

# **Inducing Agrammatic Profiles in Normals:**

evidence for the selective vulnerability of morphology under cognitive resource limitation

*Journal of Cognitive Neuroscience*, 1995, Volume 7, Number 2

Arshavir Blackwell

Elizabeth Bates

University of California, San Diego

e-mail: arshavir@andrew.cmu.edu

## **Abstract**

The selective vulnerability of morphology in agrammatic aphasia is often interpreted as evidence that closed-class items reside in a particular part of the brain (i.e., Broca's area); thus, damage to a part of the language processor maps onto behavior in a transparent fashion. We propose that the selective vulnerability of grammatical morphemes in receptive processing may be the result of decrements in overall processing capacity, and not the result of a selective lesion. We demonstrate agrammatic profiles in healthy adults who have their processing capacity diminished by engaging in a secondary task during testing. Our results suggest that this selective profile does not necessarily indicate the existence of a distinct sub-system specialized for the implicated aspects of syntax, but rather may be due to the vulnerability of these forms in the face of global resource diminution, at least in grammaticality judgment.

## **Introduction**

Agrammatism is a clinical syndrome that is often found in patients with lesions to a frontal region of the left hemisphere called Broca's area. Patients with this syndrome commonly display difficulty in using the grammatical forms of their language, in both comprehension and production. These patients' utterances tend to be halting, favoring nouns over verbs, and often involving the omission of grammatical functors and the substitution of more canonical or uninflected forms. A typical agrammatic might eliminate words such as "the" or "was," or substitute similar but grammatically incorrect forms. The archetypally difficult sentence for agrammatics to produce is "No ifs, ands, or buts", a sentence made up completely of closed-class (i.e., grammatical function) words. These patients' knowledge of content words and of world knowledge is spared, yet their ability to use even relatively simple syntax in cases where world knowledge is insufficient tends to be impaired (see Caramazza & Zurif, 1976; Heilman & Scholes, 1976; Caramazza, Berndt, Basili & Koller, 1981; Bates, Friederici, & Wulfeck, 1987a). In most patients with this profile of expressive deficits, grammatical problems are

also found in comprehension. A sentence such as 1.1 presents most agrammatic patients with no difficulty in choosing the actor, but 1.2, which has no semantic constraints to guide comprehension, often presents these subjects with difficulty.

Many researchers have argued that the selective vulnerability of particular aspects of grammar consequent to brain damage directly reveals the functional and (by extension) neuroanatomical organization of language; thus, the mapping from surface etiology to underlying architecture is relatively straightforward (the “transparency hypothesis;” Caramazza, 1986; Geschwind, 1972). If subjects have difficulty with a particular syntactic form, then we can postulate that the cognitive system has some sort of module which performs this operation. In arguing for the transparency hypothesis, Caramazza (1986) cites the selectivity of certain neurological dysfunctions (e.g., Hart, Berndt & Caramazza, 1985; Warrington & Shallice, 1984). Agrammatic aphasia has been used as support for a model of brain function wherein Broca’s area is responsible for those aspects of grammar implicated in the agrammatic syndrome. Thus, the traditional clinical view of the syndrome was of a “central syntactic deficit” (Caplan, 1981; Caramazza & Zurif, 1976; Caramazza & Berndt, 1985) in which syntactic knowledge is *lost*, affecting both production and comprehension. That is, in this view agrammatism is “a limitation on language use and language knowledge”, and “when syntactic features are absent on the level of spontaneous speech they are unlikely to be preserved at other levels of language”.<sup>1</sup>

---

1. Zurif & Caramazza, 1976, p. 290

However, later research indicated that agrammatics can make grammaticality judgments with above-chance accuracy, including many of the same sentence types that present serious problems for comprehension (Linebarger, Schwartz & Saffran, 1983; Shankweiler, Crain, Gorrell & Tuller, 1989; Wulfeck & Bates, 1991; Wulfeck, 1987). This finding challenges the central syntactic deficit or “loss of knowledge” account of agrammatic symptoms. The problem is very simple: How can a patient who has lost his syntactic knowledge make accurate judgments of grammaticality, the *sine qua non* of modern linguistic theory? Other problems for central agrammatism come from case studies of patients who display expressive agrammatism but no apparent comprehensive deficit (Kolk, van Grunsven & Guper, 1982; Kolk & van Grunsven, 1984; MacWhinney, Osmán-Sági & Slobin, 1991; Miceli, Mazzucchi, Menn & Goodglass, 1983; Naeser, Haas, Auerbach, Helm-Estabrooks & Levine, 1984; Nespoulos et al., 1988), as well as reports of individuals and groups of patients who display receptive agrammatism but no corresponding expressive deficit (Caramazza et al., 1981; Caplan, 1985; Bates, Friederici, & Wulfeck, 1987a; Bates, Friederici, & Wulfeck, 1987b; Smith & Bates, 1987). Taken together, these various lines of evidence lead to a model of agrammatism in which impaired access and processing operate over a preserved knowledge base (Bates, Wulfeck & MacWhinney, 1991; Friederici, 1988; Prather, Shapiro, Zurif & Swinney, 1991; Wulfeck & Bates, 1991).

