JOURNAL OF MEMORY AND LANGUAGE 31, 99–127 (1992)

Regulating Mental Energy: Performance Units in Language Production

KATHRYN BOCK AND J. COOPER CUTTING

Michigan State University

One of the classic puzzles of language is posed by the phenomenon of discontinuous dependency, in which the form of an element at one point in an utterance depends on the form of a noncontiguous controlling element. How do speakers use the information carried by the controller to implement the correct form of the dependent element? We contrasted two accounts of this process that differ in their assumptions about the organization of language formulation. The serial account, patterned after an augmented-transition-network model of the parsing of discontinuous dependencies, suggests that the controller is held in working memory until the point in the string at which the dependent appears. A second hypothesis, derived from a hierarchical model of language production, predicts that controllers and dependents within the same clause are specified concurrently, even when they are eventually separated in the utterance. Using a procedure to elicit verb-agreement errors in speech, we found that agreement errors were more frequent after phrases than after clauses that separated the verb from its head noun, reversing the direction of a related effect in language comprehension. When length varied, longer phrases led to more errors; longer clauses did not. These results support the hierarchical hypothesis.

By the standards of prescriptive grammar, one persistent error of English usage is the failure of a verb to agree in number with its subject. Such failures seem to be most frequent when the head subject noun is separated from its verb, as in "The only generalization I would dare to make about our customers are that they're pierced" (from an interview in a National Public Radio report on body-part piercing). In this utterance, the head noun (*generalization*) is singular, but the verb (*are*) is plural, appearing to agree with the immediate preverbal noun *customers*. In 1924, Otto Jespersen offered such errors as an example of the class of linguistic phenomena that ought to be explained in cognitive terms, venturing the hypothesis that "if the verb comes long after its subject, there is no more mental energy left to remember what was the number of the subject" (p. 345).

In the present work we will examine the occurrence of agreement errors in an exploration of what consumes "mental energy" in language production generally and in the creation of long-distance dependencies specifically. To give content to the notion of mental energy in this domain we will draw on accounts of language production which offer contrasting predictions about how the implementation of long-distance agreement could disrupt the formulation process. These issues are cast as questions about the organization of production mechanisms and about their scope over different sizes or types of production units. If the frequency of errors changes in the presence of variations in the sizes or types of language units, we may infer something about the normal organization of the mechanisms that create agreement and something about the nature of the units with which the production system usually works.

This research was supported by grants from the National Science Foundation (BNS 86-17659 and 90-09611) and the National Institutes of Health (R01 HD21011). We thank Kathleen Eberhard, Patricia Kaiser, and Carol Miller for their assistance in the collection of the data and Thomas Carr, Rose Zacks and the reviewers for comments on an earlier version of the manuscript. Send correspondence and reprint requests to Kathryn Bock at current address: Department of Psychology, 603 East Daniel, University of Illinois, Champaign, IL 61820.

0749-596X/92 $3.00

To set the stage for our research, we will survey some earlier arguments and evidence for various units of language performance and for the kinds of processing systems that deploy them. We emphasize units that immediately comprise utterances (words and clauses) rather than those that immediately comprise words (phonetic features, segments, syllables, and morphemes). Our focus is therefore on the processes that are termed *grammatical encoding* by Levelt (1989). Then we turn to a discussion of long-distance dependencies as a testing ground for hypotheses about how language processing systems work, and in that context we present the rationale for our experiments.

Units of Syntactic Encoding

The issue of performance units arises in part because speakers seem to know something about what they are going to say in advance of saying it, implying that there is a mental representation of the utterance that is separate from speech itself. One concrete indication of this is the prevalence of anticipatory errors, which arise when the speaker produces something in advance of its intended location in an utterance. Such errors seem to be much more common than perseveratory errors (Nooteboom, 1973). One question is how much conceptual or linguistic formulation can or must precede the onset of speech (see, for example, Lindsley, 1975), and closely related to this is the question of the types of units in syntactic encoding. Such units may be taken to constitute the minimal domain toward which syntactic encoding efforts can be individually directed.

One candidate unit is the word. The evidence that words constitute units with respect to language production comes from many sources, both logical and empirical, but we will mention just one. From observations of errors in spontaneous speech, it is clear that the production system manipulates words as wholes. They sometimes appear in places that violate the intentions of speakers, and they are prone to move to those places as integral elements. Among all the linguistic units that are represented in error corpora, whole words are rivaled only by individual phonemes in the frequency with which they are involved in errors (Dell, 1987; see also Stemberger, 1989), even though other types of units (e.g., phonetic features, syllables, morphemes) occur more often in the speech stream.

Despite their manifest role in production, it is not obvious that individual words constitute the units over which grammatical encoding operations are defined. Although word-to-word transitional probabilities have been found to affect the likelihood of pauses and hesitations (Goldman-Eisler, 1968; Maclay & Osgood, 1959; but see Tannenbaum, Williams, & Hillier, 1965), implying a word-by-word utterance planning process, reinterpretations of the hesitation data point toward a larger planning unit (Clark & Clark, 1977; Fodor, Bever, & Garrett, 1974; Valian, 1977). The reinterpretations hinge on the findings of Boomer (1965), who reported that the distribution of pauses and hesitations across phonemic clauses (the speech spanned by an intonation contour containing one primary stress) was very skewed, with disruptions being most common before the second words of clauses. Citing Carroll (1953), Boomer argued for a hierarchical encoding process in which larger grammatical units are selected before the words that fill them.

A chief suspect among large grammatical units is the syntactic clause (which corresponds roughly but not perfectly to the phonemic clauses examined by Boomer). Some of the evidence again comes from speech errors. Certain speech errors—those in which the mistake is traceable to another part of the intended utterance—offer a window into the planning process, indicating how much of the information that will appear in an utterance may be active before

the information is actually produced. Clauses appear to be an important delimiter. Fodor et al. (1974) reported unpublished work by Garrett and Shattuck which showed that among 170 sound exchanges, only two crossed clause boundaries. For word exchanges like *faster than the sound of speed* (when *speed of sound* was intended), Garrett (1980) reported that approximately 80% of the errors in his corpus occurred within clauses. Other reflections of the likely involvement of clause units in production planning include delays in detecting extraneous signals at the ends of clauses (Ford & Holmes, 1978), pauses at clause boundaries (Cooper, Paccia, & Lapointe, 1978; Deese, 1984; Gee & Grosjean, 1983; Goldman-Eisler, 1968), and a decreased likelihood of detecting one's own speech errors at the beginnings of constituents (Levelt, 1983).

Although there is a consensus that the clause is a unit of grammatical encoding, opinions diverge about what constitutes a clause. Boomer (1965) employed a phonological criterion, but more recent work on prosodic planning opposes the equation of syntactic and prosodic units (Ferreira, 1991). Fodor et al. (1974), interpreting Boomer's data, argued instead for syntactic or *finite* clauses as planning units, that is, clauses with tensed verbs and, usually, an explicit set of noun-phrase arguments (appearing as subjects, direct objects, and so on). In contrast, Ford and Holmes (1978; Ford, 1982; Holmes, 1988) have found that nonfinite clauses have the same consequences as finite clauses for pauses, hesitations, and latencies to respond to extraneous signals during spontaneous speech. Nonfinite clauses have untensed verbs (infinitives or gerunds), often without a full set of arguments (e.g., *going home* as in *Going home made Kathy happy* or *to program* as in *Laura loves to program*). The implication is that for purposes of explaining production demands, clauses may have to be defined over verb groups, regardless of whether those groups meet traditional criteria for clausal status. In the experiments reported below, we sidestep this question by employing only finite clauses.

The Organization of Language Production

Language processing systems may differ not only in their units of operation, but also in the ways they piece those units together or pull them apart in time. Although claims about units are historically closely related to claims about mechanisms, they are in principle independent. In this section we will compare two general approaches to language production, one a serial account and the other a hierarchical account of how utterances are created. These accounts differ in their relative emphasis on alternative ways in which the sequencing required in language performance may be controlled, whether by means of sequential connections between units of the same type, or hierarchical connections between superordinate and subordinate units.

Serial models of production have a long tradition in psychological accounts of language. Lounsbury (1965) offered a well known version in which the production of language units (phonemes, morphemes, or words) is held to be heavily influenced by the transitional probabilities between units or groups of units. On this view, what controls the likelihood of one word being produced after another one is the past frequency of their cooccurrence in the experience of the speaker. Similarly, what controls the likelihood of a particular word following a particular two-word sequence is the past frequency with which that word has followed that sequence, and so on. Though Lounsbury acknowledged the possible relevance of syntactic constituents to language use, particularly in comprehension, he hypothesized that the statistical and linguistic structures are independent, and that the statistical structure carries greater weight in language production. In line with this, disruptions of production

should be most likely to occur at points of low transitional probability which, on Lounsbury's view, need not correspond to the boundaries of higher-level linguistic constituents.

In contrast, most modern theories of language production assume a mode of operation that is heavily constrained by higher-level constituents (Bock, 1987; Dell, 1986; Fromkin, 1971; Garrett, 1988; Kempen & Hoenkamp, 1987; Levelt, 1989; MacKay, 1982; Shattuck-Hufnagel, 1979; Stemberger, 1985). Though the theories divide over how selective the mechanisms are and therefore how strong the hierarchy of control is, they all explain speech errors and other production problems within a fundamentally hierarchical architecture.

We will follow Garrett (1982, 1988) in laying out how this hierarchical control might work, since his model makes specific claims about the types of language units that are involved in different procedures and how those units are coordinated. To account for the data from studies of speech errors and hesitations, the model divides syntactic encoding into two levels of processing, the functional level and the positional level. Functional-level processing involves the integration of words (technically, representations of the semantic/syntactic properties of words) into a structural scheme that specifies certain syntactic functions (which we will take to be functions such as subject and direct object, although Garrett does not identify them as such). This structural scheme creates a clause.

As clauses are assembled, the information they contain becomes available for positional-level processing. This involves retrieving phonological representations for the words and using those representations to spell out the phonological segments. The scope of this phonological specification is roughly phrasal, limited to a single major phrase at a time.

In this model, production problems are most likely to arise from interference between similar elements that are concurrently active. For example, similarity during functional processing is defined by the semantic/syntactic properties of words, and the problems that have been attributed to this level involve interactions between words from the same grammatical categories (like *sound* and *speed* in the error *faster than the sound of speed*). Since this approach draws on the characteristics of the representations manipulated during processing rather than on such general metrics as transition probability, the proposed sources of problems are structural. In the next section we will set out some predictions about syntactic errors, specifically errors of verb agreement, that can be derived from serial and hierarchical accounts of production.

