural verbs were used; this was done (rather than using unmarked verbs) to avoid possible priming of the use of unmarked verbs in the completion task. Taking into account all sentences in the story, there were 24 singular, 9 plural, and 18 unmarked verbs in both the singular-prime and plural-prime conditions, with 20 singular, 5 plural, and 26 unmarked verbs in the neutral-prime condition. Example items are given in Table 2, and the complete story is presented in Appendix A.

The story was followed by 33 sentence fragments. The task of the participants was to type a completion to each fragment that fit with the story they had just read. Each sequence of four fragments constituted one completion "item", yielding eight such items, plus one final filler fragment. Each item contained a counterbalance fragment, a prime, a filler, and a target, in that order. The design of the prime and counterbalance fragments was analogous to that in the story, except sentence fragments rather than complete sentences were used. The filler fragment was used to minimize the effects of priming a particular verb number. For half the items the filler fragment had a singular verb, for the other half it had a plural verb. This approach was taken, rather than placing the counterbalance fragment between the prime and target, because the latter option would likely bias very strongly against the effect we sought. An example completion item is shown in Table 3. Overall, the fragments contained 13 singular and 13 plural verbs in both the singular-prime and plural-prime conditions, with nine singular and nine plural verbs in the neutral-prime condition. The complete set of fragments is presented in Appendix B.

No collective head nouns were repeated across the story primes, the completion primes and the targets. Although we might have achieved larger effects by using the same collectives in both the primes and the targets (Pickering

**Table 2**
Sample story prime and counterbalance sentences.

| | | |
|---|---|---|
| Prime | Singular | A trio of famous violinists was now scheduled for opening night |
| | Neutral | A trio of famous violinists had now been scheduled for opening night |
| | Plural | A trio of famous violinists were now scheduled for opening night |
| Counterbalance | Singular | These performances were certain to be spectacular |
| | Neutral | This event was certain to be spectacular |
| | Plural | This event was certain to be spectacular |

**Table 3**
Sample completion item.

| | | |
|---|---|---|
| Counterbalance | Singular | Some local papers were going to do a story on ... |
| | Neutral | A local paper was going to do a story on ... |
| | Plural | A local paper was going to do a story on ... |
| Prime | Singular | A caravan of buses was waiting to turn ... |
| | Neutral | A caravan of buses waited to turn ... |
| | Plural | A caravan of buses were waiting to turn ... |
| Filler | | The people at the ticket window were ... |
| Target | | The class of children ... |

& Branigan, 1998), the goal was to demonstrate that experience has a broad effect (i.e., it influences all similar constructions) rather than one limited to specific lexical items.

### 3.2. Results

All sentence completions for the target items were scored as singular verb, plural verb, or other, which included verbs that were not overtly marked for number (*played beautiful music, could not contain their excitement*) or where the response did not contain a verb. The proportion of responses in each category for the three conditions is shown in Fig. 1. The majority of completions in all conditions fell in the "other" category, reflecting the fact that most completions utilized past-tense verbs, most of which do not carry explicit number marking. However, there was also a clear increase in the number of plural completions in the plural-prime condition, relative to the singular-prime and neutral-prime conditions. That is, participants were more likely to complete a collective fragment with a plural verb if they had previously encountered collective fragments used with plural verbs. These effects were verified by planned comparisons between the singular-prime and neutral-prime conditions and between the neutral-prime and plural-prime conditions, using parallel *t*-tests by participants and by items. The singular-prime and neutral-prime conditions did not significantly differ ($t_1$ (38) < 1; $t_2$ (7) < 1). In contrast, significantly more plural verbs were produced in the plural-prime than the neutral-prime condition ($t_1$ (38) = 1.9, $p$ = 0.03, one-tailed; $t_2$ (7) = 4.9, $p$ = 0.001, one-tailed).

These differences were not due to variations in the rate of "other" responses across conditions. The rate of "other" responses did not differ between the singular-prime and neutral-prime conditions, nor between the neutral-prime and plural-prime conditions (all *t*'s < 1).

In addition to the three conditions shown in Fig. 1, we later ran a fourth condition in response to a reviewer's concern. The reviewer noted that in the fragment completion portion of the experiment, the prime fragment was closer to the target than was the counterbalance fragment (the fragment order always proceeded counterbalance, prime, filler, target, as shown in Table 3). This fragment order allowed the possibility that the increased rate of plural verbs in the plural-prime condition stemmed simply from priming of particular verb forms or a particular verb number (e.g., *were* in the plural-prime fragment primed the use of *were* in completing the target, while *was* in the counterbalance fragment, was a greater distance from the target). To ensure that lexical priming effects were not the source of the high rate of plural verbs in the plural-prime condition, we repeated this condition with the order of the prime and counterbalance fragments reversed, so that each target fragment was preceded, in order, by a plural-prime fragment, a counterbalance fragment, and a filler fragment. If the precise ordering of fragments contributed to lexical priming of target responses, then this reordering of the fragments to make the plural prime more distant from the target should remove or reduce the priming effect. Twenty-two native-English speaking undergraduates participated, completing the plural-prime condition exactly as in the original study except for the change in fragment