A number of investigators have offered a more restricted account of agrammatism called the “closed-class hypothesis” (Bradley, Garrett & Zurif, 1980; Friederici & Graetz,

1984; Grodzinsky, 1993a; Kean, 1979; Zurif & Grodzinsky, 1983). This proposal differs from initial views of agrammatism in two respects. First, the syndrome is restricted to the use of closed-class forms (inflections and function words), with sparing of word order. This partitioning of the grammar can also explain why agrammatic patients produce substitution and omission errors on function words and inflections, while preserving the order in which both open and closed-class items are expressed (that is, patients rarely make errors like “Dog the” for “The dog”, or “ing-walk” for “walking”). Second, most researchers working within this framework have adopted the assumption that knowledge of closed-class elements (i.e. grammatical competence) is preserved in agrammatic patients; the deficit is now viewed as a problem with the access and use of closed-class elements in real time (i.e. grammatical performance—see, for example, Friederici, 1988; Garrett, 1992; Prather, Shapiro, Zurif & Swinney, 1991). Later on we will consider some more recent variants of this view, and the challenge they provide to the position outlined here. For present purposes, it is fair to say that investigators working within the linguistic tradition are willing to assume or (at least) to entertain the possibility that there is a privileged and transparent relationship between the closed-class deficit and the tissue surrounding Broca's area.<sup>2</sup>

---

2. But see Zurif, Swinney, Prather, Solomon & Bushell, 1993, who state that the notion of a central syntactic deficit drives little current research, with syntactically oriented researchers—e.g., Grodzinsky, 1990; Hickok, Zurif & Canseco-Gonzales, 1993; Shapiro, Gordon, Hack & Killackey, 1993—focusing on grammatically detailed *descriptions* of the syndrome, but remaining agnostic as to particular *source* of syndrome: loss of knowledge, global capacity reduction, or more limited and specific capacity reduction.

In the present study, we will take a different approach, arguing against the notion that there is a transparent relation between specific grammatical symptoms and specific lesions sites. To set the stage, we will start by reviewing cross-linguistic data which show that the selective vulnerability of morphology interacts with the informational value of the morphology of the particular language (Bates & MacWhinney, 1987; 1989). In other words, we will show that language-specific grammatical knowledge is not lost in aphasia, even though some components are more vulnerable than others. Second, we discuss the problems inherent in assuming that selective vulnerability of morphology—to whatever extent it holds true—must only result from selective damage to some particular area of tissue, by reviewing a variety of data from normals, non-agrammatic patients and neural network simulations, which point to a possible alternative explanation: that global impairments can result in selective, specific, and seemingly modular deficits.

### **Selective vulnerability of morphology and cross-linguistic data**

One of the difficulties in concluding that morphology is selectively vulnerable in agrammatism is that much contemporary agrammatism research has been carried out in English, making it difficult to separate language-specific aspects of the syndrome (e.g., English is a language with strict subject-verb-object word order and little reliance on morphology) with universal mechanisms (Bates et al., 1991). The data in English are just as consonant with a model where those aspects of grammar that are least important are most vulnerable. Such a model would predict the same

pattern of morphological vulnerability in English as the closed-class theory, but would also predict that a language with rich morphology and little reliance on word order would show the opposite profiles (following predictions arising from the Competition Model; see MacWhinney & Bates, 1989).

The research to date indicates that, while agrammatics do show morphological impairment across languages, they also show profiles that look much like those of normals in their language, albeit noisier. Thus, an Italian agrammatic looks more like an Italian normal than an English agrammatic. For example, Italian and German aphasics produce the correct article 85% of the time, where chance would be 1/9 in Italian and 1/12 in German, a finding difficult to reconcile with the notion of closed-class elements being lost (Bates, Friederici & Wulfeck, 1987a). In a study comparing normal controls and Broca's aphasics in both Italian (a language with richer grammatical morphology than English and freer word order variation) and English (Wulfeck, Bates & Capasso, 1991), accuracy scores for Broca's aphasics in both languages showed the profile word order > agreement. However, for word order, the Broca's aphasics' profile was English > Italian, while for agreement, it was Italian > English. The selective vulnerability of morphology (in this case, comparing agreement with word order) held true in both English and Italian, but interacted with the values of those cues in the two languages. Z-score of decision time for both aphasics and controls showed the same interlinguistic difference. For English-speaking subjects (both normals and aphasics), agreement errors took longer to detect than word order errors, while for the Italian subjects the reverse pattern held. The

superiority of word order to agreement in English for both accuracy and decision time was also found in Wulfeck and Bates (1991). Similar interlinguistic differences were found when subjects were given a simple enactment task called "sentence interpretation" (described below). Italian Broca's retained more sensitivity to agreement than did English speakers, but again there was an interaction of the value of the cue in each language with the particulars of the agrammatic syndrome; in both languages, word order was less impaired than agreement, relative to normals within that language (Bates & MacWhinney, 1987; Bates & MacWhinney, 1989).