The Challenge of Long-Distance Dependencies

The serial-associative explanation of subject–verb agreement errors is an intuitively natural one: People say things such as *The time for fun and games are over* because the association between *games* and *are* is stronger than the association between *games* and *is*. However, there are two serious empirical challenges to this hypothesis, both of them reported in experiments by Bock and Miller (1991). First, failures of agreement are much more common when a plural noun separates the head noun and verb (as in the error *We know what the result of the 1930s were*) than when a singular noun intervenes (as in *The educational systems needed to correct the problem is lacking*). If errors were simply the product of associative competitions, these two types of errors should either be comparably frequent, or the latter should be even more frequent than the former (since singular nouns are more frequent than plurals). Second, agreement errors should predominate whenever the head noun and verb are separated by another noun phrase that disagrees with the head in number. Yet in Bock and Miller's results, errors consti-

tuted only a quarter of the responses even in the condition in which they occurred most frequently.

Such findings illustrate the problem that linguistic dependencies pose for simple associative accounts of language use. Regular covariation between subject and verb, even when the head noun and verb are separated in the speech stream (as in *The boy always runs fast* and *The boys always run fast, The boy from the track team runs fast* and *The boys from the track team run fast,* and so on), is unexpected from the standpoint of simple associationism because, as the distance between a head noun and verb increases, the association should grow weaker. Particularly when the interruption introduces another noun with a different number from that of the head, these weak associations should be overriden by stronger, local associations.

However, there are other ways to account for discontinuous dependences within a serial architecture. One involves the introduction of a limited capacity working memory, as in the augmented-transition-network model proposed by Wanner and Maratsos (1978) for parsing relative clauses. Wanner and Maratsos hypothesized that the head of a relative (e.g., *good-for-nothing* in *the good-for-nothing that Alice married*) is placed on a mental stack until its proper position within the relative clause is located (after *married*). This stack is held in working memory in the meantime, so that available memory resources should be reduced. Wanner and Maratsos found evidence consistent with this hypothesis in two experiments (for contrasting perspectives see Ford, 1983; Frauenfelder, Segui, & Mehler, 1980; and Holmes & O'Regan, 1981).

This more powerful serial model suggests a way in which the number of a head noun might be preserved over intervening material for the purposes of implementing subject–verb agreement in production. The head noun may be held in working memory until the main verb occurs, at which point the number of the head is used in formulating the correct verb form. If intervening material disrupts memory for the head noun or for its number, as a subject postmodifier increases in length or complexity, the probability of an agreement error should also increase. In particular, if a clause constitutes a unit toward which substantial processing efforts are directed (see Just & Carpenter, 1987, Chap. 5, for a review of the comprehension evidence for this supposition, and Ford & Holmes, 1978, for related evidence from production), the interposition of an entire clause between the subject and verb may be more likely to disrupt the agreement process than the interposition of subunits, in the same way that the interposition of a clause disrupts memory for the material preceding it (Caplan, 1972; Jarvella, 1971). The association noted above between agreement errors and plural postmodifiers could arise because plurals are semantically and morphologically more complex than singulars, creating an additional drain on working memory capacity.

The potential burden on a limited capacity memory of implementing discontinuous agreement in a serial production system offers a concrete instantiation of Jespersen's (1924) mental energy notion and is clearly in the spirit of his proposal. It contrasts with the potential demands on mental energy in a hierarchical production system. Because language production involves a massive top–down flow of information, Bock (1991) suggested that a major processing problem is regulating interference between activated elements that are similar to one another in structurally important ways. The hierarchical architecture progressively narrows the domains within which likely sources of interference can be regulated. The resulting focus on competition between information types for specific mechanisms and on structural sources of interference is compatible with selection-for-action approaches to attentional selectivity (Allport, 1989).

The hierarchical account of production

therefore offers different predictions about the distribution of agreement errors than the serial, limited-capacity memory account. The chief contrast has to do with consequences of the different types of units that can separate head nouns and verbs. This difference rests in part on a fact about subject–verb agreement: The subject and verb that agree are clausemates. Because the functional level of formulation (the level where subjects are designated and agreement must be specified) is concerned with integrating the information in clauses, and because the organization of the information at that level does not necessarily reflect the eventual order of information in the utterance, an incipient two-clause utterance such as *The claim that wolves were stealing babies was rejected* might be represented in the way schematized in the upper panel of Fig. 1. Note that the matrix clause (*The claim was rejected*) occupies a separate processing structure from the subordinate clause (*wolves were stealing babies*).

This clausal packaging—which is also hypothesized to make certain cross-clause speech errors less frequent than certain intraclause errors—partially insulates the information in one clause from the information in another. Since agreement is clause-bounded, the implication is that information from one clause should be unlikely to interfere with the specification of agreement in another clause. Accordingly, a clausal interruption between a head noun and its verb (e.g., *that wolves were stealing babies* separating *claim* and *was*) should be associated with fewer agreement errors on the verb of the matrix clause (*was*) than simple phrasal interruptions (e.g., *about the stolen babies* in *The claim about the stolen babies was rejected*). Phrasal interruptions, as in the lower panel of Fig. 1, introduce more information into a single processing structure and so create more potential sources of interference.

To summarize, the serial model (coupled with a limited capacity memory to allow it to generate discontinuous dependencies) predicts that increases in the length or syntactic complexity of the material separating a head noun from its verb should be associated with more agreement errors. Since clauses are structurally more complex than phrases, clausal separations should create more errors than phrasal separations. The hierarchical hypothesis makes the opposite prediction: Clausal separations should create fewer errors.

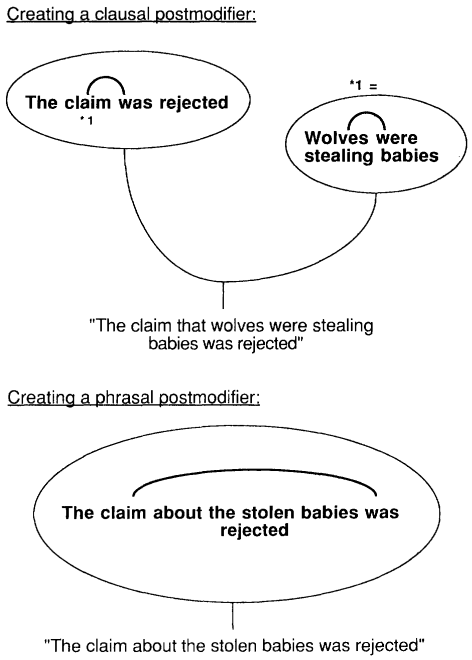

FIG. 1. The specification of agreement in a hierarchical production model for clausal (upper panel) and phrasal (lower panel) postmodifiers of the head noun phrase *The claim*. The ellipses enclose functional processing units (individual clauses), and the arcs inside the ellipses connect the agreeing elements within the clause. The starred index for the clausal postmodifier represents an embedding operation.

An Overview of the Experiments

To test these contrasting predictions, we employed a paradigm used by Bock and Miller (1991) that elicits agreement errors with properties very similar to those of naturally occurring errors. The errors all represented instances of *attraction,* or erroneous proximity concord (Francis, 1986). Proximity concord is agreement between a

verb and its immediately preceding noun phrase. It contrasts with the usual pattern of agreement or concord between the verb and the head noun of the subject phrase. In other work (Bock & Eberhard, 1991; Bock & Miller, 1991), we have found that attraction errors are not confined to a few individuals for whom proximity concord is the preferred pattern, that they are not strongly related to knowledge of prescriptive agreement rules, and that they are predominantly syntactic or inflectional rather than semantic or phonological.

The experimental task involved the creation of endings for sentences whose beginnings were specified for the speakers. The beginnings, or preambles, consisted of a complex subject that included both a head noun and a postmodifier that was either a phrase (e.g., *The advisor for the chemistry students . . .*) or a clause (*The advisor who directed the students . . .*). In all of the experiments, the phrase and clause preambles were matched in length and had the same head noun phrases and local nouns (the nouns at the ends of the preambles that immediately preceded the verb of the speaker's completion). We varied the number of the head and the local noun to create experimental conditions in which the plurality of the head and local noun mismatched. The numbers of agreement errors that occurred under these mismatch conditions were compared with control conditions in which the head and local noun matched in plurality.

In all three experiments we compared the incidence of agreement errors after phrasal and clausal postmodifiers, in order to test the contrasting predictions of the serial and hierarchical accounts of production demands. Experiment 1 employed prepositional-phrase and relative-clause postmodifiers, and Experiments 2 and 3 compared prepositional-phrase and noun-phrase complement postmodifiers. In Experiment 3 we also manipulated the length of the postmodifiers. This increased length would be expected to increase the difficulty of the agreement operation for both phrases and clauses, on the serial account, whereas the hierarchical view predicts no uniform increase in the difficulty of agreement.

Experiment 1

The contrast between prepositional-phrase and relative-clause postmodifiers that was incorporated into the first experiment was designed in part to follow up on a result reported by Bock and Miller (1991, Experiment 1). Their materials included both phrasal and clausal postmodifiers and, consonant with the hierarchical prediction, produced a reliable trend toward more errors after phrasal postmodifiers. However, there were other differences between the postmodifier types that complicated the interpretation of this result. In this experiment we eliminated those differences by creating sets of preambles in which the phrasal and clausal preambles had the same numbers of syllables and the same stress patterns, the same head nouns, and the same local nouns.

This experiment and those that follow also included a speaking span test modelled after one developed by Daneman and Green (1986). Daneman and Green showed that individuals with high speaking spans, as measured by this test, were faster at producing context-appropriate replacements for words in sentences. Conceivably, such individuals have more working memory capacity available during sentence production to devote to word finding and other formulation tasks. We included the speaking span measure as an additional test of the limited-capacity memory account of agreement in the serial production hypothesis. Applying individual-differences logic to the problem of discontinuous dependencies, the obvious prediction is that higher speaking spans should be associated with decreased numbers of agreement errors.

Method

Participants. Eighty Michigan State University undergraduates participated to ful-

fill an extra-credit option in introductory psychology courses. All were native English speakers.

Materials. The primary experimental materials consisted of 32 sentence preambles. Every preamble constituted a complex subject phrase that included two noun phrases, one the head (the first noun phrase) and the other the local (the second, preverbal noun phrase). The local noun was part of a phrase or clause constituent that served as a postmodifier of the head.

There were eight versions of each preamble corresponding to the eight conditions of the experiment, as illustrated in Table 1. Four versions had a prepositional phrase postmodifier and four had a relative clause postmodifier. Of each type of postmodifier, half had singular and half had plural heads, and of these, half had heads and local nouns that matched in number and half had heads and local nouns that mismatched in number. The verbs in the relative clauses were past tense and unmarked for number. All versions of a preamble had the same head and local noun, the same numbers of syllables, and the same stress patterns. The complete set of items is listed in Appendix A.

In addition to the experimental preambles, 56 filler preambles were constructed. All were simple noun phrases (either determiner-noun or determiner-adjective-noun). Half of the phrases were singular and half plural.

Eight 88-item lists were created from these materials. Every list included all 56 fillers, and one version of each of the 32 experimental preambles. Within each list, every experimental condition was represented by four preambles on each list. Across the eight lists, each version of the 32 experimental preambles occurred only once.

Every list started with eight fillers. The remainder of the fillers and experimental items were distributed randomly, with the constraint that no more than two experimental items could occur consecutively. The individual filler and experimental preambles occupied the same locations in all eight lists.

