Title
Effects of Semantic Integration and Advance Planning on Grammatical Encoding in Sentence Production

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 33(33)

ISSN
1069-7977

Authors
Gillespie, Maureen
Pealmutter, Neal J.

Publication Date
2011

Peer reviewed

Effects of Semantic Integration and Advance Planning on Grammatical Encoding in Sentence Production

Maureen Gillespie (gillespie.m@neu.edu), Neal J. Pearlmutter (pearlmutter@neu.edu)
Department of Psychology, 125 NI, Northeastern University
Boston, MA 02115 USA

Abstract

To examine how semantic integration and the scope of advance planning affect utterance planning and subject-verb agreement error rates, participants described picture displays using complex subject noun phrases (e.g., *the apple for the pie(s)*), with singular head nouns (*apple*) and local nouns (*pie(s)*) in prepositional phrase modifiers, and then completed them as full sentences. Semantic integration (the conceptual link between elements within a phrase; Solomon & Pearlmutter, 2004) was manipulated by varying the preposition used to link the head noun and local noun. Speech onset times and agreement error rates were recorded. Speakers were faster to initiate speech when the head and local noun were integrated than when they were unintegrated. Agreement errors were more likely when the local noun was plural than when it was singular. Supporting the scope of planning account of agreement (Gillespie & Pearlmutter, 2011), speakers who were slower to initiate speech produced more agreement errors, suggesting that speakers who do more advance planning are more likely to experience interference during agreement computation.

Keywords: language production; subject-verb number agreement; advance planning; speech errors; semantic integration; grammatical encoding

The study of language production is concerned with how speakers translate non-verbal thoughts into meaningful, grammatical utterances. Bock and Levelt (1994) separate the planning process required for language production into three main levels: the message level, which represents the speaker's intended meaning; the grammatical encoding level, which translates the meaning into a sequence of words; and the phonological encoding level, which translates the sequence of words into the sounds required to produce the utterance. Planning in language production is thought to proceed incrementally; the system takes advantage of the sequential nature of production, allowing speech to be initiated before an entire utterance is prepared for articulation. Planning in language production proceeds quickly and accurately the majority of the time; however, speech errors do occasionally occur. If the scope of advance planning at a given level is large, multiple items are likely to be simultaneously available, which increases the chance of interference and certain speech errors (Garrett, 1975).

Subject-verb agreement error production has been studied extensively in the language production literature (e.g. Bock & Cutting, 1992; Eberhard, Cutting, & Bock, 2005; Gillespie & Pearlmutter, 2011; Solomon & Pearlmutter, 2004). The most replicated finding in this domain is the mismatch effect: Agreement errors are most common when the subject of a complex noun phrase (NP) is singular and a local noun in a modifier is plural (e.g., **The key to the cabinets WERE...*; Bock & Miller, 1991). Some previous research has focused on how syntactic structure constrains agreement computation (e.g., Bock & Cutting, 1992; Franck, Vigliocco, & Nicol, 2002); however, there is increasing evidence that the timing of planning of a local noun relative to the head noun determines interference effects (Gillespie & Pearlmutter, 2010, 2011; Gillespie, Pearlmutter, & Shattuck-Hufnagel, 2010; Nicol, 1995; Solomon & Pearlmutter, 2004). Gillespie and Pearlmutter (2011) proposed that interference from local nouns is determined by the scope of advance planning in language production: Local nouns planned overlappingly with the head noun are more likely to interfere with agreement computation.

One way the head and local noun may be planned closer together in time is if they occur in a close semantic relationship. Solomon and Pearlmutter (2004) conducted experiments manipulating semantic integration (the degree to which elements within a phrase are conceptually linked) of the head and local noun. Participants read subject NP preambles with a local noun in a prepositional phrase (PP) modifier off a computer screen and then went on to complete them as full sentences. In a series of studies, they found that agreement errors were more likely to occur in highly integrated cases (e.g., *the chauffeur for the actor(s)*) than less integrated cases (e.g., *the chauffeur near the actor(s)*). They interpreted these findings as evidence that semantic integration can affect the timing of planning of elements within a phrase, independent of syntactic structure.