Results from studies of agrammatics in highly inflected languages such as Turkish and Hungarian also offer useful information about the selective vulnerability of morphology. Turkish is an agglutinative language, where even telegrams are not "telegrammatic." Turkish Broca's are non-fluent but not telegrammatic. They use noun and verb suffixes correctly. While they use appropriate nominal morphology, they employ a rather limited set of contextually appropriate forms. Word order is spared in these patients, and vowel harmony is maintained across strings of agglutinated morphemes (Slobin, 1991). In a study of another highly inflected language, Hungarian (MacWhinney & Osmán-Sági, 1991), agrammatic errors tended to involve substitution of a close semantic competitor, although often these substitutions were not wrong, but simply less canonical. As with Turkish, Hungarian aphasics displayed no word order or vowel harmony errors. Other researchers have also found that non-fluents in highly inflected languages tend towards agreement errors rather than omission errors (Grodzinsky, 1982;

Lukatela, Crain & Shankweiler, 1988; Miceli, Silveri, Romani & Caramazza, 1989; Smith & Bates, 1987; Smith & Mimica, 1984). As Slobin above, MacWhinney & Osmán-Sági, 1991 conclude that these patients' abilities are damaged and noisy but still relatively functional. Agrammatic speakers of highly inflected languages such as Turkish and Hungarian overall have profiles like normals in their language, but noisier, although they do show damage to case-marking cues as compared to the word order cue in English (MacWhinney et al., 1991).

Although it is increasingly clear that grammatical knowledge is not lost in aphasia, cross-linguistic studies provide support for a dilute, probabilistic form of the closed-class hypothesis. In particular, the pattern of morphological vulnerability with relative sparing of word order does obtain across the languages studied so far, although aphasics do tend to retain the cues of their language and preserve the language-specific ratios of closed-class morphology (Bates et al., 1987a; Bates et al., 1991). Thus, many of the underlying characteristics of Broca's aphasia may obtain cross-linguistically but manifest themselves in ways which interact with the cue values of the language.

### **Does selective vulnerability imply selective loss?**

Assuming that the case for selective vulnerability of morphology has been made, what does this fact imply about localization? Even the most ardent proponent of the transparency hypothesis would agree that there are cases where a spared ability next to an impaired ability does not mean separate and dissociable modules. For example, if a neurological pa-

tient displays difficulty adding four-digit numbers but not two-digit numbers, the most logical conclusion is of a global processing deficit affecting the harder computation, not a separate module for computing four-digit numbers. In evaluating dissociation data, we must somehow distinguish between cases that indicate distinct processes and cases that simply result from more general decrements (e.g., some global processing resource) having a different effect on different tasks. For this reason, neuropsychologists and psychologists often rely on the "double dissociation" as proof of modularity of function. If process A is spared in one patient, but process B is impaired, while a second patient shows impairment of process A but not of process B, many investigators are willing to conclude that A and B are processed by two distinct modules, each with its own neurological base. An example of this logic can be seen in a recent paper by Hillis and Caramazza (1994):

"We report the performance of a neurologically impaired patient who makes far more errors on nouns than on verbs in spoken output tasks, but makes far more errors on verbs than on nouns in written input tasks. This double dissociation within a single patient with respect to grammatical category provides evidence for the hypothesis that phonological and orthographic representation of nouns and verbs are processed by discrete neural mechanisms. Furthermore, the opposite dissociation in the verbal output modality...has also been reported. This double dissociation across patients on the same task indicates that results cannot be ascribed to "greater difficulty" with one type of stimulus, and provides further evidence for the view that grammatical category information is an important organizational

principle of lexical knowledge in the brain.” (Hillis & Caramazza, 1994, p. 2).