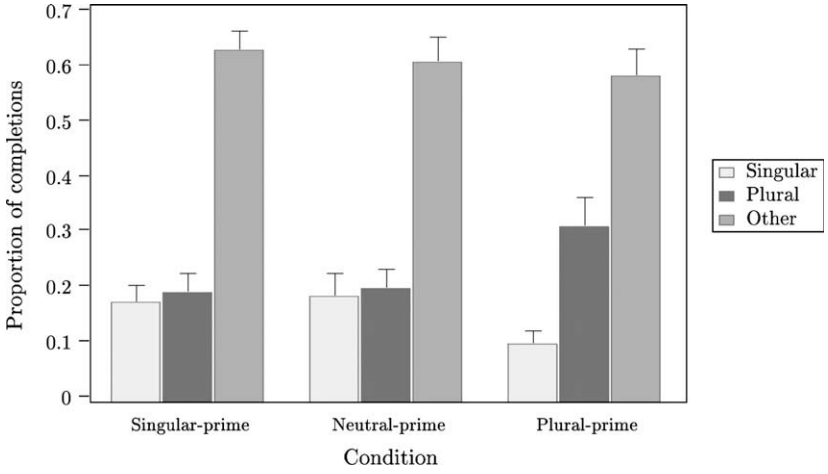


**Fig. 1.** Mean proportion of singular, plural and other completions in each condition of Experiment 1, with standard error bars computed across participants.

ordering. This group produced 37.5% plural responses, 9.1% singular responses, and 53.5% other responses. This rate of plural responses is numerically slightly larger than the 31.3% plural responses in the plural-prime condition in the original experiment. We therefore conclude that the effects observed in this study were not due to lexical priming or any other factors related to the exact ordering of prime and counterbalance fragments.

### 3.3. Discussion

The results of Experiment 1 show that encountering a plural verb used with a collective expression increased the likelihood of using a plural verb with other collective constructions in subsequent productions, even when the prime and target constructions contained different lexical items. This effect seems analogous to demonstrations of structural priming, where exposure to one construction increased production of that construction or a similar construction in subsequent utterances (e.g., Bock, 1986; Bock & Loebell, 1990). It should be noted that the effect was asymmetrical: Experience with singular nouns and plural verbs increased the likelihood of using plural verbs again with singular nouns, but experience with singular nouns with singular verbs had no measurable effect relative to the neutral condition. Further discussion of this issue will be deferred until after the results of Experiment 2 have been presented.

Although early studies of structural priming focused on cases where speakers produced both the prime and target sentences, later work has shown that priming also occurs if the speaker only hears the prime sentence, without having uttered it (Bock, Dell, Garnsey, Kramer, & Kubose, 2007; Branigan, Pickering, & Cleland, 2000 ; Hartsuiker, Pickering, & Veltkamp, 2004). The present findings provide further evidence for such comprehension-to-production influence. The existence of this kind of effect is significant because it suggests that the representations and mechanisms utilized in language production are at least partially shared with those used for comprehension (Branigan et al., 2000).

Our results also extend previous structural priming research addressing the longevity of priming and the effect of the spoken versus written medium on the strength of priming. Some studies have found structural priming effects to be robust over several intervening sentences (Bock & Griffin, 2000; Bock et al., 2007), while others have found that structural priming effects dissipate rapidly with sentences intervening between prime and target, especially in written presentations and productions (Branigan, Pickering, & Cleland, 1999).

Hartsuiker, Bernolet, Schoonbaert, Speybroeck, and Vanderelst (2008) compared structural priming effects in spoken and written production both with and without lexical overlap between prime and target. They found that while effects of lexical overlap dissipate rapidly, structural priming effects persist in both written and spoken modalities. They took these findings as evidence that structural priming effects were partially due to implicit learning processes. The results of Experiment 1, in which the effects of experience in the written modality persisted over intervening sentences, extend this account to agreement production.

Having established that the distributional pattern of agreement in a language user's input can influence subsequent agreement behavior in collective constructions, we can now broaden our attention to the agreement asymmetry more generally. To support the claim that implicit learning of agreement in collective constructions affects error rates in agreement more generally, it is necessary to show that the effects of experience extend not just from one collective construction to another, but from collective constructions to non-collective constructions. Thus Experiment 2 investigates whether encountering collective expressions like *a number of other analysts are recommending them* actually does increase the likelihood of producing true agreement errors such *as the key to the cabinets were*.

## 4. Experiment 2

### 4.1. Method

#### 4.1.1. Participants

Ninety-four undergraduate students at the University of Wisconsin–Madison participated in the experiment (30 in the singular-prime condition, 31 in the neutral-prime condition, and 33 in the plural-prime condition). Participants received course credit in an introductory psychology class or were paid for their participation. All participants reported that they were native speakers of English.

#### 4.1.2. Materials and procedure

The materials and procedure were exactly the same as in Experiment 1, with the exception of the target fragments. Instead of containing a collective noun followed by a prepositional phrase (*The class of children...*), fragments contained a non-collective head noun, followed by two prepositional phrases (e.g., *The pencil in the gift bags for the organizers...*). Two prepositional phrases were used, instead of a single phrase as in *the key to the cabinets*, to increase the difficulty of the agreement computation and ensure that participants would make a measurable number of errors. This choice also makes the structure of the target phrases more distinct from the structure of the primes, thus providing a stronger test of the generality of the effects. The complete set of target fragments is presented in Appendix B. Scoring was as in Experiment 1.