Another way the head and local noun may be planned with more overlap is if they appear close together within the utterance. Gillespie and Pearlmutter (2011) showed that plural local nouns that are linearly closer to the head noun are more likely to create interference during agreement computation than local nouns appearing more distally, when controlling structural distance. These results suggest that planning proceeds in the order in which items are to be produced, with semantic integration shifting the relative timing of planning. Thus, given Gillespie and Pearlmutter's findings, the degree to which speakers plan local nouns in advance should affect agreement error rates.

There is experimental evidence that the production system does a considerable amount of advance planning (e.g., Smith & Wheeldon, 1999; Wagner, Jescheniak, & Schriefers, 2010). Smith and Wheeldon examined the scope of conceptual and grammatical encoding during language production using a picture description task. They designed visual arrays of three pictures in which one, two, or all three pictures moved, and participants had to describe the movement. The movements were designed to elicit sentences where the subject NP was either complex (e.g., *the dog and the foot*) or simple (e.g., *the dog*). They found that speech onset times for single-clause sentences with complex subject NPs were longer than for single-clause sentences with simple subject NPs. They argued from this that the entirety of the subject NP must be grammatically encoded prior to speech onset, and that the scope of grammatical encoding is the first full phrase of an utterance.

Wagner et al. (2010) found evidence that the scope of advance planning of lexical items is flexible rather than determined by grammatical structure. In a series of studies, participants described picture displays, using simple sentences (e.g., *the frog is next to the mug*) and complex sentences (e.g., *the red frog is next to the blue mug*). An auditory distractor word that was semantically related to either the first noun or second noun was played at varying stimulus onset asynchronies. In general, when a semantically related distractor was co-present with a picture to be named, speech onset times were longer than when the distractor was unrelated to the picture, due to both lexical items being active and in competition for production. This effect is known as semantic interference. Wagner et al. found a reliable semantic interference effect for the second noun in the simple sentences and a weaker semantic interference effect for complex sentences, suggesting that speakers plan the second noun in their utterance farther in advance when cognitive demands are low. In addition, they found that speakers who initiated speech faster (those with shorter average speech onset times) showed reduced semantic interference effects relative to speakers who were slower to initiate speech, indicating the possibility of individual differences in planning scope. These findings suggest that individuals' planning scopes may influence how susceptible they are to agreement errors.

The current study examined the effect of advance planning on agreement error rates, by combining methods from the scope of planning literature and subject-verb agreement error elicitation studies. All previous studies of agreement production have had a comprehension component: Participants have either read or listened to preambles prior to using them to complete a full sentence. As mismatch effects are also observed in comprehension studies (e.g., Pearlmutter, Garnsey, & Bock, 1999), it is possible that interference from the comprehension system could be driving effects in production due to the two systems sharing at least some representations (e.g., Garrod & Pickering, 2004). Using a presentation method that requires speakers to formulate their utterance from the message level should better approximate the natural production process and avoid potential confounds from comprehension.

In the current study participants described picture displays using complex subject NPs (subject head nouns followed by PP modifiers) and then went on to complete these descriptions (preambles) as full sentences. Semantic integration was manipulated by varying the preposition used to link the head and local noun in the display. Responses were recorded, and speech onset times, which can reflect advance planning in language production (e.g. Smith & Wheeldon, 1999), were measured; and responses were examined for subject-verb agreement errors. This study aimed to replicate the mismatch and semantic integration effects on error rates from earlier studies, but using pictorial stimuli. Furthermore, because Solomon and Pearlmutter suggested that semantically integrated nouns are planned with more overlap than unintegrated pairs of nouns, planning of the subject NP may be completed more quickly when the head noun and local noun are semantically integrated than when they are unintegrated, predicting shorter speech onset times for integrated conditions than for unintegrated conditions.