As Shallice (1988) has pointed out, such conclusions are not always warranted. For example, drawing upon the distinction between “data-limited” and “resource-limited” processes (Norman & Bobrow, 1975), it is possible to imagine some task “A” that is more vulnerable to data limitations (e.g., decrements in acoustical acuity) and some task “B” that is more vulnerable to resource limitations (e.g., a decline in short-term memory due to aging). Both of these types of decrements are relatively global, yet investigation of these two patient populations using these two tasks would reveal a double dissociation. Indeed, Plaut (in press) has shown that double dissociations can arise through random or gross-grained lesions to a neural network that has no modular structure at all. The danger lies not in the use of double dissociations per se, but in the investigators’ willingness to assume a transparent relationship between performance on two tasks and the regions of the brain responsible for each one. With this in mind, we now review evidence from a variety of subject populations that suggests that the selective vulnerability of morphology may be due in some cases to global resource limitations rather than selective neurological damage.

**Agrammatism in non-agrammatic patients:** Bates and her colleagues have reported evidence for task profiles similar to that of agrammatics in a variety of non-agrammatic patient populations (Bates et al., 1987a). Their task was “sentence interpretation” (SI), an enactment task in which subjects are presented two nouns and a verb (e.g., “The pencil is kicking the cow” or “Are kicking the cows

the pencils”) and instructed to indicate whether the first or the second noun is the actor. Stimuli vary in word order, agreement morphology of the nouns with the verb, and animacy of the nouns (for details on the technique, its applications, and responses to various criticisms, see Bates & MacWhinney, 1987; Bates, McNew, MacWhinney, Devescovi & Smith, 1982; MacWhinney & Bates, 1989; MacWhinney, Pléh & Bates, 1985). The subjects of main interest for our purpose are Italian-speaking Broca’s, anomics, older neurological patient controls, older non-neurological patient controls from the orthopedic ward, and younger controls.<sup>3</sup>

---

3. The anomics were fluent patients suffering from isolated word-finding deficits with clinically normal comprehension. The older neurological patient controls included cases of poliomyelitis and myasthenia gravis, but had no focal lesions, although some of these patients were on central nervous system drugs. The older non-neurological patient controls were from the orthopedic ward, none had any diagnosed nervous system damage, and none were on any central nervous system drugs. This last group’s only similarity to the Broca’s was that both groups were enduring the stresses of a hospital stay.

The Broca's—as expected—were found to be impaired in their use of morphology on the SI task, with word order and semantic cues displaying the normal Italian profile. Figure 1 shows that for young Italian controls, when the verb agrees with the first noun only, the first noun is almost always chosen as the actor, while when it agrees with the second noun only, the first noun is almost never chosen as the actor (recall that these conditions are collapsed over the other factors of word order and animacy). Thus, agreement had a very strong effect for the younger controls in this task. As the flattened profile for the Broca's shows, when the verb agrees with the first noun only, the first noun is chosen as the actor only somewhat above chance (about 67%), while when it agrees with the second noun the first noun is chosen only somewhat *below* chance (about 36%). When the Broca's were compared to the anomics, the older neurological controls, and the older non-neurological controls in three

separate analyses of variance, in all three cases there were no significant interactions with the group factor and no significant main effects of group. Although the older patient controls do appear to be in between the Broca's and the younger controls (Figure 1 also shows the non-neurological, orthopedic patient controls), the group effect is contributed mainly by the younger controls. In comparing just the three control groups (neurological, non-neurological, and younger), the group  $\times$  agreement interaction strongly favors the younger controls. In German, performing a similar comparison of young controls and older non-neurological patients found a similar group  $\times$  agreement interaction favoring the younger controls. Thus, the older German controls were like the German Broca's tested in the same study. The German Broca's, like the Italian Broca's, showed less reliance on agreement morphology than younger controls in their language (with, in the German case, a compensatory reliance upon word order). Results showing similar profiles in comparing Broca's and anomics in Serbo-Croatian have also been reported (Smith & Bates, 1987).

As Bates et al. Themselves caution, there are several caveats to note about these results before concluding that receptive agrammatism is simply due to global stress, whether that be due to the stress associated with neurological damage, hospitalization, or aging. There is no claim that the non-agrammatic patient controls are identical to the Broca's. First, the anomics, neurological patient controls and non-neurological patient controls all showed non-significant trends ( $p = 0.07$ ) towards greater use of animacy cues than the Broca's. Second, inspection of the cell means for the non-neurological control groups showed that, despite