The lists were recorded on audio tape by a female speaker. Each preamble was spoken as rapidly as possible without compromising clarity.

Procedure. The participants were run one at a time. They were told that they would hear a series of sentence beginnings, and that their task was to repeat each one along with a completion for the sentence. They were asked to respond quickly with the first complete sentence that came to mind, and to speak as fast as they were able. No other restraints were put on the forms or contents of the completions. Two

TABLE 1
AN EXAMPLE OF THE VERSIONS OF A SENTENCE PREAMBLE FROM EXPERIMENT 1

Number match of head and local noun	Number of local noun	Preamble
Prepositional-phrase preambles		
Mismatch	Singular	The editors of the history book
	Plural	The editor of the history books
Match	Singular	The editor of the history book
	Plural	The editors of the history books
Relative-clause preambles		
Mismatch	Singular	The editors who rejected the book
	Plural	The editor who rejected the books
Match	Singular	The editor who rejected the book
	Plural	The editors who rejected the books

practice preambles (both of them simple singular noun phrases) were presented for the participants to complete.

The recorded preambles were presented one by one, and the participant repeated each one along with a completion. In the rare cases of failures to hear a preamble, the experimenter repeated the preamble aloud. Whenever the participants' speech rates dropped noticeably the experimenter encouraged them to talk faster. The experimental sessions were recorded on audio tape.

Following each session a speaking span test (adapted from Daneman & Green, 1986) was administered. The test required the participants to construct sensible and grammatical sentences for every word in a list. The numbers of words in each list ranged from two to five, and all the words in a list were presented in one trial. Since each word was to appear in one sentence, the number of sentences produced on a trial could also range from two to five.

The test included 70 unrelated words, each two syllables long, with concreteness values greater than or equal to 5.00 and Thorndike–Lorge frequencies of A or AA in the Paivio, Yuille, and Madigan norms (1968). Each word was randomly assigned to one of 20 lists. The set of 20 lists contained five lists of each of four lengths (two, three, four, or five words).

The lists were arranged in five cycles, each cycle including a two-, a three-, a four-, and a five-word list, presented in progression. All the words were presented auditorily. The participants had to maintain the words from each list in memory until the end of the list, when they could begin to produce sentences.

A single cycle of four lists, using different words from those in the test, was given as practice before the speaking span test itself was administered.

Scoring the preamble completions. The completions were transcribed and then assigned to one of four scoring categories. A completion was scored as a correct response when a participant repeated the preamble correctly, said it only once, produced an inflected verb as the first word of the completion, and used a verb form that correctly marked the singular or plural (all third-person present tense forms and the past tense forms of *to be*). If two different forms were produced in succession, only the first was scored. Completions were scored as agreement errors when all the above criteria for correct responses were met, except that the verb form failed to agree in number with the subject of the sentence. The third category contained completions that met all the criteria for correct responses except that the verb form was one that has no inflection for number, and so it failed to differentiate plural from singular (we will call these uninflected-number responses). Application of these criteria yielded 1426 corrects (55.7% of all responses), 75 agreement errors (2.9%), and 714 uninflected-number responses (27.9%).

The remainder of the responses were classified as miscellaneous. There were 345 of these, and most of them (82%) contained preamble reproduction errors. The others included multiple repetitions of the preamble (12%), had no verb (2%), were noncompletions or responses such as "I don't know" (2%), or omitted the preamble (1%). The rest did not have the verb as the first word or were in a language other than English.

Appendix B lists random samples of four correct and four error responses from the phrase and clause conditions of this and the two subsequent experiments.

Design and data analyses. The three experimental factors were (1) type of postmodifier (prepositional phrase versus relative clause); (2) the number of the local noun (singular versus plural); and (3) the correspondence between the number of the head and the local noun (match versus mismatch). Orthogonal combinations of these three factors yielded eight conditions. Every participant received four preambles representing each of the eight conditions,

and every sentence preamble was presented to ten participants in each of the same eight conditions.

The numbers of agreement errors constituted the dependent variable for the major statistical tests. Two analyses of variance were performed, one with participants and another with items as random factors. The test statistics for these analyses are designated F_1 and F_2, respectively. All effects that achieved significance were reliable at or beyond the .05 level.

Since the results for correct responses were generally complementary to those for the errors, these analyses are not reported. The most important error results are shown in figures as net proportions of the total number of inflected verbs produced, both correct and incorrect. The net proportions are the proportions of errors that remain in the mismatch conditions after subtracting the proportions of errors in the corresponding match conditions. These figures provide information about error base rates and aid comparisons of major results across experiments.

TABLE 2
NUMBERS OF RESPONSES BY SCORING CATEGORY AND CONDITION IN EXPERIMENT 1

Number of local noun	Type of postmodifier: Prepositional phrase	Type of postmodifier: Relative clause
Correct responses		
	Number mismatch	
Singular	180	177
Plural	140	167
	Number match	
Singular	184	208
Plural	174	196
Agreement errors		
	Number mismatch	
Singular	8	3
Plural	29	22
	Number match	
Singular	3	1
Plural	6	3
Uninflected-number responses		
	Number mismatch	
Singular	90	87
Plural	90	84
	Number match	
Singular	115	78
Plural	88	82
Miscellaneous responses		
	Number mismatch	
Singular	42	53
Plural	61	47
	Number match	
Singular	18	33
Plural	52	39

Results

The numbers of responses in each condition are given in Table 2. Agreement errors occurred more often in the mismatch than in the match condition and more often after plural than singular local nouns. This produced main effects of number match ($F_1(1,79) = 23.02$; $F_2(1,31) = 34.22$) and of local-noun number ($F_1(1,79) = 11.45$; $F_2(1,31) = 8.43$) in the analyses of variance. Local-noun number made a difference primarily in the mismatch condition, with most of the agreement errors (68%) occurring when the subject was singular and the local noun was plural. This yielded a significant interaction between match and the number of the local noun ($F_1(1,79) = 20.06$; $F_2(1,31) = 22.36$).

Errors were reliably more frequent after prepositional-phrase postmodifiers (overall, 61.3% of the errors) than after relative-clause postmodifiers ($F_1(1,79) = 4.65$; $F_2(1.31) = 4.52$). None of the interactions involving postmodifier type achieved reliability (all Fs < 1). This result is summarized in Fig. 2.

The distribution of uninflected-number responses is also shown in Table 2. The effect of postmodifier type was significant with participants random but marginal with items random ($F_1(1,79) = 5.38$; $F_2(1,31) =$

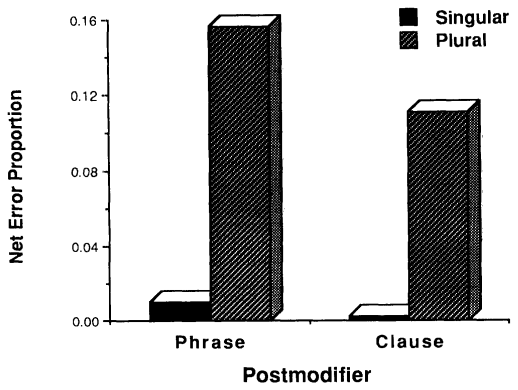

FIG. 2. The net proportions of agreement errors in the mismatch conditions in Experiment 1. The net proportions are the proportions of errors that occurred when the head and local noun mismatched in number minus the proportions of errors that occurred when they matched.

TABLE 3
PROPORTIONS OF AGREEMENT ERRORS AMONG THE MISCELLANEOUS RESPONSES IN EXPERIMENT 1

Number of local noun	Postmodifier type: Prepositional phrase	Postmodifier type: Relative clause
	Number mismatch	
Singular	.12 (32)	.13 (30)
Plural	.47 (15)	.20 (20)
	Number match	
Singular	.08 (24)	.00 (34)
Plural	.03 (39)	.07 (38)

Note. The entries represent the proportions of errors among all the correctly and incorrectly agreeing number-inflected verbs in the miscellaneous responses. The total number of inflected verbs is shown in parentheses.

3.10, $p < .10$). No other differences were significant.

Miscellaneous responses, shown in the bottom panel of Table 2, were more frequent in the mismatch than in the match condition ($F_1(1,79) = 11.99$; $F_2(1,31) = 9.42$). There were significant interactions between match/mismatch and local noun number ($F_1(1,79) = 10.43$; $F_2(1,31) = 4.28$), and among all three factors ($F_1(1,79) = 10.28$; $F_2(1,32) = 8.63$). The interactions together reflect a difference in the patterns for phrasal and clausal postmodifiers. Local noun number had comparatively little effect on the occurrence of miscellaneous responses for relative clauses, but for prepositional phrases, there were disproportionately few miscellaneous responses after singular local nouns in the match condition. Overall, however, there were similar numbers of miscellaneous responses associated with the phrase and clause conditions, 173 and 172, respectively.

It is possible to divide the miscellaneous responses into those for which an agreeing verb was correct or incorrect relative to the form in which the preamble was actually produced (which, for these responses, was generally different from the form in which it was presented). Table 3 presents the results of this classification, giving agreement errors as proportions of the correct and incorrect responses. It omits responses with uninflected-number verbs and responses that did not contain verbs. The general pattern, particularly in the mismatch condition, is similar to the results for the errors after correctly repeated preambles.

Speaking-span performance was evaluated by examining the total number of appropriate sentences that the participants produced (Daneman & Green, 1986). An appropriate sentence was one that was grammatical and contained one of the target words from the immediately preceding word set. If more than one target word was used in the same sentence, credit was given for only one. Likewise, if a single target word was used in more than one sentence, credit was given for only one.

The mean total-performance score on the speaking span test was 51.1, with a standard deviation of 5.6 and a range of 38 to 63. Correlations were calculated between the participants' total-performance scores and their (1) agreement errors (as a proportion of the total of correct and incorrect agreements); and (2) the number of miscellaneous responses they produced. The correlations were $r = -.15$ between speaking

span and the proportion of agreement errors and $r = -.40$ between speaking span and miscellaneous responses. Only the latter was significantly different from zero, $t(78) = 3.81$.

Discussion

The primary result of Experiment 1 was that clausal postmodifiers yielded fewer agreement errors than phrasal postmodifiers. This replicates Bock and Miller (1991, Experiment 1) under better-controlled conditions, and is more in line with the hierarchical than the serial prediction. However, the interpretation of the finding remains problematic in one important way. The clausal postmodifiers were all relative clauses, and relative clauses have properties that may make them less vulnerable to agreement problems than other clause types. We will take this up in the next experiment.

The outcome of the speaking span analysis was also inconsistent with the serial model's limited-memory account of agreement implementation. Although measured speaking span bore a detectable relationship to the speakers' ability to correctly repeat the sentence preambles from memory, the correlation with correct agreement performance was negligible.