Method

Participants One hundred fifty-three Northeastern University undergraduates participated in the experiment for course credit. After excluding data from non-native speakers, participants who had language/cognitive disorders or could not follow instructions, recording failures, and a counterbalancing error, 127 participants' data were included in the error analyses. Speech onset time (SOT) data were available for 118 of these participants.

Materials and design Twenty-four picture pairs like that shown in Figure 1 were created. The computer display screen was split in half vertically with one or two pictures arranged vertically on each side. A picture of the head noun, enclosed by a colored outline, was presented on one side of the screen, and a picture of the local noun was presented on the other side of the screen. In half the trials the head noun appeared on the left side of the screen (as in Figure 1), and in the other half the head noun appeared on the right. Local noun plurality was manipulated by presenting either a single picture (as in Figure 1a, 1c) or a pair of identical pictures on the same side of the screen (as in Figure 1b, 1d). Participants were required to construct NP PP preambles from the pictures presented. Participants linked the head noun and local noun with one of two prepositions that manipulated semantic integration; the color of the outline around the head noun determined which preposition was to be used. A blue outline indicated *for*, the preposition that created a more integrated relationship between the head and local noun; and a green outline indicated *near*, the preposition that created a less integrated relationship between the nouns (see Solomon & Pearlmutter, 2004). Examples of the preambles to be constructed for the displays in Figure 1 are provided in the second column. Local noun number, linking word, and head noun position were varied within picture pair, creating eight conditions. Each of eight counterbalanced lists included 24 fillers and exactly one version of each of the 24 experimental items. All fillers had the same basic structure as the critical items, but with plural head nouns.

Picture norming Seventy-six participants not involved in any other part of the experiment were presented with 122 black-and-white line drawings of inanimate objects and were

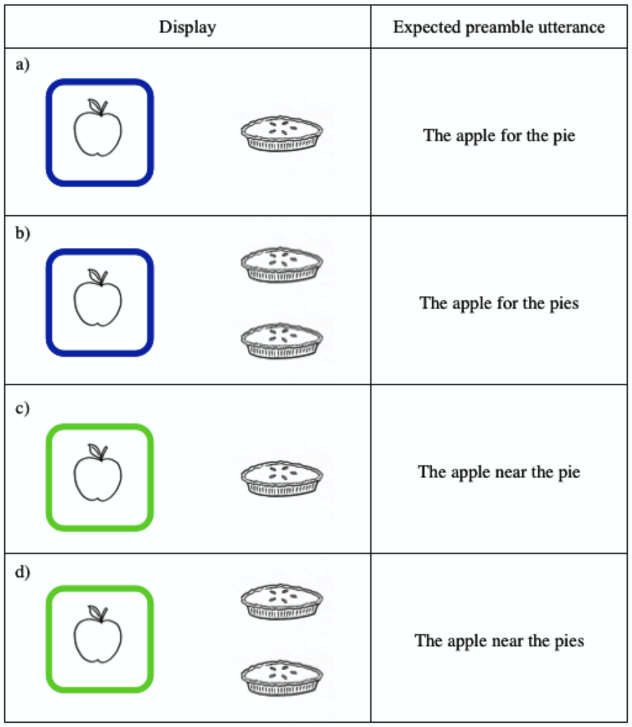

Figure 1: Examples of the picture displays representing four of the eight conditions. The remaining four conditions were created by switching the head noun and local noun positions on the screen. Preambles that were intended to be constructed from each display are presented in the second column.

required to provide a one-word label for the picture as quickly as possible. Data from one participant were excluded. Names and naming times were recorded for each picture. The 48 pictures used in the critical items were highly codable (i.e., only elicited one or two nearly synonymous labels), their labels were singular count nouns that had regular plural forms, and the average naming latency for critical items was 914ms (SD = 94ms). Pictures for filler and practice trials were also taken from this set.