### 4.2. Results

The proportion of responses in each category for the three conditions is shown in Fig. 2. In comparison with the results of Experiment 1, singular verbs were much more common overall in the current experiment. This is not surprising, as the constructions used as target fragments required the use of a singular verb. However, as in Experiment 1, there was a clear increase in the number of plural completions in the plural-prime condition, relative to the singular-prime and neutral-prime conditions. These effects were verified by planned comparisons between the singular-prime and neutral-prime conditions

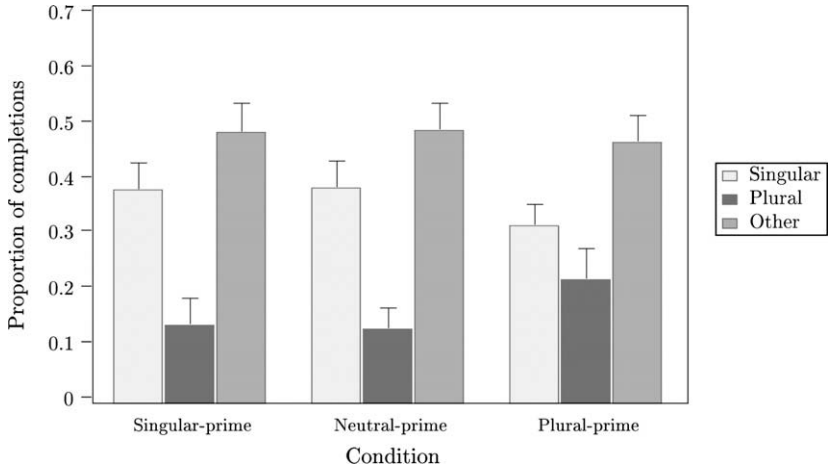


**Fig. 2.** Mean proportion of singular, plural and other completions in each condition of Experiment 2, with standard error bars computed across participants.

and between the neutral-prime and plural-prime conditions, using parallel *t*-tests by participants and by items. The singular-prime and neutral-prime conditions did not differ ($t_1$ (59) < 1; $t_2$ (7) < 1). In contrast, significantly more plural verbs were produced in the plural-prime condition than in the neutral-prime condition ($t_1$ (62) = 1.8, $p < 0.05$, one-tailed; $t_2$ (7) = 2.6, $p < 0.05$, one-tailed).

Again, these differences were not due to variations in the rate of "other" responses across conditions. The rate of "other" responses did not differ between the singular-prime and neutral-prime conditions, nor between the neutral-prime and plural-prime conditions (all *t*'s < 1).

### 4.3. Comparisons across experiments

Given that the two experiments were identical except for the target sentence fragments that participants completed, we conducted an additional analysis across both data sets, with experiment as a between-participants and within-items factor. This analysis allowed us to compare the effect of experience in the two studies. The primes and targets were much more similar in Experiment 1 than in Experiment 2, and so it is interesting to ask whether the effect of varying prime type was larger in Experiment 1. Such a difference would be evident in an interaction between experiment and prime type. Some studies have found increases in priming of noun phrase structure when there is semantic overlap in the prime and target noun phrases (Cleland & Pickering, 2003). As the repetition of collective terms from prime to target in Experiment 1 entails some degree of semantic overlap, there might be both semantic and syntactic (more similar structures) factors that could yield larger effects in Experiment 1 than in Experiment 2.

The analysis showed a marginal main effect of experiment, such that there were overall more plural responses with collective target fragments in Experiment 1 than with the non-collective targets in Experiment 2. This effect was significant by participants but not by items ($F_1$ (1, 144) = 6.0, $p = 0.02$; $F_2$ (1, 7) = 1.2, n.s.). This result is not surprising, since the use of plural verbs is generally considered acceptable with collective phrases but is an error with the non-collective phrases. There was also a main effect of prime type ($F_1$ (2, 144) = 3.9, $p = 0.02$; $F_2$ (2, 14) = 7.3, $p < 0.01$), reflecting the greater use of plural verbs in the plural-prime condition (25%) than in the other two conditions (both 16%), consistent with the individual experiment analyses. With respect to the question of whether the effect of prime type was larger in Experiment 1 than in Experiment 2, the interaction between prime type and Experiment was not reliable (both *F*'s < 1). This result suggests that the collective fragments with plural verbs increased plural verb usages in both new collective expressions and new non-collective expressions to about the same degree.

### 4.4. Discussion

As in Experiment 1, exposure to the use of collectives with plural verbs increased the likelihood of erroneously using plural verbs with target fragments such as *The pencil in the gift bags for the organizers…*. Thus, the effect demonstrated in Experiment 1 does not depend on the prime and target having exactly the same structure, but extends to cases where the prime and target have somewhat different structures. This result supports the claim that experience with collective phrases could generalize to other similar phrases, and thus contribute to the well-documented agreement error asymmetry.

Experiment 2 also replicated the asymmetry found in Experiment 1, such that the plural-prime condition yielded a reliable increase in plural verbs over the neutral baseline, while the singular condition did not increase singular verbs over baseline. This asymmetry in the effects of singular versus plural primes likely reflects the relative frequency of the alternative forms, in that singular verbs are the overwhelming choice for agreement with singular head nouns in the language, and exposure to a few more sentences consistent with this pattern is unlikely to have much effect. Similar frequency-based priming asymmetries can be seen in several domains. For example, Hartsuiker and Westenberg (2000) examined structural priming

in Dutch subordinate clauses. Dutch allows subordinate clauses to have a participle-final or an auxiliary-final word order; a baseline measure showed the participle-final order to be more common. Hartsuiker and Westenberg found that participle-final primes had little tendency to increase the rate of the already-common participle-final responses, whereas auxiliary-final primes substantially increased the rate of auxiliary-final responses. Complementary results have been found in structural priming studies in which speakers produced either active or passive sentences. Passive primes have been shown to be effective in increasing passive sentences, which are fairly rare constructions in the language as a whole, while active primes have a weaker and less consistent effect on production of the already-common active construction (e.g., Bock, 1986; Hartsuiker et al., 2004). Similarly, Ferreira (2003) found more robust priming for production of a low frequency form of sentential complement structure (the form that omits the complementizer *that*) than for the higher frequency form.