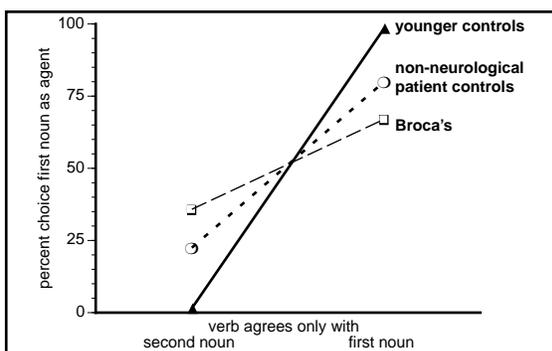


Figure 1: Sentence interpretation. Group  $\times$  agreement interaction for Italian speakers, adapted from Bates, Friederici, and Wulfeck, 1987a. Data are collapsed over the other two factors of word order and animacy. The slope of the line is proportional to the influence of agreement on choosing the agent. As slope increases, the influence of agreement increases. For younger controls, agreement is almost deterministic. For Broca's aphasics, agreement is minimal but extant. For non-neurological (orthopedic) patients, agreement is also impaired compared to younger controls. See text for details.

that lack of a significant interaction with group or a significant main effect of group, this group as a whole did tend to pay more attention to agreement contrasts than Broca's. Third, not all orthopedic, non-neurological patients showed a selective vulnerability of morphology. Five of the ten non-neurological patients looked like the younger controls in their (almost deterministic, as Figure 1 shows) preference for morphology, while the other five non-neurological patients were indistinguishable from the agrammatics. Nonetheless, these results strongly suggest that an alternative—or at least additional—interpretation of receptive agrammatic performance is necessary, beyond the “specific damage site) impaired morphology” explanation, to one in which selective impairment may result from global resource diminution. Consonant with this is the work of Dronkers et al. (Dronkers, Shapiro, Redfern & Knight, 1992) that reports on neurological patients without damage to Broca's area who show agrammatic profiles, as well as patients with damage to Broca's area who do *not* show agrammatic profiles. It is also consonant with reports on the effects of damage to neural networks (McClelland & Rumelhart, 1986; Rumelhart & McClelland, 1986), collections of simple, interconnected units that do not have any distinctive internal structure (i.e., nothing that would seem to correspond to a specific module), yet that can display seemingly modular selective dissociations in the face of global damage (Bullinaria & Chater, 1993; Farah & McClelland, 1991; Harley, 1993; Marchman, 1993).

If the conclusions that we have drawn from these results are valid, then it should be possible to obtain similar “agrammatic-like” profiles in normals by inducing the same

resource-diminished state that we believe the focal lesion patients, the non-aphasic neurological patients, and some of the orthopedic patients to be suffering from. We now briefly review several experiments in which receptive agrammatism has been induced in normals either by diminishing cognitive resources (i.e., a dual task) or by diminishing the availability of information to the subject (i.e., a partial noise mask).

### Inducing agrammatism in normals

King and Just (King & Just, 1991) investigated the differences between high- and low-verbal-capacity subjects (assessed using the reading span task<sup>4</sup> of Daneman & Carpenter, 1980) in comprehension of relative clause sentences, both with and without a concurrent memory load task. In the King and Just experiment, subjects saw sentences that were either subject relative sentences, such as 2.1, where the subject of the main clause, “the reporter,” is also the subject of the relative clause, or object relative sentences, such as 2.2, where the subject of the main clause is the object of the relative clause.

Object relative sentences, which are also harder for aphasics (Caplan, Baker & Dehaut,

**2.1. The reporter that attacked the senator admitted the error publicly after the hearing.**

**2.2. The reporter that the senator attacked admitted the error publicly after the hearing.**

---

4. We also used this technique to account for individual differences. However, the test had little to tell us about the results in these experiments, and for the sake of brevity it is not reported on here.

1985; Hickok et al., 1993), showed greater vulnerability to a concurrent memory task than did subject relative sentences. In addition, low-verbal-capacity subjects were poorer on object relatives than high-verbal-capacity subjects. As King and Just point out, low-span subjects are not globally less accurate, but they are particularly poor on the more computationally demanding object relative sentences.

Miyake, et al. (Miyake, Carpenter & Just, in press) performed a similar experiment in the visual modality, where cognitive stress was induced by having the words presented at a “fast” rate (120 msec/word). Miyake et al. demonstrated that if severity of the simulated aphasic deficit is taken to be a conjoint function of span and rate of presentation, the interaction of severity of “deficit” with the complexity of the sentence type produces profiles quite similar to that of aphasics with similar test materials. In a second experiment, individual differences in normals were explored using cluster analysis, based upon which the subjects were divided into six subgroups, which overall demonstrated double dissociations. That is, the individual groups differed in ways that could not be accounted for strictly as a function of sentence complexity and severity of deficit, in a manner similar to aphasics on the analogous Caplan et al. Task (Caplan et al., 1985). These two studies suggest that global resource deficits may affect comprehension of different sentence types differently, rather than resulting in the same patterns of comprehension deficits over all sentence types, that there may be individual differences in vulnerability to different sentence types, and that the severity of the “deficit” in normals is also affected by the verbal capacity of the subject. Aphasics may not rep-

resent a sharp break from normal processing, but may merely be at the far end of a continuum that includes both high, medium, and low normal processors.