Semantic integration norming A separate set of 123 participants produced the preamble aloud and rated the picture displays for semantic integration on a 7-point scale with higher numbers indicating a tighter conceptual link. As in Solomon and Pearlmutter (2004), the *for* versions were referred to as the integrated condition and the *near* versions were referred to as the unintegrated condition. Data from seven participants were excluded due to participants being non-native speakers or recording failure. Only trials for which participants correctly constructed the preamble with the expected noun labels (according to picture norming) were included, and these remaining data were analyzed in a linear mixed-effect model with head position, integration, local noun number, and their interactions as fixed effects; and with participant and item intercepts as random effects. Items linked with *for* were rated as more integrated than items linked with *near* ($t = 13.12$, $p < .001$)[1]. Items with plural local nouns were rated as more integrated than items with singular local nouns ($t = 2.13$, $p < .05$). An interaction of linking word and local noun number was marginal ($t = 1.94$, $p = .06$), with the effect of plurality stronger for *for* items than *near* items[2]. No other results were significant.

Procedure Each participant was run individually in the main experiment. On each trial, a fixation cross was presented in the center of the screen for 1000ms to focus attention. The picture display immediately followed. 500ms later the colored outline appeared, indicating which picture was the head noun. Participants were required to name the head noun picture (including the determiner *the*), produce the linking word indicated by the color of the outline, name the local noun picture (including the determiner *the*), and then produce an ending that formed a complete sentence. The pictures disappeared after 3000ms, and a prompt to continue to the next trial was presented. Participants were not instructed as to how they should formulate a completion, only that they should form a complete sentence. Speech onset times (SOTs) were measured from the onset of the picture display until the speaker triggered the voice key. Nine practice items preceded the 48 trials.

Apparatus A Macintosh computer running the PsyScope (v. 1.2.5) software package (Cohen, MacWhinney, Flatt, & Provost, 1993) controlled stimulus presentation, with a button box to record SOTs. Responses were recorded to CD for analysis.

Response scoring All responses were transcribed and assigned to one of four coding categories: (1) correct, if the participant said the preamble correctly, exactly once, produced an inflected verb immediately after the preamble, and used a verb form that was correctly marked for number; (2) error, if all the criteria for correct responses were met, but the verb form did not agree in number with the subject; (3) uninflected, if all the criteria for correct responses were met, but the verb was uninflected; and (4) miscellaneous, if the participant made an error in the preamble, if a verb did not immediately follow the preamble, or if the response did not fall into any of the other categories. Trials in which a participant made no response were excluded from all analyses. If the participant produced a disfluency (e.g., pauses, coughs) before, during, or immediately after producing the preamble and went on to produce a correct, error, or uninflected response, the scoring category and the disfluency were recorded.

---

[1] Regression analyses were performed in R (R Development Core Team., 2009), using the languageR package (Baayen, 2008). Models were fit using the lme4 package (v. 0.999375-32), and p-values were obtained using MCMC sampling in the coda package (v. 0.13-4).

[2] The semantic integration ratings for the plural local noun conditions are the most critical for the integration manipulation as only plural local noun conditions should show reliable error rates. Thus, integration was manipulated as desired.

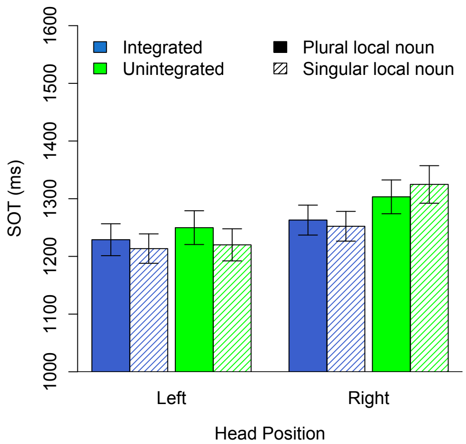

Figure 2: Mean SOTs as a function of head position, integration, and local noun number. Error bars indicate ±1 *SEM*.