The effect can also be seen outside of language production; for example, studies going at least back to Becker (1979) have found greater priming for low frequency words than for high frequency words in lexical decision tasks. The asymmetry in the results from Experiments 1 and 2 is therefore consistent with previous findings. Furthermore, connectionist networks, which have been used to model structural priming as emergent from implicit learning (Chang et al., 2000), exhibit exactly this sort of behavior.

## 5. General discussion

In the introduction, we noted a parallel between patterns in agreement error rates on the one hand, and distributional patterns in the English language on the other. Specifically, SP phrases (*the key to the cabinets*) are far more likely to elicit number agreement errors than PS phrases (*the keys to the cabinet*). It also happens that there are certain constructions in the English language in which singular heads appear with plural verbs. The opposite pattern, constructions systematically pairing plural heads with singular verbs, is not attested. The central claim of this article is that this asymmetry in the language helps cause the asymmetry in the agreement errors. On this view, language users have encoded this asymmetry via statistical learning over extensive experience with the language, and the distributional regularities shape agreement production in subsequent utterances.

In our experience-based account, the higher error rate typically observed with SP phrases is a result of a less consistent mapping between singular heads and singular verbs, as compared with plural heads and plural verbs. Because singular heads sometimes correctly appear with plural verbs, this type of error is, in a sense, less bad and consequently more likely to be produced. This finding parallels the results of Dell et al. (2000), who found that the distributional phonotactics of words presented in an experiment affected the types of speech errors participants made. We are not claiming that implicit learning of distributional patterns of collective expressions is the only factor responsible for the existence of the agreement error asymmetry. Rather, this distributional information is one of several cues that affects the choice of what verb form to produce (Haskell & MacDonald, 2005). In previous work we have argued that some of the other cues include number semantics, morphophonology, and verb–noun plausibility relationships, some of which may themselves be subject to statistical learning as well (Haskell & MacDonald, 2003; Thornton & MacDonald, 2003). In this framework, the effect of a number-marked determiner such as *one* on agreement production is viewed as part of a larger constellation of factors that modulate competition between alternative forms.

In the remainder of this paper, we address four additional issues. First, we offer some thoughts on the agreement elicitation methods we used here. Second, we discuss the implications of our work for theories of agreement production. This is followed by discussion of two important questions for future research: Whether some other constructions may also contribute to the error asymmetry in English, and whether our arguments could be extended to other languages in which the agreement error asymmetry has been demonstrated (Fayol et al., 1994; Hartsuiker et al., 2003 ; Vigliocco et al., 1995, 1996).

### 5.1. Methodological limitations

We are confident that our data support the central claim of this article – namely, that agreement behavior is systematically affected by experience. However, there are also certain limitations related to the agreement error methodology and stimuli we used in our experiments. We note a few issues for those who might want to pursue experience manipulations in future studies.

First, as Haskell and MacDonald (2003) noted, the use of proportions as a dependent measure creates the possibility of floor effects. Specifically, as the overall proportion of plural verbs becomes smaller, the more difficult it is to observe differences between conditions. Since the overall proportion of plural verbs is generally higher with collective phrases (as in Experiment 1) than with SPP phrases (as in Experiment 2), floor effects are more likely to be a problem with the latter type of phrase. Indeed, some of our attempts to demonstrate priming with SPP phrases were hampered by low error rates, around half of what is reported above for Experiment 2.

A second issue also relates to the use of SPP phrases. We chose to use this type of phrase because we believed that the complexity of such phrases would result in a higher rate of errors than if we used the simpler SP phrases that are more familiar in agreement research. However, in some cases this complexity may have discouraged some participants from fully attending to the target fragments. In one version of an experiment with SPP phrases we halted data collection half way through when the error rate with the neutral primes reached 28% and the nature of the completions indicated that at least some participants were not reading the stories or fragments very carefully.

A third issue involves the use of written responses, especially typewritten responses. Typing provides participants an opportunity to edit and correct their responses

before they are submitted. This could conceivably reduce the overall rate of errors in the data. It is unknown whether these corrections would affect all conditions equally.

Every error elicitation experiment requires a careful balancing of task and materials such that participants make a reasonable number of errors without finding the task impossibly difficult. We found that balance for these studies, but ultimately, the generality of our results will benefit from additional experiments utilizing a range of modalities and phrase types.

### 5.2. Implications for theories of agreement

The present work attempts to bring together two different strands of research on language production. Work on phenomena such as structural priming and the acquisition of phonotactic constraints has considered how past experience with language shapes subsequent performance. This research has led to proposals for a language production system that learns and adapts with each new utterance it encounters (Bock & Griffin, 2000; Chang, 2009; Chang et al., 2000; Dell et al., 2000). Work on agreement production has largely focused on identifying the various factors that influence the agreement process and devising an architecture that can account for the behavioral data (e.g., Eberhard et al., 2005). In combining these two threads, we have proposed an account for certain agreement phenomena that is motivated by findings in other production domains. We believe that this is a major strength of the current approach.

We conceive of the effects of experience on agreement as one part of a broader constraint satisfaction account of agreement (Haskell & MacDonald, 2003, 2005; Thornton & MacDonald, 2003). On this view, the verb form used with a particular subject is determined by the interaction of multiple graded, probabilistic constraints, including past experience with similar constructions. The mechanisms of implicit learning that we believe are important in the current case have also been appealed to account for other agreement phenomena within a constraint satisfaction framework (Haskell & MacDonald, 2005).