The variation in the plurality of the head and local noun produced a striking asymmetry, as in earlier work (Bock & Miller, 1991). Substantially more errors occurred when the head was singular and the local noun plural than when the head was plural and the local noun singular. In the latter condition, the number of errors was not much different from conditions in which the head and local noun matched in number. This suggests that the singular verb form is a default that is overridden or suppressed in the presence of a plural, rather than an actively triggered consequence of singular noun marking. Because the error mechanism seems to be keyed to local-noun plurality and because our immediate interest centers on the scope of the mechanism, we simplified the design of the next two experiments by omitting the conditions that assessed singular attraction errors.

Experiment 2

The second experiment was undertaken to determine whether a different type of clause, the noun-phrase-complement clause, would yield results comparable to those for relative clauses in the first experiment. Relative clauses are dependent upon the noun phrase that they modify, in that the relative pronoun (*that, who,* etc.) stands for the modified noun phrase. So, in *The advisor who directed the students . . . , who* stands for the advisor, making it a part of the proposition conveyed by the relative clause itself as well as part of its syntactic structure, since *who* serves as the subject of the verb *directed.* Ultimately, then, the verbs in the higher and lower clauses agree with the same constituent. In contrast, a complement clause is complete without the participation of an element from the higher clause. So, in *The report that they controlled the fire . . . ,* the clause *that they controlled the fire* specifies the report without incorporating *the report* into the lower clause's structure.

This difference may have consequences for production. Because the relative pronoun in a relative clause reinstantiates the head noun phrase, it may constitute a reminder about the number of the subject, thereby making agreement with the main-clause verb simpler than it would otherwise be (and reducing the number of errors in comparison to phrasal postmodifiers, as in the first experiment). There are several results in the parsing literature which suggest that readers and listeners implicitly fill such slots during sentence comprehension (Bever & McElree, 1988; Garnsey, Tanenhaus, & Chapman, 1989; MacDonald, 1989; Nicol & Osterhout, 1989; Swinney, Nicol, Ford, Frauenfelder, & Bresnan, 1987); if they are also filled during production, the link between the head noun and main verb may be reinforced. Some evidence that rel-

TABLE 4
AN EXAMPLE OF THE VERSIONS OF A SENTENCE PREAMBLE FROM EXPERIMENT 2

Type of postmodifier	Number of local noun	Preamble
Prepositional phrase	Singular	The report of the destructive fire
	Plural	The report of the destructive fires
Complement clause	Singular	The report that they controlled the fire
	Plural	The report that they controlled the fires

ative clauses may indeed be easier to produce can be found in Rochester and Gill (1973), who reported that relatives were less likely to be accompanied by hesitations than were complement clauses.

To determine whether the reduction in agreement errors that we observed in Experiment 1 was only due to the use of relative clauses, in this experiment we employed complements instead. The materials were in most other ways comparable to those in the first experiment, in that we matched preambles for their length, stress patterns, and heads and local nouns. However, for the reasons we noted above, we used only singular head nouns and varied the number of the local noun. This created a match condition in which the head and local nouns were both singular, and a mismatch condition in which the head was singular and the local noun was plural.

Method

Participants. The participants were 80 Michigan State University undergraduates, all native English speakers, fulfilling an extra-credit option in introductory psychology classes. Two were replaced because they had unusual difficulty repeating the preambles correctly.[1]

Materials. The primary experimental materials consisted of 32 sentence preambles. Four versions of each preamble were generated, as illustrated in Table 4. All of the versions contained the same head and local noun and had the same numbers of syllables and the same stress patterns. The heads of all the preambles were singular in number. Two versions contained prepositional phrase postmodifiers and two contained complement clause postmodifiers. For each type of postmodifier, one version had a singular local noun and the other a plural local noun. Within the complement-clause postmodifiers, the subject of the clause was singular for half of the items and plural for the other half. The clause verbs were uninflected in 26 of the 32 items, singularly inflected in 2, and plurally inflected in 4.

A new set of 56 filler preambles was designed to lessen the contrast in length between the filler and experimental items. These were equally divided among four types of noun phrases, determiner-noun, determiner-adjective-noun, determiner-adjective-adjective-noun, and determiner-noun-relative clause. Half the fillers of each type, except the third, were singular and half were plural. In the determiner-adjective-adjective-noun fillers, six were singular and eight plural. All of the relative clause fillers were subject relatives, half with singular and half with plural objects.

Four 88-item lists were created from these materials following the scheme described for Experiment 1, with the additional constraint that no experimental preambles could occur consecutively. The lists were recorded by a male speaker who produced each preamble as clearly and rapidly as possible.

[1] The agreement error rates for these two subjects were not particularly high. One produced two errors out of 24 scorable responses, and the other produced one error out of 12 scorable responses.

Procedure. The procedure from Experiment 1 was followed in presenting the lists. The speaking-span test was again administered after the preamble completion task.

Scoring. Application of the scoring criteria described for Experiment 1 produced the following distribution of responses: 1823 corrects (71.2%), 54 agreement errors (2.1%), 314 uninflected-number (12.3%), and 369 miscellaneous (14.4%). Of the miscellaneous responses, 77% were reproduction errors.

Design and data analysis. Orthogonal combinations of the factors of postmodifier type (prepositional phrase versus complement clause) and number of the local noun (singular or plural), created four conditions. In each one, every participant received eight preambles. Every sentence preamble occurred in every cell of the design, presented to 20 participants.

The major statistical tests were again performed on the numbers of agreement errors, as in Experiment 1.

Results

The net proportions of agreement errors are presented in Fig. 3 and the raw numbers in Table 5. There were significantly more errors in the mismatch than in the match condition ($F_1(1,79) = 16.90$; $F_2(1,31) = 16.50$). However, this pattern was more clearly in evidence for the phrasal than for the clausal postmodifiers, yielding an interaction between number and postmodifier type that was reliable for participants and marginal for items ($F_1(1,79) = 4.44$; $F_2(1,31) = 3.81$, $p < .06$). There were no other significant effects.

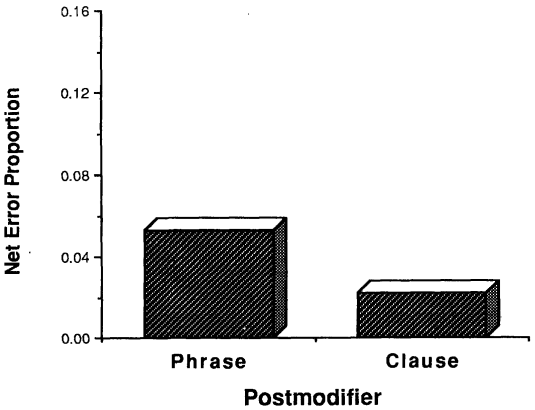

FIG. 3. The net proportions of agreement errors in the mismatch conditions in Experiment 2. In these conditions, the head noun was always singular and the local noun always plural.

TABLE 5
NUMBERS OF RESPONSES BY SCORING CATEGORY AND CONDITION IN EXPERIMENT 2

Number of local noun	Type of postmodifier: Prepositional phrase	Type of postmodifier: Complement clause
Correct responses		
Singular	499	455
Plural	433	436
Agreement errors		
Singular	4	6
Plural	28	16
Uninflected-number responses		
Singular	83	74
Plural	89	68
Miscellaneous responses		
Singular	54	105
Plural	90	120

A breakdown of the agreement errors that occurred after complement-clause postmodifiers revealed that the number of errors also varied according to the number of the complement-clause subject within the preamble. Table 6 gives those data and shows that for both singular and plural local nouns there were more errors after plural subjects (16 total) than after singular sub-

TABLE 6
AGREEMENT ERRORS AFTER SINGULAR AND PLURAL COMPLEMENT CLAUSE SUBJECTS IN EXPERIMENT 2

Number of local noun	Number of complement clause subject: Singular	Number of complement clause subject: Plural
Singular	1	5
Plural	5	11

jects (6 total). The totals differ significantly by a binomial test. For the six items in which the complement verb was inflected, there was no discernible impact of the verb's inflection on the number of errors. When the clause subject was plural and its verb was uninflected, the proportion of errors per item (relative to the number of corrects) increased by .022 compared to when the clause subject was singular and its verb was uninflected. In comparison, when the the clause subject was plural and its verb was inflected, the proportion of errors per item increased by .021 compared to when the subject was singular and its verb was inflected.

There were no significant differences among the conditions for the uninflected-number responses. In the miscellaneous category there were more responses for preambles with clausal than with phrasal postmodifiers ($F_1(1,79) = 19.23$; $F_2(1,31) = 13.24$) and more for preambles with plural than with singular local nouns ($F_1(1,79) = 8.60$; $F_2(1,31) = 4.74$).

Among the miscellaneous responses the proportions of agreement errors were as shown in Table 7. When the head noun was correctly produced in singular form, the results were comparable to those of the main analysis, with proportionally more errors after phrases than clauses, particularly after plural local nouns. However, when the head noun was incorrectly produced as a plural, the pattern was quite different. We will consider this further in the General Discussion.

TABLE 7
PROPORTIONS OF AGREEMENT ERRORS AMONG THE MISCELLANEOUS RESPONSES IN EXPERIMENT 2

Number of local noun	Postmodifier type: Prepositional phrase	Postmodifier type: Complement clause
Singular head		
Singular	.04 (56)	.03 (71)
Plural	.15 (47)	.07 (83)
Plural head		
Singular	.00 (2)	.79 (14)
Plural	.17 (12)	.58 (12)

Note. The entries represent the proportions of errors among all the correctly and incorrectly agreeing number-inflected verbs in the miscellaneous responses. The total number of inflected verbs is shown in parentheses.

On the speaking-span test, the mean total-performance score was 51.2, with a standard deviation of 5.7 and a range from 34 to 63. The correlation between speaking span and the proportion of agreement errors was $r = -.23$, and the correlation between speaking span and the number of miscellaneous responses was $r = -.41$. Both values differ significantly from zero, $t(78) = 2.09$ and 3.97, respectively.

Discussion

As in Experiment 1, agreement errors were more frequent after phrasal than after clausal postmodifiers. This occurred even though the clause type was one in which the verb of the postmodifying clause agreed with a different subject from the verb in the main clause. This further supports the claim that agreement may be specified within clauses before the components of the clauses are arrayed as a sequence of words, in line with the hierarchical hypothesis. Clausal constituency appears to insulate noun phrases from interactions with verbs that are not members of the same clause. Phrases within the same clause are freer to influence the verb of their clause.

Even so, clausal insulation is incomplete. Two pieces of evidence point to an influence of the information in one clause on the verb of another clause. First, there *were* errors after clausal postmodifiers. Second, there was a sizeable difference in the number of agreement errors that occurred after singular and plural complement-clause subjects, with more errors occurring when the subjects were plural. Clearly, the number of the interior clause's subject affected the

verb of the main clause. This would not happen if clauses were planned singly, but some speech error evidence suggests that functional-level processing may encompass roughly two clauses (Garrett, 1980).