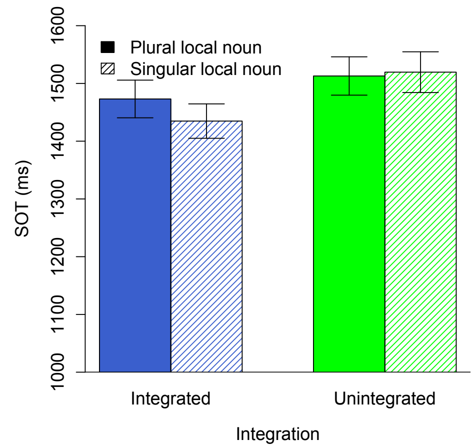

Figure 3: Mean SOTs for slow initiators as a function of integration and local noun number. Error bars indicate ±1 *SEM*.

Results

Speech onset time A linear mixed-effect model was used to analyze untransformed SOT data from 118 participants (Baayen, 2008). Analyses were also conducted for log-transformed SOTs; statistical patterns were identical to the untransformed data unless otherwise noted. Fixed effect predictors were head position, integration, noun number, and their interactions; random effects were participant and item intercepts. Only trials for which participants correctly constructed the preamble with the expected noun labels (according to norming) were included. Trials where the SOT was more than ±2.5 *SD*s from the participant's mean SOT were not included, nor were trials where the voice key was triggered before the appearance of the outline.

Mean SOTs for each condition are shown in Figure 2. Participants were faster to initiate speech when the head noun appeared on the left side of the screen than when it appeared on the right (t = -4.87, $p < .001$). Speech was initiated faster for integrated than unintegrated versions (t = -2.20, $p < .05$; this effect did not reach significance for log-transformed SOTs, t = -1.60, p = .11). No other results were significant ($|t|$s < 1.63, ps $> .13$).

If SOTs reflect advance planning of upcoming material, the speakers who are slower to initiate speech should show stronger integration effects than speakers who are faster to initiate (see Wagner et al., 2010). Linear mixed-effect models (as described above) were run on the data from the fastest half of participants ($N = 59$) and the slowest half of participants ($N = 59$). The fastest half of participants were faster to respond when the head noun appeared on the left than when it appeared on the right ($t = -4.68$, $p < .001$), but no other main effects nor interactions were significant. The slowest half of participants also showed the head position effect ($t = -2.94$, $p < .01$). In addition, they were faster to respond to integrated than unintegrated versions ($t = -2.77$, $p < .01$). The interaction of integration and local noun number was marginal ($t = 1.89$, $p = .08$). In the integrated condition, slower participants were slower to respond to the plural local noun condition than the singular local noun condition[3]; this difference was not observed for the unintegrated condition. This pattern is displayed in Figure 3.

Response analyses Analyses were performed for error rates that included disfluencies (patterns were identical if disfluency cases were excluded). Performing ANOVAs on proportion data is problematic and may produce spurious results; Jaeger (2008) instead suggested analyzing such data using logit mixed-effect models. However, the error rates produced in subject-verb agreement studies are extremely low, creating problems in applying the logit link function during model fitting (the log odds of proportions near 0 approach negative infinity). Thus, following Barr (2008), the data were analyzed using empirical logit weighted linear regression, aggregating separately over participants and items. By-participant and by-item weighted linear regressions on transformed error rates were performed with head position, linking word, noun number, and all interactions as sum-coded fixed effects (t-tests of parameter estimates are identified as t_1 for by-participant analyses and as t_2 for by-item analyses).

Untransformed mean error rates are shown in Figure 4. Agreement errors were more likely when the local noun was plural than when it was singular ($t_1 = 6.81$, $p < .01$; $t_2 = 10.2$, $p < .01$), but there were no other main effects nor interactions ($|t|$s < 1).

To determine if speakers' average SOTs predicted their error rates, a second by-participant error rate model was

[3] A linear mixed-effect model equivalent to a t-test comparing the local noun number conditions in the integrated version showed that this effect was marginal ($t = 2.31$, $p = .052$).