Kilborn (1991) investigated the effects of a low-level noise mask on the SI task for both German- and English-speaking normals. In the no-noise condition, the English speakers’ agent-choice strategies were heavily driven by word order, while the German speakers relied heavily on morphology and semantics. In the noise condition, the English speakers’ agent-choice strategies were unaffected, remaining heavily driven by word order. The German speakers’ agent-choice was affected, as they displayed less use of morphology and more reliance upon word order than in the no-noise condition, similar to the profiles displayed by German Broca’s in the SI task (Bates et al., 1987a, see above).

These studies suggest that the profiles of receptive processing displayed in agrammatism (and other patient populations, as we have seen) can be induced in normals by externally limiting either data or processing ability. The following experiment seeks to extend these results by demonstrating that normal subjects will display agrammatic profiles in grammaticality judgment under a dual task. The experiments on normals discussed above only represent two (highly related) paradigms, in which subjects must map nouns to their proper roles. We also know that agrammatics display difficulty in making grammaticality judgments, compared to normals, but that they are certainly above chance, and that they often are able to judge the grammaticality of a sentence which they cannot completely comprehend (Linebarger et al., 1983; Shankweiler et al., 1989; Wulfeck & Bates, 1991; Wulfeck,

1987). Thus, it is possible that in some way grammaticality judgment is special, particularly resilient, or at least separate in some way from the normal processes in which grammatical knowledge is used to extract meaning. If so, perhaps the effects of global stress on normals will not create the same agrammatic profiles in grammaticality judgment that they do in comprehension. If this is the case, then the profiles that agrammatics display in grammaticality judgment may not in fact be due to global resource difficulties, and we may be forced to a more task-specific (and less satisfying) explanation of the interactions between task, cognitive stress, and specific lesion type. If, however, we do find that cognitive stress in normals causes the same types of agrammatic profiles for grammaticality judgment that it does for comprehension tasks, then this is additional support for the hypothesis that the receptive agrammatic deficit is indeed due to some form of global stress, at least across these two tasks which have been such workhorses for aphasia research in the past twenty years.

## The Current Experiment

### Overview

If our claim that receptive agrammatism is in some cases due to the effects of global stress and not to damage to some particular “closed-class” module, then it should be possible to induce behavior consonant with receptive agrammatism in normal subjects by diminishing working-memory capacity. We have already discussed the greater sensitivity of agrammatics to syntax (i.e., in the case of the grammaticality judgment task, transposition errors) than morphology (i.e., omission and agreement errors; Wulfeck and Bates, 1991). This same difference occurs in their production, in that omission and agreement errors are common in aphasic speech, while word order violations such as “dog the” and morpheme order violations such as “ing-walk” are rare (Bates et al., 1991). This symmetry may come about because the same aspects of the processor implicated in the receptive deficit are also responsible for self-monitoring (see Discussion section, below).

With this in mind, we designed our stimuli to test these same contrasts. We presented subjects with auditory sentences containing one of three different error types: TRANSPOSITION errors (e.g., “She is selling books” ⇒ “She selling \* is books”), OMISSION errors (e.g., “She selling \* books”) or AGREEMENT errors (e.g., “She are \* selling books.”). One group hears these sentences in a baseline, single-task condition, the others are exposed to a simultaneous memory-load task, at varying levels of difficulty. We predicted that this memory-load manipulation would not cause an equal decrement in accuracy across all error types, but that errors would be differentially affected in a pro-

file that resembles the patterns of vulnerability displayed by agrammatic patients. Specifically, agreement errors would be most vulnerable, followed by errors of omission, with the best performance displayed on errors of transposition.

These stimuli were quite similar to those used in an earlier grammaticality judgment study (Blackwell, Bates & Fisher, 1993) with normals in the visual domain, which, in turn, were based upon the Wulfeck and Bates (1991) auditory grammaticality judgment study. The Wulfeck et al. (1991) study, which only employed transposition and agreement errors, demonstrated that normal English listeners are faster at detecting transposition errors than agreement errors.

The Blackwell et al. study, which in addition employed omission errors, discovered differences between grammaticality judgment in the auditory and visual domains, but again found that transposition errors were superior to other error types in that normal English speakers showed the greatest accuracy in detecting them<sup>5</sup>. This study also demonstrated significant differences between errors that occur early in the sentence and those that occur later, at least over these stimuli. As the sentence progresses and information builds up, subjects seem to be less willing to entertain alternative potential sentence completions and they decide sooner after the error point that the sentence is ungrammatical. In other words, earlier in the sentence the subject's structural

---

5. This study differs from Blackwell et al., 1993 primarily in that the current paper investigates (in the *auditory* domain) the effect of *global resource diminution* on response to these types of stimuli in an *on-line* task, while Blackwell et al., 1993 investigated (in the *visual* domain) processing of these stimuli under normal (i.e., unstressed) conditions, using a combination of on-line and off-line techniques.

representations are less definite and well-formed, and as a consequence these early errors may be more vulnerable to our cognitive stress manipulation.