A comparison with the correct agreement scores in Experiment 1 reveals that correct agreement was much more common in the present study. This was in turn responsible for a lowering of the proportion of agreement errors. However, the number of agreement errors did not change noticeably; rather, the number of uninflected verbs decreased. We suspect that this was due to the head nouns in the preambles. All the heads in Experiment 2 were abstract, factive nouns (because only factive nouns take complements), whereas those in Experiment 1 included a large number of concrete, mostly animate nouns. Bock and Miller (1991) found that animate subjects were strongly associated with the use of verbs that are uninflected for number in the past tense (such verbs being good descriptors of actions). Abstract nouns, in contrast, elicit many copulas (forms of *to be*) in descriptions of states, and the copula marks number in the past as well as the present tense.

The miscellaneous responses diverged from those in Experiment 1 in an important way. Specifically, there were more such responses when the preambles contained clauses than when they contained phrases. Since preamble reproduction was an immediate memory task and since most of the miscellaneous responses involved reproduction mistakes of various kinds, it appears that complement clauses disrupted immediate memory more than prepositional phrases did. Consistent with this, a number change on the head of the preamble (from singular to plural) was more likely after clauses than after phrases, and the agreement errors that accompanied this change were more frequent after clauses than after phrases, counter to the pattern that prevailed in the main error analysis.

The results of the speaking span test were similar to those of the first experiment, with a stronger association between speaking span and the ability to correctly repeat the preambles than between speaking span and the ability to implement correct agreement. In this experiment, however, unlike the previous one, the correlation between span scores and agreement performance was also reliable.

Experiment 3

In the third experiment we investigated the effect of increasing the length of the postmodifier, in order to further test the limited-capacity memory component of the serial hypothesis. If an increase in the number of words separating the head noun and its verb leads to an overall increase in the occurrence of verb errors, it would argue that agreement procedures may draw on a memory capacity that is sensitive to the quantity of information being manipulated during language performance, regardless of the linguistic status of that information.

The parallel computation implied by the hierarchical account predicts no general length effect on verb errors. To the extent that the complement clause is immaterial to agreement in the main clause, its length should have no impact on agreement in the main clause. So, as in previous experiments, the clause postmodifiers should interfere less with agreement than the phrase postmodifiers, but also be indifferent to clause-postmodifier length. However, lengthened phrase postmodifiers may adversely affect agreement. Because phrase postmodifiers occupy the same processing structure as the target verb, being in the same clause with it (unlike clausal postmodifiers; see Fig. 1), longer phrase postmodifiers increase the size of the clause in which the target verb occurs. Since parallel formulation puts a premium on the avoidance of interference and since increasing the amount of material within a single clause increases the amount of information that must be held in check, there may be more opportunities for errors to arise with

TABLE 8
AN EXAMPLE OF THE VERSIONS OF A SENTENCE PREAMBLE FROM EXPERIMENT 3

Length of postmodifier	Number of local noun	Preamble
Prepositional-phrase preambles		
Short	Singular	The report of the destructive fire
	Plural	The report of the destructive fires
Long	Singular	The report of the destructive forest fire
	Plural	The report of the destructive forest fires
Complement-clause preambles		
Short	Singular	The report that they controlled the fire
	Plural	The report that they controlled the fires
Long	Singular	The report that they controlled the forest fire
	Plural	The report that they controlled the forest fires

longer than with shorter phrase postmodifiers.[2]

We adapted the materials from Experiment 2 for this study, adding an adjective to each phrasal and clausal postmodifier to create a long condition. The resulting preambles are illustrated in Table 8.

Method

Participants. Seventy-two Michigan State University undergraduates participated to fulfill an extra-credit option in introductory psychology classes. All were native English speakers.

Materials. The experimental materials consisted of the preambles used in Experiment 2 and four new versions of each preamble. An example is presented in Table 8. The new versions were created by adding to each postmodifier an adjective that ranged from one to four syllables in length, with a median of two syllables. Four of the Experiment 2 preambles were slightly modified to accommodate these additions. The full set is listed in Appendix A.

[2] This prediction is specific to the verb that follows the local noun. Different predictions would apply to the verb within the complement clause. In particular, the length of the complement clause could have an effect on verb agreement *within* the complement. But because our focus was on the main-clause verb, the form of the complement-clause verb was controlled and dictated as part of the preamble. Most of those verbs were uninflected, so there were few opportunities for agreement errors to arise.

Eight lists of 88 items were formed from these preambles and the fillers used in Experiment 2. The lists were comparable to those of the previous experiment with two minor differences. There was a change in the order of the experimental preambles and there were eight rather than four experimental conditions represented among the experimental preambles in each list. There were an equal number of items from each condition in every list.

The tape-recorded lists were prepared with the aid of a computerized speech-editing system (MacSpeech Lab II, GW Instruments). The fillers and all versions of the experimental preambles were digitally recorded by a male speaker at a sampling rate of 20 kHz and stored on disk. The sound files were edited using a spectrographic display accompanied by audio playback, removing extraneous material from the beginnings and ends of the preambles and shortening long internal pauses. Finally, the preambles were played back in analog form and recorded on audio tape in the orders specified for each list.

Procedure. The procedure was the same as in the previous experiments.

Scoring. Application of the scoring guidelines from the previous experiments yielded 1621 corrects (70.4%), 61 agreement errors (2.6%), 271 uninflected-number responses (11.8%), and 351 miscellaneous responses (15.2%). Of the miscellaneous responses, 81% were reproduction errors.

Design and data analyses. The three factors were (1) the type of postmodifier (complement clause or prepositional phrase), (2) the length of the postmodifier (long or short), and (3) the number of the local noun (singular or plural). Orthogonal combinations of these factors produced eight conditions. In each one, every participant received four preambles, and every preamble was presented to nine participants. Statistical tests were performed as in the first two experiments.

TABLE 9
NUMBERS OF RESPONSES BY SCORING CATEGORY AND CONDITION IN EXPERIMENT 3

Number of local noun	Postmodifier type: Prepositional phrase	Postmodifier type: Complement clause
Correct responses		
	Short postmodifier	
Singular	219	208
Plural	206	205
	Long postmodifier	
Singular	212	194
Plural	190	187
Agreement errors		
	Short postmodifier	
Singular	2	3
Plural	15	7
	Long postmodifier	
Singular	1	1
Plural	27	5
Uninflected-number responses		
	Short postmodifer	
Singular	42	28
Plural	49	24
	Long postmodifier	
Singular	38	27
Plural	33	30
Miscellaneous responses		
	Short postmodifier	
Singular	25	49
Plural	18	52
	Long postmodifier	
Singular	37	66
Plural	38	66

Results

The numbers of responses of each type in each condition are presented in Table 9, and the net proportions of agreement errors are shown in Fig. 4. Looking first at the agreement errors, we see that more of them again occurred when the local nouns were plural than when they were singular ($F_1(1,71) = 25.83$; $F_2(1,31) = 44.61$). There were also more errors after prepositional-phrase postmodifiers than after complement-clause postmodifiers ($F_1(1,71) = 13.10$; $F_2(1,31) = 13.88$).

Most of the errors (69%) occurred after the plural prepositional-phrase postmodifiers, resulting in a significant interaction between plurality and the type of postmodifier ($F_1(1,71) = 13.61$; $F_2(1,31) = 23.89$). The long prepositional-phrase postmodifiers were especially likely to elicit errors, and this is reflected in a significant interaction between the length and the type of postmodifier ($F_1(1,71) = 3.97$; $F_2(1,31) = 5.74$). This tendency was obvious only for plurals, as Table 9 shows, although the interaction of length, postmodifier type, and local-noun plurality did not achieve conventional levels of significance for participants ($F_1(1,71) = 3.42$, $p < .10$) or for items ($F_2(1,31) = 2.57$, $p < .12$). However, planned comparisons indicated that the short–long difference after plural phrases was significant, whereas the short–long difference after plural clauses was not. For these comparisons, the 95% confidence interval was 7.03.

As in Experiment 2, agreement errors after complement-clause postmodifiers were likelier when the subject of the complement clause was plural than when it was singular, 13 errors to 3. This difference was significant by a binomial test. Most of these errors occurred in the short condition, nine after

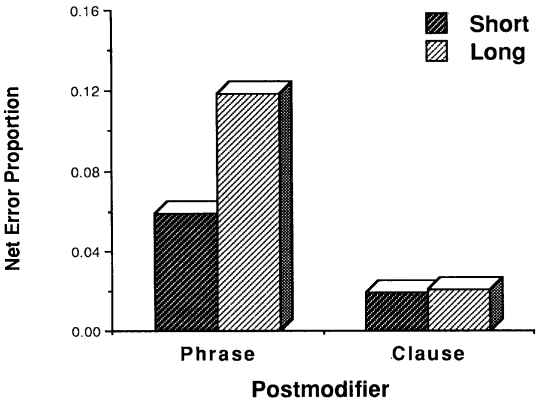

FIG. 4. The net proportions of agreement errors in the mismatch conditions in Experiment 3. The length designations (short, long) indicate the length of the postmodifier.

TABLE 10
PROPORTIONS OF AGREEMENT ERRORS AMONG THE MISCELLANEOUS RESPONSES IN EXPERIMENT 3

Number of local noun	Postmodifier type: Prepositional phrase	Postmodifier type: Complement clause
Singular head		
Singular	.02 (44)	.04 (80)
Plural	.15 (33)	.06 (81)
Plural head		
Singular	.00 (2)	.50 (12)
Plural	.14 (7)	.81 (16)

Note. The entries represent the proportions of errors among all the correctly and incorrectly agreeing number-inflected verbs used in the miscellaneous responses. The total number of inflected verbs is shown in parentheses.

plural subjects and one after a singular subject. A plural inflection on the complement-clause verb, in the four items in which this occurred, slightly magnified the error tendency, in contrast to Experiment 2. When the clause subject was plural and its verb was uninflected, the proportion of errors per item (relative to the number of corrects) increased by .022 compared to when the clause subject was singular and its verb was uninflected. When the clause subject was plural and its verb was inflected, however, the proportion of errors per item increased by .050 compared to when the subject was singular and its verb was inflected.

The completions with uninflected-number verbs were more frequent after prepositional phrases than after clauses ($F_1(1,71) = 10.65$; $F_2(1,31) = 4.30$). In the miscellaneous-response category, there were more responses after long than after short postmodifiers ($F_1(1,71) = 17.48$; $F_2(1,31) = 10.31$) and after clause than after phrase postmodifiers ($F_1(1,71) = 43.57$; $F_2(1,31) = 9.71$). There were no other significant effects in either of these categories.

Table 10 presents the results of the classification of the miscellaneous responses into those with correctly and incorrectly agreeing verbs. The two preamble lengths are pooled because the length differences in the original preambles were not reliably preserved in these responses. The table shows that the agreement error pattern for singular-subject responses was very similar to the pattern in the main analysis. As in Experiment 2, this pattern changed when the speakers produced plural subjects. In these (few) cases, errors constituted a greater proportion of the responses after clausal than after phrasal postmodifiers, and in the match than in the mismatch condition.