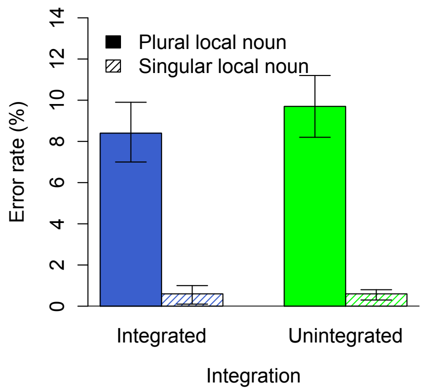

Figure 4: Untransformed mean error rates as a function of linking word and local noun number. Error bars represent ±1 *SEM* calculated from the by-participant analysis.

run with data from participants whose SOTs were available (N =118). Fixed effect predictors were local noun number, participant mean speech onset latency (obtained from the SOT analyses), and their interaction; the only random effect was the participant intercept. Agreement errors were more likely when the local noun was plural than when it was singular ($t = 6.68$, $p < .001$). The odds of producing agreement errors increased with speech onset latency ($t = 1.11$, $p < .05$). There was a marginal interaction of noun number and speech onset latency, such that the mismatch effect was larger for speakers with longer average SOTs ($t = 1.40$, $p = .06$).

For uninflected and miscellaneous responses, the only significant result was that there were more miscellaneous errors when the head appeared on the right than on the left ($t_1 =$ -2.01, $p < .05$; $t_2 =$ -2.50, $p = .19$).

Discussion

This agreement error elicitation paradigm allowed speech onset time (SOT) to be measured in addition to error rates, while approximating the natural production process. The SOT results reflect the nature of the grammatical encoding process in this task. Speakers were faster to initiate speech when the head noun appeared on the left. This result is possibly due to native English speakers in this task being biased to attend to the left side of the screen because English is written left to right, and speakers prefer to begin their utterances with the first item they look at (Griffin & Bock, 2000). Speakers were faster to initiate their responses when the head and local noun were integrated (linked with *for*) than when they were unintegrated (linked with *near*). This result suggests that speakers were sensitive to the semantic integration manipulation and completed planning of the subject NP more quickly when the preposition created a tighter semantic link between the head and local noun (see also Gillespie et al., 2010).

Interestingly, the integration effect was only present for the slowest initiators. The slower initiators also showed a trend toward a local noun number effect in the integrated condition, with no such effect in the unintegrated condition. If the slow initiators plan the full subject NP prior to speech onset (see Smith & Wheeldon, 1999), the head and local noun should be planned with some degree of temporal overlap. When the head and local nouns mismatched in number (plural local noun conditions), the number mismatch could introduce competition or interference, increasing SOTs. Thus, the speeded planning of the integrated condition may result in the head and local noun being planned with some overlap which would introduce competition in the number mismatch cases, predicting a local noun number effect. In the slower-to-plan unintegrated condition, the head and local noun may be planned with relatively little overlap, resulting in little or no interference in the number mismatch cases, resulting in no local noun number effect.

The scope of planning account of agreement computation (Gillespie & Pearlmutter, 2011) predicts that differences in timing of planning due to semantic integration (as observed in SOTs) should lead to a difference in error rates, but the integration effect (Solomon & Pearlmutter, 2004) was not replicated in the error data. It is possible that the overall timing difference between integrated and unintegrated cases was not large enough to have a reliable effect on error production. Or, speakers may have been less sensitive to the semantic integration manipulation than they were in previous studies. In previous studies, there were fillers that contained a variety of structures which could have forced participants to pay attention to the message-level properties expressed in the preambles; in the current task, the preambles had identical structures that only varied in the preposition, allowing the speakers to do less conceptual processing. Notably, speakers in this study often produced sentences that reflected a tight semantic link between the head and local noun even in unintegrated conditions (e.g., *The apple near the pies should be used in them*), which also could have decreased sensitivity to the semantic integration manipulation.

Even though the semantic integration effect was not replicated in the error analyses, these results suggest that the new paradigm can reliably reproduce the mismatch effect observed in many other agreement error elicitation studies relying on text or auditory stimuli (e.g. Bock & Miller, 1991). Interestingly, the slower initiators showed marginally stronger mismatch effects than the faster initiators. These results are consistent with the scope of planning account of agreement production (Gillespie & Pearlmutter, 2011): Speakers who plan more of their utterance in advance are more likely to experience interference from plural local nouns during agreement computation.