TO SUMMARIZE OUR PREDICTIONS FOR PERFORMANCE BY NORMAL SUBJECTS UNDER STRESS: (1) Accuracy to transposition errors will be less affected by the stress manipulation than will omission and agreement errors; (2) This effect will be focused in errors that occur early in the sentence, while late errors will either not show the effect or will show it in a dilute form.

The experiment involves imposition of a secondary task, in this case keeping a series of digits in memory. Subjects listen to a sentence and are instructed to judge the grammaticality of the sentence, pressing the appropriate key ("good" or "bad") as soon as they make their decision—even if the sentence is still running. In the digits condition, subjects see either two, four, or six "target" digits on the screen in front of them (only one number of digits per group, throughout the experiment). Immediately after the to-be-memorized digit string, an individual sentence is presented auditorily. After the sentence is over and they have made their grammaticality judgment, they then see another series of digits (the same number as in the target set), and are asked to decide whether this string is the "same" or "different" from the sequence presented before the sentence. In cases of a mismatch, the pre-sentence and post-sentence digit strings only differ by one number. Thus, in the three different digit conditions (two, four or six numbers), all subjects are forced to keep unrelated and arbitrary ma-

**Table 1: Criteria for Rejection from No-Digits Condition**

measure	cut-off	number dropped
reaction time:		
ungrammatical	1661 ms	0
grammatical	4185 ms	2
A'	95.09	2
hit	89.29	1
false alarm	12.28	0
no response	—*	0
* Subjects ranged from 0 to 2 “no responses” for the entire experiment, with one subject as high as 4.		

**Table 2: Criteria for Rejection from Digits Conditions**

measure	cut-off	number dropped
sentences:		
RT ungrammatical	1766 ms	3
A'	93.27	6
hit	84.15	1
false alarm	11.18	3
digits:		
percent correct	52.21	7
reaction time	2809 ms	0
No responses	16	7

terial in memory while they were making their grammaticality judgment.

Digit load (0, 2, 4 or 6 digits) was treated as a between-subjects variable. The various linguistic manipulations are all within-subject variables (error type—agreement, omission and transposition; part of speech—auxiliary vs. determiner; location of the error—early vs. late in the sentence).

## Results

### Subjects dropped

**Without digits:** Seven subjects were dropped based upon the “two standard deviations from the mean” criteria displayed in Table 1. An additional subject was dropped for not following directions during the experiment. New subjects were re-run to take their place in the design, bringing the total in this condition to 28.

**With digits:** Twenty-seven subjects were dropped based upon the “two-standard deviations from the mean” criteria displayed in Table 2. An additional subject was dropped for not following directions during the experiment. New subjects were re-run to take their place in the design, bringing the total to 28 for each of the three digit conditions.

**Overall scores:** Table 3 shows the overall A' for subjects in all non-filler sentences, with no sentences removed due to error on the digits task. It also shows the overall percent correct to digits for all ungrammatical sentences, with no data points removed due to error on the grammaticality judgment task. In general, performance on the judgment task was very good, but it varied over conditions, as follows.

### No-digits Condition Alone

First we examine accuracy without stress, to ascertain if there are any significant differences between error types in the baseline case. The 28 subjects' data were submitted to a  $2 \times 2 \times 3$  analysis of variance, with three within-subject variables (location, part of speech, and type of error). Subjects was the random factor. All the main effects were significant, as were 2 two-way interactions: location  $\times$  type and part of speech  $\times$  type. Auxiliary errors were more accurately detected than determiners, and late errors were better than early errors. A Newman-Keuls post-hoc on the main effect of type showed a selective vulnerability for agreement errors, with transpositions and omissions not significantly different from each

other (i.e., transposition = omission > agreement).

We next explore the two significant interactions, in which we find that they are due again to the selective vulnerability of agreement interacting with location (early agreement errors are more vulnerable than late agreement errors) and part of speech (determiner agreement errors are more vulnerable than auxiliary agreement errors). Figure 2 shows a graph of the type effect separated by location of error. Newman-Keuls post-hocs at each level of location showed the same pattern of significant results as the main effect: transposition = omission > agreement. The significant interaction indicates that the vulnerability for agreement errors is greater for early agreement than late agreement errors. Thus, this pattern of accuracy can be summarized as transposition = omission > late agreement > early agreement.