The mean total-performance score on the speaking-span test was 49.7, with a range from 34 to 63 and a standard deviation of 5.5. The correlation between these scores and the proportions of agreement errors was a nonsignificant $r = -.14$. The correlation with the numbers of miscellaneous responses was weak but reliable ($r = -.22$; $t(70) = 2.08$).

Discussion

The disadvantage for phrasal interruptions appeared again in Experiment 3, and was magnified when the phrases were longer. So, more separation between the

head noun and verb did lead to more errors, but only after phrasal postmodifiers. Since it was only after phrasal postmodifiers that the agreeing verb was in the same clause with the postmodifying material, the conclusion is similar to the one from the two previous experiments. Clause membership delineates the scope of agreement, and changes in the amount of material within a clause may affect the implementation of agreement between the subject and verb of that clause. However, a change in the amount of material in a clause to which the agreeing verb did not belong, even when that clause immediately preceded the agreeing verb in the spoken sentence, had relatively little impact on its form. This aligns with the predictions of the hierarchical account of production.

The results for the phrasal postmodifiers appear to conflict with Bock and Miller's finding (1991, Experiment 1) that lengthening the separation between the head noun and verb failed to increase the number of agreement errors. In that experiment, half the items had phrasal and half had clausal postmodifiers. We would expect, in light of the present results, that only the lengthened phrasal postmodifiers would yield increases in errors, and inspection of the data in the mismatch conditions in Bock and Miller's experiment reveals such a trend. In fact, the number of errors after longer clauses decreased slightly as the number of errors after longer phrases increased slightly, nullifying the length effect.

In most other respects, the last experiment produced results similar to its predecessors. Again there were more agreement errors after plural than after singular complement-clause subjects and a slightly stronger association between speaking span scores and correct preamble repetition than between speaking span and correct agreement. As in Experiment 2, there were more miscellaneous responses associated with clausal than with phrasal preambles and more number changes on the head nouns that were followed by clausal than by phrasal postmodifiers.

In all three experiments there was a tendency for uninflected-number verbs to follow phrasal more often than clausal preambles, though the tendency achieved full reliability only in this experiment. Our explanation for this trend is the same as one offered by Bock and Miller (1991) to account for an increased use of uninflected-number verbs after short postmodifiers. Most of the verbs used in the completions were past-tense forms, and only forms of *be* are inflected for number in the past tense. The uninflected verbs were therefore semantically rich, so-called lexical verbs like *kissed, jumped,* and *called.* Because of their heavy semantic burden, the use of these verbs is more constrained by the meaning of the subject than is a verb such as *to be.* Accordingly, the uninflected-number verbs are more likely to be used with semantically simple subjects.

General Discussion

Uniformly, the three experiments showed that phrase interruptions interfered with subject–verb agreement more than clause interruptions: Number agreement failed more often when the verb was separated from the head noun by a prepositional phrase postmodifier (as in the error *The report of the destructive fires were accurate*) than when the verb was separated from the head by a complete clause (as in the error *The report that they controlled the fires were printed in the paper*). In Experiment 3, lengthening the interruption by increasing the number of words in the phrase or clause produced no overall increase in the number of errors. Only with phrasal postmodifiers did a longer interruption lead to more errors.

The most straightforward explanation of these results is that clause boundaries delimit and help to insulate the specification of verb agreement, so that the processes responsible for agreement are most sensi-

tive to noun phrases and verbs in the same clause. As depicted in Fig. 1, with a clause postmodifier only the head noun of the main clause was in the same clause as the agreement-target verb (the verb that followed the local noun), whereas with a phrasal postmodifier, both the head and the local noun were in the same clause as the target verb.

The clause-boundedness of agreement specification also helps to explain why agreement errors were no more likely for verbs that followed short clauses than for verbs that followed long clauses. In neither instance was the target verb a member of the clause whose length varied. However, length did matter in the phrase conditions. Notice that in the phrase conditions, the target verb was *within* a clause whose size varied. And in these instances, verb errors were more frequent after longer phrases (that is, when the verb itself was in a larger clause) than after shorter ones (when the verb was in a shorter clause). As the size or complexity of a clause increases, the efficiency of the grammatical processes that operate within that clause may decrease.

With respect to the usual implementation of agreement, one implication of these results is that verb number is specified at some point prior to that at which the verb is actually spoken or put into place in the word string. This is consistent with a hierarchical account of sentence formulation, and contrary to a serial account. If verb number remained open until the point of utterance, we would have expected more pronounced effects of the immediate context of utterance as the distance (either in words or phrasal constituents) to the head noun increased.

An alternative to preplanning is a sophisticated search that looks only at features of already-produced elements that are relevant to the number of the verb in the current clause. Though possible, the latter strategy presupposes a backward-looking mode of operation that seems to have relatively little impact on normal speech.[3] If backward search were common (and it would have to be very common, since verb agreement is demanded in every sentence that a speaker of English ever produces), it would be at odds with the observed dominance of forward-looking (anticipatory) over backward-looking (perseveratory) speech errors.

The correlations between agreement errors and speaking span offer further evidence that the ability to maintain information in working memory during production—an ability that a serial processing system must have in order to create discontinuous dependencies—is at best weakly linked to the normal implementation of agreement. In only one of the experiments, Experiment 2, was the correlation between speaking span scores and agreement performance significantly greater than zero. At the same time, the relationship between speaking span scores and the ability to correctly repeat the preambles, though modest, was reliable in every experiment. The conclusion, in line with the results from the preamble completions themselves, is that immediate memory for prior speech is not the usual source of information for the specification of number agreement.

Still, there may be limits to the endurance of hierarchically prepared information that we did not probe. Because we did not manipulate the number of clauses separating the verb from its head noun, we do not know whether the superior agreement accuracy that we found after clausal separa-

[3] Everyone is familiar with the experience that lends some credence to the backward search account: In the midst of a long and complex sentence, one arrives at a verb whose number is unspecified, prompting a frantic search for the head noun. That this occurs tells us that number specification does fail (as well as err) and that speakers do resort to a memory search when failures happen. It is important to note that failures seem to be rather rare (in comparison to the number of occasions when correctly or incorrectly agreeing verbs are fluently produced) and may be remedied in ways that are unrelated to normal agreement marking.

tions would persist as the number of clauses increased. An experiment reported by Mann (1982, Experiment 7) is relevant to this question, albeit inconclusive. Mann asked his subjects to insert verbs into printed sentences that varied the amount of information between the head noun and the slot provided for the verb. Although Mann was not explicitly concerned with the syntactic characteristics of his materials, in one condition the interruption contained a single clause and in other conditions two or more clauses. There were no reliable or even consistent differences attributable to increasing the number of intervening clauses (in one condition the number of errors actually decreased). However, Mann's written presentation method permitted a much more strategic approach to formulation than is used in normal speech.

Clause Effects in Production and Comprehension

The clause boundedness of agreement is a linguistic fact. By itself, this fact does not determine the organization of language processing: Clauses could define structural boundaries without defining processing boundaries. Yet the present data, with others, suggest that the language processor makes individual clauses a focus of its operations. Because of this focus, the material within a current clause has processing priority over material outside the current clause. In both comprehension and production, the consequence is reduced availability for information prior to or outside of the immediate clause.

This reduction may nonetheless manifest itself in a very different way in language production than it does in language comprehension. In comprehension, there is an array of research showing that listeners experience difficulty when trying to recover the wording of material that precedes the beginning of the most recently heard clause or sentence, even when the recency of the material is equated in terms of numbers of words (Caplan, 1972; Gernsbacher, Hargreaves, & Beeman, 1989; Glanzer, Dorfman, & Kaplan, 1981; Jarvella, 1971; Perfetti & Lesgold, 1977; von Eckardt & Potter, 1985). For instance, in Caplan's (1972) work, listeners heard two-clause sentences such as (a) *Whenever one telephones at night, rates are lower* or (b) *Make your calls after six, because night rates are lower.* The last few words in both sentences (*night rates are lower*) were acoustically identical, spliced into recordings of the sentence beginnings. A probe word (*night*) immediately followed each sentence, and the subjects decided whether it had occurred within the sentence. Although the target word was always the same distance from the probe in the experimental sentences, decisions were consistently faster for the (b) sentences, presumably because the clause boundary preceded the target word in those sentences but followed it in the (a) sentences.

In our production data, in contrast, information on the far side of a clause boundary in the spoken string of words seemed to be *more* accessible than information in a similar serial position within a current clause. This paradox may be a consequence of the different ways in which production and comprehension processes deal with information. If production were a process of ordered word retrieval in the way that comprehension is a process of ordered word apprehension, one might expect similar clause-boundary effects for both. But this ignores the likelihood that the information that begins the production process is formatted differently from the information that begins the comprehension process.

Listeners are slaves to the order in which the elements of an utterance are produced, requiring a sequential attack on constituents and special strategies to deal with potential discontinuities. The minimal attachment heuristic (Frazier & Fodor, 1978; Frazier & Rayner, 1982) is one example of a strategy that comprehenders may employ to integrate something into a developing parse tree without introducing a deep break (e.g., a clause break). If it turns out that

such a break is required, it may be easier to detach the material and restructure it than to maintain it in memory in a raw, unstructured form (Miller, 1951).

In production, the processing organization can be different because of the speaker's firsthand knowledge of the propositional content to be communicated. The number of a verb may be specified before the verb is produced, making it possible for higher-level formulation to proceed on a more nearly clause by clause basis (at least with respect to the planning of clausal constituents and the assignment of syntactic functions), leaving it to lower-level processes to establish the structured sequence of words. The consequence is that the information in one clause can be somewhat insulated from the information in others.

The results for the miscellaneous responses in the last two experiments highlight the apparent disparity in clause processing between comprehension and production. Most of the miscellaneous responses were failures to reproduce the preambles correctly, reflecting failures of immediate memory. In Experiments 2 and 3, there were more miscellaneous responses for clause than for phrase preambles, more failures to remember the number of the head nouns in clause than in phrase preambles, and, following failures to remember the number of the head noun, more agreement errors after clause than phrase postmodifiers. All of these differences are likely consequences of problems in understanding the clausal preambles. Without accurate preamble comprehension, the conceptual representation that supports normal production (and immediate memory performance; Potter & Lombardi, 1990) would have been absent from these responses, leaving only a superficial trace to guide preamble repetition. It is in such circumstances that the clausal-interruption deficit found in the literature most often arises (i.e., when verbatim memory of wording is required), and the deficit is enhanced when the to-be-remembered material is not comprehensible (e.g., when it has syntactic structure but is semantically anomalous; Marslen-Wilson & Tyler, 1976).

All of these miscellaneous-response patterns were absent from the first experiment, in which relative clauses served as the clausal postmodifiers. Notably, relative clauses also appear to have penetrated further into the processing of the main clause than did the complement clauses, judging from a somewhat smaller difference between phrase and clause preambles in agreement error induction (compare the error difference between phrases and clauses for plural mismatch local nouns in Table 2 with the corresponding differences for plural local nouns in Tables 5 and 9). Relative clauses are semantically and syntactically better integrated with their matrix clauses than are complements (since one of the constituents of the relative clause must be identified with a constituent of the matrix clause), and this may demand more integration of processing than between complements and their matrix clauses. Such differences in integrity may help to account for Rochester and Gill's (1973) finding that relatives were produced more fluently than complements.