Building on the account outlined in Haskell and MacDonald (2003), we hypothesize that the production of agreeing forms is a lexical choice in which alternative agreeing forms (e.g., *eat* versus *eats*) are partially activated and compete for selection. Competition processes of this sort are already familiar in other areas of language production, and our suggestion here is a first step in extending them to subject–verb agreement. First, in languages in which determiners agree with nouns in gender and number, a number of studies have demonstrated the existence of competition processes in the production of determiners (Alario & Caramazza, 2002; Janssen & Caramazza, 2003; Miozzo & Caramazza, 1999; Schriefers, 1993). Most researchers have assumed that the locus of competition here is at the level of phonological form rather than at an abstract gender representation level. Importantly, Spalek and Schriefers (2005) showed that the nature of this competition was modulated by past frequency of usage, such as the relative frequency of the singular versus plural form of the noun. Second, other studies have observed competition effects between synonyms (Bock & Levelt, 1994; Peterson & Savoy, 1998). Synonyms, such as *sofa* and *couch*, tend to vary in their use in different discourse contexts. For example, a Google search revealed that in English phrases referring to a sofa or couch as a possession, examples of *couch* outnumber examples of *sofa* only slightly, at a ratio of about 5:4 (as in *buying a couch, owning a sofa*, etc.). However, in references to a surface for reclining (*sleeping on the couch, lying on the sofa*, etc.), *couch* strongly outnumbers *sofa* by about 3:1.[1] Thus certain messages favor one synonym more than another. We suggest that inflectional variants like *bark/barks* and *is/are* are activated and compete in ways analogous to those observed for synonyms and determiner forms, modulated by both the frequency of previous productions and the current message context. Thus in the environment of a message about dogs (plural) participating in a barking event, the plural verb *bark* will have an advantage over the singular form *barks*. This state is hypothesized to owe to experience, so that the competition between *is* and *are* in the context of *government* will be different for American and British English speakers, as a function of their different experiences with terms like *government, management*, etc.

This view of inflectional morphology contrasts with the more traditional position that inflection is represented by abstract number, gender and other morphological nodes (e.g., Levelt, Roelofs, & Meyer, 1999). In this alternative framework, lexical properties such as number and gender emerge from the process of learning the mapping between semantics and phonology (see, e.g., Mirkovic & MacDonald, 2003). Clearly these two very different conceptions of morphological representation and lexical production have extensive implications for research that are well beyond the scope of this paper, and the debate will not be resolved here. Our more immediate claim is that a reconceptualization of morphological representations along these lines can help offer an account of the agreement error asymmetry.

Given these observations, our results suggest two possible reasons why Eberhard's (1997) manipulations of number marking failed to equalize the error rates in the SP and PS conditions. One possibility is that there are two factors that are jointly responsible for the agreement error asymmetry, markedness and distributional patterns in the language. Consequently, even when singular and plural number are both explicitly marked, other distributional patterns in the language continue to promote the use of plural verbs with singular nouns, so that a substantial portion of the error asymmetry remains. On this view, it might be possible to accommodate our findings within a theory like Marking and Morphing by incorporating an explicit role for experience into the theory.

For example, the computational model presented by Eberhard et al. has a number of free parameters, the appropriate values of which were determined by fitting the model to behavioral data. Perhaps speakers also engage in this type of model-fitting, continuously fine-tuning the param-

[1] The reported statistics reflect means over a number of Google searches of exact gerund phrases such as *purchasing a sofa*. Tense changes (*will purchase, purchased*) and adjective additions (e.g. *new sofa*) did not change these proportions.

eters of the system on the basis of the utterances the speaker encounters. On this view, the effects we observed would result from such small-scale adjustments in the system. What we are suggesting here is the need for a learning component over distributional regularities in agreement processing, and the adjustments in agreement behavior as a function of this learning could be implemented in many different computational frameworks, including symbolic ones as in Eberhard et al. (2005).

Turning to the second possibility, traditionally it is often assumed that markedness leads to certain distributional regularities in languages. However, it has also been argued that distributional patterns could lead to markedness (e.g., Haspelmath, 2006), with markedness being an emergent property. This approach is similar to proposals in inflectional morphology that a "default" form for the plural or past tense could emerge from the distributional properties of the language (e.g., Hare, Elman, & Daugherty, 1995). On this view, rather than markedness and distributional patterns being separate factors affecting agreement, they would be two aspects of the same phenomenon.

Haskell, Mansfield, and Brewer (submitted for publication) have begun to explore the idea of markedness as an emergent property in a series of artificial language learning studies, examining the extent to which factors of frequency and variability of gender-like categories lead to treating one category as the unmarked default. On this view, the same sort of learning mechanisms that we have appealed to here would also underlie learning to treat the singular category as the unmarked default in English. Our current findings do not distinguish between these possibilities, and this will clearly be an important topic for future research.

### 5.3. Beyond collective phrases

The focus of the current work has been on the effect of experience with collective constructions. However, there are several other constructions in English that could also promote the use of plural or plural-sounding verbs with singular nouns. One such construction is conjunctions, such as *Susan, Malcolm, and Javier are cooking dinner* and *Love and hatred were one*. In most cases, conjoined noun phrases require a plural verb, regardless of the number of the individual nouns. Both the current Experiment 2 and prior work (Bock & Loebell, 1990; Chang, 2009) suggest that the effects of experience do not require exact structural or lexical matches between prior experience (the prime) and subsequent behavior (the nature of the target). Thus it is at least possible that experience with plural agreement with conjunctions could promote plural agreement with non-conjoined complex noun phrases, much as plural agreement with collective phrases yielded plural agreement with non-collective phrases in Experiment 2.