Of course it is possible that the speakers who were slower to initiate speech were just not very good at the task, which could lead to agreement errors as well as other types of speech errors such as disfluencies and preamble errors. To determine if this was the case, regression analyses predicting participant disfluency rates and preamble error rates (both empirical-logit transformed) from mean SOT were performed. There was no reliable relationship between mean SOT and disfluency rate ($t = -0.22$, $p = .82$), and the numerical pattern showed that speakers who were faster to initiate speech were more likely to be disfluent. Wagner et al. (2010) found a similar numerical pattern, with speakers who did less advance planning producing more disfluencies. Thus, the pattern of disfluencies observed in this study is consistent with the SOT and error data, suggesting that faster initiators did less advance planning than slower initiators.

There was a nonsignificant positive relationship between SOT and preamble error rates ($t = 1.10$, $p = .28$); however, with slower initiators more likely to produce errors when creating the preambles, this suggests that the slower initiators may have had more difficulty than the faster initiators. These preamble errors were of many types, including exchange errors (e.g., *the pies for apple... I mean the apple for the pies*) and anticipations (e.g., *the apples for the pies, no... the apple for the pies*). These types of errors have been hypothesized to arise due to multiple items being simultaneously planned and prepared for production (Dell, 1986; Garrett, 1975; Pearlmutter & Solomon, 2007). Unfortunately, unambiguous exchange and anticipation errors were very rare in the current experiment (10 out of 592 preamble errors), so there was not enough data to determine if slower initiators were also more susceptible to these specific types of speech errors.

Acknowledgments

We thank our research assistants for running participants and transcribing, and Keith Levin for creating the pictures.

References

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge, UK: Cambridge University Press.

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

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

Bock, K., & Levelt, W. (1994). Language production: Grammatical encoding. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics* (p. 945-984). San Diego, CA: Academic Press.

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

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

Dell, G. S. (1986). A spreading activation theory of retrieval in language production. *Psychological Review*, *93*, 283-321.

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

Franck, J., Vigliocco, G., & Nicol, J. (2002). Subject-verb agreement errors in French and English: The role of syntactic hierarchy. *Language and Cognitive Processes*, *17*, 371-404.

Garrett, M. F. (1975). The analysis of sentence production. In G. Bower (Ed.), *Psychology of learning and motivation* (Vol. 9, p. 133-177). New York: Academic Press.

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, *8*, 8-11.

Gillespie, M., & Pearlmutter, N. J. (2010). *Against structural constraints in subject-verb agreement production.* Manuscript submitted for publication.

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

Gillespie, M., Pearlmutter, N. J., & Shattuck-Hufnagel, S. (2010, March). *Prosodic phrasing reflects planning processes in sentence production.* Paper presented at the 23rd Annual CUNY Conference on Human Sentence Processing, New York, NY.

Griffin, Z., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, *11*, 274-279.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*, 434-446.

Nicol, J. L. (1995). Effects of clausal structure on subject-verb agreement errors. *Journal of Psycholinguistic Research*, *24*, 507-516.

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

Pearlmutter, N. J., & Solomon, E. S. (2007, March). *Semantic integration and competition versus incrementality in planning complex noun phrases.* Paper presented at the 20th Annual CUNY Conference on Human Sentence Processing, San Diego, CA.

R Development Core Team. (2009). *R: A language and environment for statistical computing.* Vienna: R Foundation for Statistical Computing, http://www.R-project.org.

Smith, M., & Wheeldon, L. (1999). High level processing scope in spoken sentence production. *Cognition*, *73*, 205-246.

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

Wagner, V., Jescheniak, J. D., & Schriefers, H. (2010). On the flexibility of grammatical advance planning during sentence production: Effects of cognitive load on multiple lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 323-340.