The difference between relatives and complements reinforces other evidence from the experiments that a verb's protection from the contents of another clause is not complete. In Experiments 2 and 3 we found that the contents of the complement clauses, though less likely to cause interference with the matrix verb than the contents of the matrix clause itself, nonetheless penetrated matrix-clause processing. Specifically, when the clausal postmodifiers had plural subjects, the number of erroneous plural verb forms in the completion increased.

There are two important caveats about the differences between phrases and clauses that we observed in these experiments. We have attributed these differences to the structure of the processing system that supports normal production. How-

ever, because we controlled only the head and local noun phrases, there were meaning differences between the phrase and clause postmodifiers that may have influenced the results. Although the only systematic difference was one that is intrinsic to the distinction between phrases and clauses (i.e., the clauses had a predicate—a verb—and the phrases did not), it is possible that some other semantic disparity between the phrases and clauses led to the contrasting performance patterns. We cannot dismiss this possibility. However, in other work (Bock & Miller, 1991; Bock & Eberhard, 1991) we have uncovered no semantic variables that strongly or systematically influence the occurrence of agreement errors. These variables have included the animacy of the local noun, its concreteness, and even its intrinsic number meaning (as opposed to its grammatical number marking). So, a grammatically singular but semantically plural collective (such as *choir*) is no more likely to create errors than a grammatically and semantically singular individual (such as *singer;* Bock & Eberhard, 1991).

The second caveat is that we had no control over the contents of the completions that the speakers produced, and because of their diversity, we performed no assessment of the contents of those completions beyond our analyses of the verbs. It may be true that those contents influence agreement. Indeed, if the account of production that we have endorsed is correct, it must be true. The danger in this is that differences which we have attributed to features of the preambles (their phrase and clause features, in particular) may be consequences of correlated features of the completions. Until this can be ruled out, our conclusions are necessarily rather inconclusive.

Conclusions

The results suggested that in language production, errors in agreement are more likely to represent a failure to resolve interference between concurrent features of linguistic information than a failure to remember previously produced speech. The major evidence came from a finding that agreement errors were more common when a verb and its controller (the head of the subject noun phrase) were separated by a prepositional phrase than when they were separated by an equivalently long clause. This reverses the pattern that would be expected under extrapolations from tasks that require remembering language, in which material prior to a just-completed clause tends to be less well recalled than material within a just-completed clause. The results are consistent with an account of language production according to which the formulation of an utterance proceeds in hierarchical rather than sequential fashion, with the planning of clauses preceding the sequencing of words. The contrast between comprehension and production can be explained in terms of a basic difference in the problems that these language performance systems must solve. In comprehension, clauses (and the propositions that underlie them) must be put together on the heels of words that are heard one at a time. In production, words must be spoken one at a time on the heels of propositions that are, perhaps, thought all at once.

APPENDIX A: PREAMBLES

Experiment 1 Preambles (Key: Head noun phrase [prepositional phrase/relative clause] local noun phrase; (s) designates location of plurality manipulation):

The demo tape(s) [from the popular/that promoted the] rock singer(s)

The pamphlet(s) [from the British/that described the] government agency(s)

The fire-eater(s) [in the carnival/who enlivened the] sideshow(s)

The bright light(s) [in Doctor Smith's/that lit the small] examination room(s)

The security force(s) [at the giant/that patrolled the] manufacturing plant(s)

The confession(s) [of the famous/that involved the] television preacher(s)

The charismatic leader(s) [of the left-wing/who addressed the] dissident group(s)

The advisor(s) [for the chemistry/who directed the] student(s)
The superintendent(s) [of the technical/who inspected the] school(s)
The memo(s) [from the junior/that concerned the] executive(s)
The neutral zone(s) [around/that split] the Arcturian solar system(s)
The traffic jam(s) [on the Okemos/that obstructed the] street(s)
The office(s) [of the certified/that belonged to the] accountant(s)
The rebel soldier(s) [in the furious/who deserted the] battle(s)
The actor(s) [in the blockbuster/who directed the] film(s)
The consultant(s) [for the growing/who advised the] firm(s)
The teaching assistant(s) [for the physics/who prepared the] lab(s)
The courier(s) [with the diplomat's/who delivered the] message(s)
The star(s) [of the Broadway/who revived the] musical(s)
The corporation(s) [with the banking/that dissolved the] monopoly(s)
The picture(s) [of the prominent/that embarrassed the] politician(s)
The composer(s) [of the modern/who rewrote the] opera(s)
The teacher(s) [with the special/who received the] education certificate(s)
The speaker(s) [at the union/who addressed the] meeting(s)
The director(s) [of the forthcoming/who created the] motion picture(s)
The applicant(s) [for the corporate/who accepted the] scholarship(s)
The editor(s) [of the history/who rejected the] book(s)
The laboratory(s) [with the analog/that invented the] computer(s)
The demonstrator(s) [at the political/that interrupted the] rally(s)
The student(s) [in the Spanish/who had failed the] class(s)
The Peace Corps volunteer(s) [in the African/who assisted the] village(s)
The ruler(s) [of the Roman/who enslaved the] city-state(s)

Experiment 2 Preambles (Key: Head noun phrase [prepositional phrase/complement clause] local noun phrase; (s) designates location of plurality manipulation):

The report [of the destructive/that they controlled the] fire(s)
The dream [about the mystifying/that Anne inherited the] castle(s)
The story [of the homeless/that they lost the] goat(s)
The myth [about the magic/that David wore the] boot(s)
The notion [of the wild roaming/that the man raced the] horse(s)
The suspicion [of the foreign/that they killed the] taxi driver(s)
The tradition [of the Christmas/that she lights the] tree(s)
The announcement [at the weekly football/that the king had banned the] game(s)
The regulation [about the delivery/that drivers examine the] truck(s)
The impression [of the circus/that they liked the] clown(s)
The message [from the excited/that they expelled the] student(s)
The illusion [of the exotic tropical/that the explorer sighted the] island(s)
The view [of the deep enormous/that the river made the] canyon(s)
The thought [of the power/that they paid the] bill(s)
The proposal [for the personal/that Jane program the] computer(s)
The assumption [in the physics/that he caused the] problem(s)
The claim [about the newborn/that wolves had raised the] baby(s)
The custom [at the union/that they hold the] meeting(s)
The fear [of the powerful laser/that the prisoners hid the] weapon(s)
The idea [of the wise and wealthy/that the patient sued the] doctor(s)
The hope [of the wet and hungry/that the army caught the] fugitive(s)
The opinion [of the mysterious/that he should honor the] donor(s)
The belief [about the unusual/that aliens raided the] planet(s)
The saying [about the autumn/that Zeus controls the] day(s)
The fantasy [of the ancient haunted/that they never left the] room(s)
The conclusion [of the long-awaited/that the runner lost the] race(s)
The delusion [about the delicious/that peacocks consumed the] apple(s)
The theory [about the melted/that penguins built the] igloo(s)
The fact [about the plastic/that Gibson caught the] ball(s)
The concept [behind the killer/that comets tow the] satellite(s)
The statement [about the unpopular/that sewage polluted the] lake(s)
The fallacy [about the savage/that lions rule the] jungle(s)

Experiment 3 Preambles (Key: Head noun phrase [prepositional phrase/complement clause] (adjective present = long condition, absent = short condition) local noun phrase; (s) designates location of plurality manipulation):

The report [of the destructive/that they controlled the] (forest) fire(s)

The dream [about the mystifying/that Anne inherited the] (German) castle(s)

The story [of the homeless/that they lost the] (mountain) goat(s)

The myth [about the magic/that David wore the] (riding) boot(s)

The notion [of the wild roaming/that the man raced the] (talking) horse(s)

The suspicion [of the foreign/that they killed the] taxi (cab) driver(s)

The tradition [of the Christmas/that she lights the] (holly) tree(s)

The announcement [at the weekly football/that the king had banned the] (all-star) game(s)

The regulation [about the insecticide/that drivers examine the] (delivery) truck(s)

The impression [of the silly/that they liked the] (circus) clown(s)

The message [from the excited/that they expelled the] (college) student(s)

The illusion [of the exotic Mexican/that the explorer sighted the] (tropical) island(s)

The view [of the deep enormous/that the river made the] (desert) canyon(s)

The thought [of the whopping/that they paid the] (power) bill(s)

The proposal [for the personal/that Jane program the] (IBM) computer(s)

The assumption [in the second/that he caused the] (technical) problem(s)

The claim [about the newborn/that wolves had raised the] (Indian) baby(s)

The custom [at the union/that they hold the] (election) meeting(s)

The fear [of the Soviet laser/that the prisoners hid the] (powered) weapon(s)

The idea [of the wise and wealthy/that the patient sued the] (country) doctor(s)

The hope [of the wet and hungry/that the army caught the] (Bulgarian) fugitive(s)

The opinion [of the mysterious/that he should honor the] (millionaire) donor(s)

The belief [about the unusual/that aliens raided the] (inhabited) planet(s)

The saying [about the rainy/that Zeus controls the] (autumn) day(s)

The fantasy [of the ancient haunted/that they never left the] (dungeon) room(s)

The conclusion [of the long-awaited/that the runner lost the] (Olympic) race(s)

The delusion [about the delicious/that peacocks consumed the] (caramel) apple(s)

The theory [about the melted/that penguins built the] (Alaskan) igloo(s)

The fact [about the plastic/that Magic caught the] (basket-) ball(s)

The concept [behind the killer/that comets tow the] (nuclear) satellite(s)

The statement [about the unpopular/that sewage polluted the] (suburban) lake(s)

The fallacy [about the savage/that lions rule the] (African) jungle(s)

APPENDIX B: SAMPLE CORRECT AND ERROR RESPONSES, EXPERIMENTS 1–3

Experiment 1

Corrects

Prepositional preambles:

The courier with the diplomat's messages was killed by assassins.

The teacher with the special education certificate was fired.

The pamphlet from the British government agencies was informative.

The traffic jam on the Okemos streets was patrolled by a policeman.

Clause preambles:

The composers who rewrote the opera were were yelled at.

The rebel soldiers who deserted the battle were shot for treason.

The speakers who addressed the meetings were long winded.

The Peace Corps volunteer who assisted the villages was named Barbara.

Errors

Prepositional preambles:

The offices of the certified accountants has regular internships.

The rulers of the Roman city-state is Caesar.

The directors of the forthcoming motion picture was of Star Wars.

The demo tapes from the popular rock singers was playing on the radio.

Clause preambles:

The security force that patrolled the manufacturing plant were stringent.

The bright light that lit the small examination rooms were scaring the patients.

The consultants who advised the firms was very helpful.

The laboratory that invented the computers were famous world-wide.