Additionally, singular nouns often occur with verbs that are grammatically singular but homophonous with a plural form in English. For example, the verb form *love* is used not only with plural subjects (e.g., *they love*), but also with 1st and 2nd person singular subjects (*I love, you love*). Similarly, the verb form *were* is not only used with plural subjects in the past tense, but is also the singular subjunctive form, e.g., *if I were to love you...* In our framework, this would mean that in the presence of a singular subject, both "love" and "loves" would become partially activated because both phonological forms have been used with singular subjects in the past, while in the presence of a plural subject only "love" would become activated. In other words, encountering a sentence with a subjunctive verb such as *If the key to the cabinets were missing, what would we do?* might increase the likelihood of saying *The key to the cabinets were missing* in non-subjunctive contexts.

Franck, Vigliocco, and Nicol (2002) put forward a related proposal to account for different patterns of agreement errors in French and English. They noted that in their experiments, errors with constructions requiring a plural verb were relatively more common in French than in English. Franck et al. attributed this pattern to differences in the morphophonology of English and French verbs, specifically the fact that plural verbs are morphophonologically more complex than singular verbs in French, while the reverse is true in English. In their view, greater complexity is likely to lead to more errors. Our suggestion regarding homophony is essentially analogous, but focuses on the complexity of the mapping between syntax and form, rather than the complexity of form itself. In English, the mapping between singular number and verb forms is one-to-many, whereas the mapping between plural number and verb forms is one-to-one. Following the same reasoning as Franck et al., one would expect this complexity asymmetry to lead to more errors when a singular verb is required than when a plural verb is required.

The effects of homophony in verb forms are not yet known, but the plausibility of such a hypothesis is increased by evidence for effects of homophony in the case of nouns. English has a number of nouns for which the singular and plural forms are the same, such as *fish* and *sheep*. In isolation, such nouns have ambiguous number. There is evidence that surface ambiguity of this sort can affect agreement processing. Vigliocco et al. (1995), working in Italian, compared rates of agreement errors for invariant plurals and ordinary plurals. In the experiment, the nouns were presented along with a number-marked determiner, so that the number of the noun was never truly ambiguous (e.g., *la città* 'the town' versus *le città* 'the towns'). Nevertheless, more errors occurred with the invariant nouns than with ordinary nouns. Similarly, Vigliocco and Zilli (1999) found that gender agreement errors in Italian were more common when a noun carried ambiguous gender marking than when it carried clear gender marking. Hartsuiker et al. (2003), working in German and Dutch, manipulated whether the determiners used with head nouns were ambiguous or unambiguous for number. More agreement errors occurred with the ambiguous than the unambiguous determiners. In light of these findings, the extent to which homophony and morphophonological complexity influence agreement will likely be an interesting topic for future research.

### 5.4. The agreement asymmetry in other languages

On our account, a major reason that speakers produce a large number of agreement errors with SP phrases is that

there is a distributional asymmetry in the language that promotes such errors. In making this argument, we have focused on the specific distributional properties of English. Of course, the number of constructions that contribute to such an asymmetry, and crucially their frequency of use in the language, can vary across languages. Thus, our account makes interesting cross-linguistic predictions.

For example, all else being equal, languages in which singular nouns appear less frequently with plural verbs than in English should exhibit a reduced error asymmetry, while languages in which singular nouns appear more frequently with plural verbs than in English should exhibit a stronger error asymmetry. If there were a language in which plural nouns regularly appeared with singular verbs, it might even exhibit a reverse asymmetry, depending on the relative contributions of markedness and distributional factors.

Testing these predictions will be an interesting but challenging avenue for future work. It will be necessary to identify the relevant constructions in a given language, which could be the same structures we have identified for English, or a completely different set, and then obtain information on the frequency with which these constructions actually occur. A language could have a large number of relevant structures, but if they are rarely used, they might have little impact. Conversely, a language could have a small number of relevant structures, but if used frequently enough, they might have a larger impact. Moreover, exactly what counts as a relevant structure is not known. For example, we have suggested that noun conjunctions, first and second person forms and subjunctive forms may contribute to the error asymmetry in English, but there is as of yet no evidence supporting this possibility.

In addition, to our knowledge there has been no systematic examination of the relative strength of the error asymmetry across languages (but see Franck et al., 2002). Obtaining such data is made more difficult by the fact that the size of the error asymmetry depends in part on the overall rate of errors. Anecdotally, it appears that agreement errors in fragment completion tasks are much less common in highly inflected languages such as Serbian, which could make detection of any error asymmetry more difficult (J. Mirković, personal communication).

A final part of the difficulty in comparing the asymmetry cross-linguistically is that even within a language, the size of the asymmetry will vary with the stimulus set used to test it. For example, it has been shown that plausibility information significantly mediates the magnitude of the asymmetry in English (Thornton & MacDonald, 2003). As discussed earlier, number ambiguity of nouns and determiners has been shown to affect error rates in several languages (Hartsuiker et al., 2003; Vigliocco et al., 1995). Thus, cross-linguistic differences might be observed across studies due to differences in stimuli rather than a real language-wide difference in the magnitude of the asymmetry.

Despite these challenges, the ultimate success of our approach will depend on how well it generalizes to languages other than English. The most fruitful approach may well be to manipulate distributional information within the context of an experiment itself, as was done in the present experiments.

## Acknowledgements

This research was supported by NSF Grant SBR-9511270 to the University of Southern California, NIH Grant HD047425 to the University of Wisconsin–Madison, and NIH Grant AG025195 to Pomona College. The authors would like to thank Kazunaga Matsuki, Leslie May, and Gary Oppenheim for assistance in running the experiments, and Jelena Mirković for helpful comments on an earlier version of this manuscript.

## Appendix A. Story used in the experiments

Material that varied across conditions is shown in parentheses, in the order singular-prime condition/neutral-prime condition/plural-prime condition.

### *A.1. Story*

Eleanor Sarno is the enthusiastic and talented orchestra director at Westfield High, and she is always thinking of how to inspire her students. Eleanor knows that the (teachers are/principal is/principal is) behind her in her efforts, so she really tries to be creative. When she sat down to read the newspaper this morning, she felt that some of (those creative juices were/those creative juices were/that creative energy was) flowing. A front page article reported several new additions to the local music festival. A trio of famous violinists (was now/had now been/were now) scheduled for opening night. Eleanor was excited because her current crop of violin students (was/seemed/were) very promising, and they would love to hear this concert. She also read that the well-known soprano Dominique LaBelle was making an appearance. (These performances were/This event was/This event was) certain to be spectacular. Eleanor picked up the phone to call her colleague the choral director.

"Carol!" Eleanor said excitedly when the choral teacher answered, "A gadzook of really great musicians (is/will be/are) coming to the music festival! I have got to organize a field trip for our students!"

"I know!" said Carol, "I'll help. I just heard that Dominique LaBelle and Hector Rodriguez are giving a special concert and workshop for vocal students!"

"That's great", said Eleanor, "who is Hector Rodriguez?"

"Oh, I adore him", said the choral director. "He's an up-and-coming tenor who is really into music education, not just performing. A whole constellation of critics (thinks/wrote/think) that he's going to be the next Placido Domingo, but they worry that his (talents aren't/talents aren't/talent isn't) getting enough attention. A committee of opera directors (was/had been/were) saying last year that he spent too much time on education

stuff and not enough time performing. (Those comments were/Those comments were/That sentiment was) not well-received in the voice education community."

"Wow", said the orchestra director, "A group of fabulous violinists (is/will be/are) coming too. I wonder if (they are going to be doing/there is going to be/there is going to be) a violin workshop, too. This is so exciting!"

"Yes", Carol replied. "But Eleanor, if (we don't/someone doesn't/someone doesn't) get moving fast, there won't be any field trip. With LaBelle, Rodriguez and those violinists coming, an army of fans (is going to be/will be/are going to be) storming the ticket windows. (Aren't tickets going on sale today?/Aren't tickets going on sale today?/Isn't today the day tickets went on sale?) A huge portion of the tickets (is probably/must be/are probably) gone already!"

"OK, OK, Carol, I'm on it! I'll give you a call back later when I find out more". The orchestra director hung up the phone and immediately started writing down ideas.

## Appendix B. Sentence fragments used in the experiments

*CB = counterbalance, P = prime, F = filler, T1 = Experiment 1 target, T2 = Experiment 2 target*

| | | |
|---|---|---|
| (1) | CB | Although many of Eleanor's students play the violin, she also has (some students who play/one student who plays/one student who plays) the ... |
| (2) | P | After getting off the phone with Carol, a jumble of wonderful possibilities (was already/had already started/were already) racing through Eleanor's ... |
| (3) | F | Eleanor's students were going to meet for class ... |
| (4) | T1 | Every year, a series of concerts ... |
| (4) | T2 | The schedule in the envelopes on the students' seats ... |
| (5) | CB | (Some local papers were/A local paper was/A local paper was) going to do a story on... |
| (6) | P | A caravan of buses (was waiting/waited/were waiting) to turn ... |
| (7) | F | The people at the ticket window were... |
| (8) | T1 | The class of children ... |
| (8) | T2 | The pencil in the gift bags for the organizers ... |
| (9) | CB | During the long bus ride to the music festival, the (singers were/singers were/bus driver was) ... |
| (10) | P | A fleet of limousines (was/had been/were) parked outside the ... |
| (11) | F | A parent accompanying the students on the field trip was ... |
| (12) | T1 | Near the gate a bunch of police officers ... |
| (12) | T2 | The address on the programs at the ticket windows ... |
| (13) | CB | (Volunteers were/Volunteers were/A volunteer was) helping the concert-goers to ... |
| (14) | P | A wide variety of snacks (was available for purchase/could be purchased/were available for purchase) at the ... |
| (15) | F | A local radio station was ... |
| (16) | T1 | The audience in the crowded stands ... |
| (16) | T2 | The design on the banners above the speakers ... |
| (17) | CB | Because (tickets for the concert were/tickets for the concert were/admission to the concert was) quite expensive, the students' seats ... |
| (18) | P | A cluster of reporters (was gathering/had gathered/were gathering) at the entrance to the ... |
| (19) | F | Apparently LaBelle was about to ... |
| (20) | T1 | The mob of spectators ... |
| (20) | T2 | The stamp on the tickets to the VIP boxes ... |
| (21) | CB | (Eleanor and Carol were/Eleanor was/Eleanor was) surprised that the orchestra ... |
| (22) | P | A panel of music experts (was going to/would/were going to) be giving an award for the best ... |
| (23) | F | After Hector Rodriguez performed in the workshop, Carol's choral students were ... |
| (24) | T1 | Above the stage, a flock of birds ... |
| (24) | T2 | The label on the box lunches for the judges ... |
| (25) | CB | (Some opera singers who Carol had known in college/An opera singer who Carol had known in college was/An opera singer who Carol had known in college was) sitting ... |
| (26) | P | Carol's entourage of budding virtuosos (was impatient/could not wait/were impatient) for ... |
| (27) | F | All afternoon dark clouds were passing overhead, but fortunately ... |
| (28) | T1 | The row of violinists ... |
| (28) | T2 | The logo on the backs of the chairs ... |
| (29) | CB | The (organizers of the concert were/organizers of the concert were/organizer of the concert was) happy that ... |
| (30) | P | A collection of autographed pictures (was/had been/were) displayed near the ... |
| (31) | F | The music that was performed during the workshop ... |
| (32) | T1 | For the final act, a whole family of trumpet players ... |
| (32) | T2 | The warning on the signs by the emergency exits ... |
| (33) | F | At the end of the day, Carol and Eleanor ... |

## References


Alario, F.-X., & Caramazza, A. (2002). The production of determiners: Evidence from French. *Cognition, 82*, 179–223.