Experiment 2

Correct

Prepositional preambles:

The custom at the union meetings was to sign in first and then do the minutes.

The theory about the melted igloos was not true.

The proposal for the personal computer was not well-written.

The story of the homeless goal was a stupid one.

Clause preambles:

The report that they controlled the fires was difficult to read.

The saying that Zeus controls the day is held by many people that believe in gods.

The fact that Gibson caught the ball was important in the game.

The belief that aliens raided the planets was stirred by Orson Welles.

Errors

Prepositional preambles:

The impression of the circus clowns are happy impressions.

The fear of the powerful laser weapons are understandable.

The report of the destructive fires were accurate.

The statement about the unpopular lakes were that they were polluted.

Clause preambles:

The proposal that Jane program the computer were rejected.

The claim that wolves had raised the babies are untrue.

The report that they controlled the fires were printed in the paper.

The theory that penguins built the igloos were told many years ago.

Experiment 3

Correct

Prepositional preambles:

The theory about the melted Alaskan igloo is probably incorrect.

The concept behind the killer nuclear satellite is a difficult one to understand.

The belief about the unusual planets was true.

The custom at the union election meetings was to drink coffee and doughnuts.

Clause preambles:

The story that they lost the mountain goats was boring.

The regulation that drivers examine the delivery truck was posted.

The statement that sewage polluted the lakes was correct.

The notion that the man raced the talking horses was unbelievable.

Errors

Prepositional preambles:

The story of the homeless mountain goats were told to many children.

The saying about the rainy autumn days are not true.

The opinion of the mysterious donors were just their opinions only.

The statement about the unpopular suburban lakes were that they were beautiful.

Clause preambles:

The assumption that he caused the technical problems were correct.

The regulation that drivers examine the truck are enforced every day.

The claim that wolves had raised the babies were disproved.

The concept that comets tow the satellite are unbelievable.

References

ALLPORT, A. (1989). Visual attention. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 631–682). Cambridge, MA: MIT Press.

BEVER, T. G., & MCELREE, B. (1988). Empty categories access their antecedents during comprehension. *Linguistic Inquiry,* **19,** 35–43.

BOCK, J. K. (1987). Coordinating words and syntax in speech plans. In A. Ellis (Ed.), *Progress in the psychology of language* (pp. 337–390). London: Erlbaum.

BOCK, J. K. (1991). A sketchbook of production problems. *Journal of Psycholinguistic Research,* **20,** 141–160.

BOCK, J. K., & EBERHARD, K. M. (1991). *Meaning, sound, and syntax in English number agreement.* Manuscript submitted for publication.

BOCK, J. K., & MILLER, C. A. (1991). Broken agreement. *Cognitive Psychology,* **23,** 45–93.

BOOMER, D. S. (1965). Hesitation and grammatical encoding. *Language and Speech,* **8,** 148–158.

CAPLAN, D. (1972). Clause boundaries and recognition latencies for words in sentences. *Perception & Psychophysics,* **12,** 73–76.

CARROLL, J. B. (1953). *The study of language.* Cambridge, MA: Harvard University Press.

CLARK, H. H., & CLARK, E. V. (1977). *Psychology and language: An introduction to psycholinguistics.* New York: Harcourt Brace Jovanovich.

COOPER, W. E., PACCIA, J. M., & LAPOINTE, S. G. (1978). Hierarchical coding in speech timing. *Cognitive Psychology,* **10,** 154–177.

DANEMAN, M., & GREEN, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language,* **25,** 1–18.

DEESE, J. (1984). *Thought into speech: The psychology of a language*. Englewood Cliffs, NJ: Prentice–Hall.

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

DELL, G. S. (1987, July). *Phonological encoding in production*. Lecture in the course "Language comprehension and production," Linguistic Society of America Summer Institute, Stanford University, Stanford, CA.

DELL, G. S., & REICH, P. A. (1981). Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior*, **20**, 611–629.

FERREIRA, F. (1991). Effects of length and syntactic complexity on initiation time for prepared utterances. *Journal of Memory and Language*, **30**, 210–233.

FODOR, J. A., BEVER, T. G., & GARRETT, M. F. (1974). *The psychology of language*. New York: McGraw–Hill.

FORD, M. (1982). Sentence planning units: Implications for the speaker's representation of meaningful relations underlying sentences. In J. Bresnan (Ed.), *The mental representation of grammatical relations* (pp. 797–827). Cambridge, MA: MIT Press.

FORD, M. (1983). A method for obtaining measures of local parsing complexity throughout sentences. *Journal of Verbal Learning and Verbal Behavior*, **22**, 203–218.

FORD, M., & HOLMES, V. M. (1978). Planning units and syntax in sentence production. *Cognition*, **6**, 35–53.

FRANCIS, W. N. (1986). Proximity concord in English. *Journal of English Linguistics*, **19**, 309–318.

FRAUENFELDER, U., SEGUI, J., & MEHLER, J. (1980). Monitoring around the relative clause. *Journal of Verbal Learning and Verbal Behavior*, **19**, 328–337.

FRAZIER, L., & FODOR, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, **6**, 291–325.

FRAZIER, L., & RAYNER, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, **14**, 178–210.

FROMKIN, V. A. (1971). The non-anomalous nature of anomalous utterances. *Language*, **47**, 27–52.

GARNSEY, S. M., TANENHAUS, M. K., & CHAPMAN, R. M. (1989). Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research*, **18**, 51–60.

GARRETT, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production* (pp. 177–220). London: Academic Press.

GARRETT, M. F. (1982). Production of speech: Observations from normal and pathological language use. In A. Ellis (Ed.), *Normality and pathology in cognitive functions* (pp. 19–76). London: Academic Press.

GARRETT, M. F. (1988). Processes in language production. In F. J. Newmeyer (Ed.), *Linguistics: The Cambridge survey, III: Language: Psychological and biological aspects* (pp. 69–96). Cambridge, UK: Cambridge University Press.

GEE, J. P., & GROSJEAN, F. (1983). Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology*, **15**, 411–458.

GERNSBACHER, M. A., HARGREAVES, D. J., & BEEMAN, M. (1989). Building and accessing clausal representations: The advantage of first mention versus the advantage of clause recency. *Journal of Memory and Language*, **28**, 735–755.

GLANZER, M., DORFMAN, D., & KAPLAN, B. (1981). Short-term storage in the processing of text. *Journal of Verbal Learning and Verbal Behavior*, **20**, 656–670.

GOLDMAN EISLER, F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. London: Academic Press.

HOLMES, V. M. (1988). Hesitations and sentence planning. *Language and Cognitive Processes*, **3**, 323–361.

HOLMES, V. M., & O'REGAN, J. K. (1981). Eye fixation patterns during the reading of relative-clause sentences. *Journal of Verbal Learning and Verbal Behavior*, **20**, 417–430.

JARVELLA, R. J. (1971). Syntactic processing of connected speech. *Journal of Verbal Learning and Verbal Behavior*, **10**, 409–416.

JESPERSEN, O. (1924). *The philosophy of grammar*. London: Allen & Unwin.

JUST, M. A., & CARPENTER, P. A. (1987). *The psychology of reading and language comprehension*. Boston: Allyn & Bacon.

KEMPEN, G., & HOENKAMP, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science*, **11**, 201–258.

LEVELT, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition*, **14**, 41–104.

LEVELT, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

LINDSLEY, J. R. (1975). Producing simple utterances: How far ahead do we plan? *Cognitive Psychology*, **7**, 1–19.

LOUNSBURY, F. G. (1965). Transitional probability, linguistic structure, and systems of habit-family hierarchies. In C. E. Osgood & T. A. Sebeok (Eds.), *Psycholinguistics: A survey of theory and research problems* (pp. 93–101). Bloomington: Indiana University Press.

MACDONALD, M. C. (1989). Priming effects from gaps to antecedents. *Language and Cognitive Processes,* **4,** 35–56.

MACKAY, D. G. (1982). The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behavior. *Psychological Review,* **89,** 483–506.

MACLAY, H., & OSGOOD, C. E. (1959). Hesitation phenomena in spontaneous English speech. *Word,* **15,** 19–44.

MANN, J. W. (1982). *Effects of number: Experimental studies of the grammatical atmosphere effect.* Johannesburg: Witwatersrand University Press.

MARSLEN-WILSON, W., & TYLER, L. K. (1976). Memory and levels of processing in a psycholinguistic context. *Journal of Experimental Psychology: Human Learning and Memory,* **2,** 112–119.

MILLER, G. A. (1951). *Language and communication.* New York: McGraw–Hill.

NICOL, J., & OSTERHOUT, L. (1989). *Re-activating antecedents of empty categories during sentence processing.* Manuscript submitted for publication.

NOOTEBOOM, S. G. (1973). The tongue slips into patterns. In V. A. Fromkin (Ed.), *Speech errors as linguistic evidence* (pp. 144–156). The Hague: Mouton.

PAIVIO, A., YUILLE, J. C., & MADIGAN, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monograph Supplement,* **76** (1, Pt. 2).

PERFETTI, C. A., & LESGOLD, A. M. (1977). Discourse comprehension and sources of individual differences. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension* (pp. 141–182). Hillsdale, NJ: Erlbaum.

POTTER, M. C., & LOMBARDI, L. (1990). Regeneration in the short-term recall of sentences. *Journal of Memory and Language,* **29,** 633–654.

ROCHESTER, S. R., & GILL, J. (1973). Production of complex sentences in monologues and dialogues. *Journal of Verbal Learning and Verbal Behavior,* **12,** 203–210.

SHATTUCK-HUFNAGEL, S. (1979). Speech errors as evidence for a serial-ordering mechanism in sentence production. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett* (pp. 295–342). Hillsdale, NJ: Erlbaum.

STEMBERGER, J. P. (1985). An interactive activation model of language production. In A. Ellis (Ed.), *Progress in the psychology of language* (pp. 143–186). London: Erlbaum.

STEMBERGER, J. P. (1989). Speech errors in early child language production. *Journal of Memory and Language,* **28,** 164–188.

SWINNEY, D. A., NICOL, J., FORD, M., FRAUENFELDER, U., & BRESNAN, J. (1987, November). *The time course of co-indexation during sentence comprehension.* Paper presented at the meeting of the Psychonomic Society, Seattle, WA.

TANNENBAUM, P. H., WILLIAMS, F., & HILLIER, C. S. (1965). Word predictability in the environments of hesitations. *Journal of Verbal Learning and Verbal Behavior,* **4,** 134–140.

VALIAN, V. (1977). Talk, talk, talk: A selective critical review of theories of speech production. In R. O. Freedle (Ed.), *Discourse production and comprehension* (pp. 107–139). Norwood, NJ: Ablex.

VON ECKARDT, B., & POTTER, M. C. (1985). Clauses and the semantic representation of words. *Memory & Cognition,* **13,** 371–376.

WANNER, E., & MARATSOS, M. (1978). An ATN approach to comprehension. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 119–161). Cambridge, MA: MIT Press.

(Received October 2, 1990)

(Revision received April 17, 1991)