Barlow, M. (1992). *A situated theory of agreement*. New York: Garland.

Becker, C. A. (1979). Semantic context and word frequency effects in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 5*, 252–259.

Bock, K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18*, 355–387.

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

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

Bock, K., Dell, G. S., Garnsey, S., Kramer, A., & Kubose, T. T. (2007). Car talk, car listen. In A. S. Meyer, L. R. Wheeldon, & A. Krott (Eds.), *Automaticity and control in language processing. Advances in behavioural brain science* (pp. 21–42). New York, NY: Psychology Press.

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

Bock, 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., & Levelt, W. (1994). Language production: Grammatical encoding. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics*. San Diego, CA: Academic Press.

Bock, K., & Loebell, H. (1990). Framing sentences. *Cognition, 35*, 1–39.

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.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (1999). Syntactic priming in written production: Evidence for rapid decay. *Psychonomic Bulletin and Review, 6*, 635–640.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition, 75*, B13–B25.

Chang, F. (2009). Learning to order words: A connectionist model of heavy NP shift and accessibility effects in Japanese and English. *Journal of Memory and Language, 61*, 374–397.

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.

Cleland, A. A., & Pickering, M. J. (2003). The use of lexical and syntactic information in language production: Evidence from the priming of noun–phrase structure. *Journal of Memory and Language, 49*, 214–230.

Corbett, G. G. (1983). Resolution rules: Agreement in person, number, and gender. In G. Gazdar, E. Klein, & G. K. Pullum (Eds.), *Order concord and constituency*. Dordrecht, The Netherlands: Foris Publications.

Corbett, G. G. (1986). Agreement: A partial specification, based on Slavonic data. *Linguistics, 24*, 995–1023.

Dell, G. S., Reed, K. D., Adams, D. R., & Meyer, A. S. (2000). Speech errors, phonotactic constraints, and implicit learning: A study of the role of experience in language production. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26*, 1355–1367.

Eberhard, K. M. (1997). The marked effect of number on subject–verb agreement. *Journal of Memory and Language, 36*, 147–164.

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

Fayol, M., Largy, P., & Lemaire, P. (1994). Cognitive overload and orthographic errors: When cognitive overload enhances subject–verb agreement errors: A study in French written language. *The Quarterly Journal of Experimental Psychology, 47A*, 437–464.

Ferreira, V. S. (2003). The persistence of optional complementizer production: Why saying 'that' is not saying 'that' at all. *Journal of Memory and Language, 48*, 379–398.

Ferreira, V. S., & Bock, K. (2006). The functions of structural priming. *Language and Cognitive Processes, 21*, 1011–1029.

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

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.

Hare, M., Elman, J. L., & Daugherty, K. G. (1995). Default generalisation in connectionist networks. *Language and Cognitive Processes, 10*, 601–630.

Hartsuiker, R. J., Bernolet, S., Schoonbaert, S., Speybroeck, S., & Vanderelst, D. (2008). Syntactic priming persists while the lexical boost decays: Evidence from written and spoken dialogue. *Journal of Memory and Language, 58*, 214–238.

Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Syntactic co-ordination in dialogue. *Psychological Science, 15*, 409–414.

Hartsuiker, R. J., Schriefers, H. J., Bock, K., & Kikstra, G. M. (2003). Morphophonological influences on the construction of subject–verb agreement. *Memory and Cognition, 31*, 1316–1326.

Hartsuiker, R. J., & Westenberg, C. (2000). Word order priming in written and spoken sentence production. *Cognition, 75*, B27–B39.

Haskell, T. R., Mansfield, C. D., & Brewer, K. M. (submitted for publication). Language and category learning: The case of subject–verb agreement.

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

Haskell, T. R., & MacDonald, M. C. (2005). Constituent structure and linear order in language production: Evidence from subject–verb agreement. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 891–904.

Haspelmath, M. (2006). Against markedness (and what to replace it with). *Journal of Linguistics, 42*, 25–70.

Janssen, N., & Caramazza, A. (2003). The selection of closed-class words in noun phrase production: The case of Dutch determiners. *Journal of Memory and Language, 48*, 635–652.

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1–75.

Mallinson, G., & Blake, B. (1981). *Language typology: Cross-linguistic studies in syntax*. Amsterdam: North-Holland.

Marcus, M. P., Santorini, B., & Marcinkiewicz, M. A. (1993). Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics, 19*, 313–330.

Miozzo, M., & Caramazza, A. (1999). The selection of determiners in noun phrase production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 907–922.

Mirkovic, J., & MacDonald, M. C. (2003). *The role of morphophonological factors in agreement production: When singular and plural are both grammatical.* Poster presented at the 16th CUNY conference on human sentence processing, Boston, MA.

Peterson, R. R., & Savoy, P. (1998). Lexical selection and phonological encoding during language production: Evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory and Cognition, 24*, 539–557.

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language, 39*, 633–651.

Schriefers, H. (1993). Syntactic processes in the production of noun phrases. *Journal of Experimental Psychology: Learning, Memory and Cognition, 19*, 841–850.

Spalek, K., & Schriefers, H. (2005). Dominance affects determiner selection in language production. *Journal of Memory and Language, 52*, 103–119.

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

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

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

Vigliocco, G., & Nicol, J. (1998). Separating hierarchial 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.