Next we turn to the part of speech  $\times$  type interaction, where we find a parallel pattern of results, with agreement selectively vulnerable at both levels of part of speech. A Newman-Keuls post-hoc analysis of error type at each level of part of speech showed the same pattern of significant results as the main effect: transposition = omission > agreement. The significant interaction indicates that the vul-

**Table 3: Overall A' to Sentences and Percent Correct to Digits as a Function of Digit Load**

	number of digits			
	0	2	4	6
sentences	98.0	97.5	97.6	96.8
digits	—	92.4	88.4	71.0

nerability for agreement errors is greater for determiner agreement than auxiliary agreement errors. Thus, this pattern of accuracy can be summarized as transposition = omission > auxiliary agreement > determiner agreement.

All of these results point to the same conclusions. Even in normal (i.e., no dual task) grammaticality judgment, agreement errors are more vulnerable than either transposition or omission errors. This selective vulnerability is heightened for early errors and for determiner errors. Although the three-way interaction did not reach significance, the pattern of vulnerability lead us to perform two additional sets of Newman-Keuls comparisons. First, we compared the four levels of location  $\times$  part of speech (i.e., early auxiliary, early determiner, late auxiliary, late determiner) at each level of type. These four means were only significantly different for agreement errors. The pattern of means was late auxiliary > early auxiliary > late determiner > early determiner (see Figure 3). Late auxiliary errors were detected with significantly greater accuracy than both early and late determiner errors, and early determiner

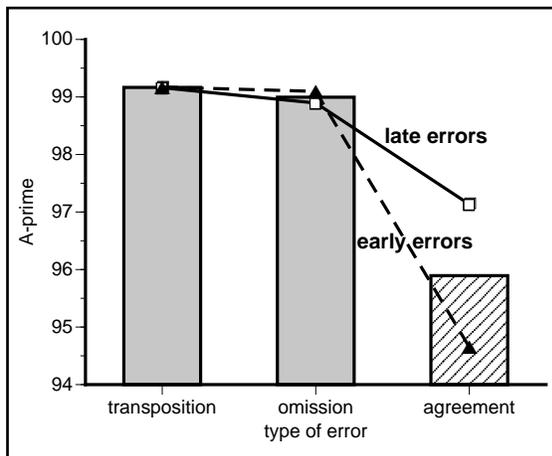


Figure 2: A' Scores without Dual Task. By error type (bars) and by error type and location. Transposition = omission > late agreement > early agreement.

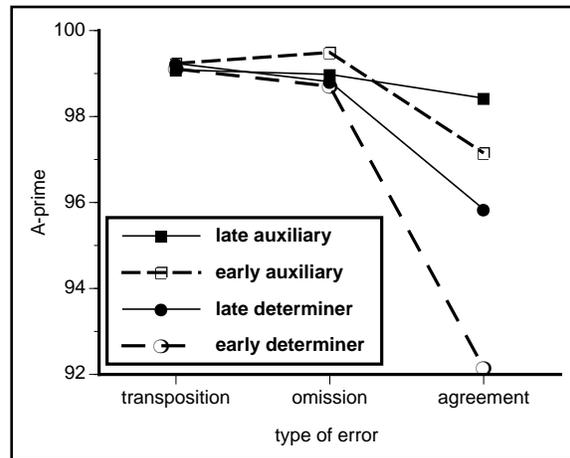


Figure 3: A' Scores without Dual Task. By error type, location, and part of speech. The differences within transposition and omission are not significant. The type difference for late auxiliary errors is not significant.

**For agreement errors:** late auxiliary agreement errors are significantly more accurate than early and late determiner agreement errors, early determiner agreement errors are significantly less accurate than the other three agreement errors, and early auxiliary errors are not significantly different from either late auxiliary or late determiner agreement errors.

er errors were detected with significantly less accuracy than all of the other three means. Early auxiliary errors, while numerically less than late auxiliary errors and greater than late determiner errors were not significantly different from either. Second, we analyzed the type effect at each of the four levels of location  $\times$  part of speech. We found the transposition = omission > agreement pattern to hold at 3 of these 4 levels, while for late auxiliary errors there was no significant effect of error type.

In summary, in on-line auditory grammaticality judgment for normal native English speakers, agreement errors are selectively vulnerable. This vulnerability was not demonstrated for late auxiliary errors. Changes in both location and part of speech have a cumulative, diminishing effect on the detection of agreement errors.

## No-digits and dual task conditions together

Our earlier work in the visual domain, and now these results in the auditory domain, strongly suggest that the structural differences between early and late errors may interact in quite different ways with the other experimental factors. Therefore, we will break our analysis of the dual task results into two separate analyses, one for early errors, and one for late. Each analysis has one between-subjects factor (digit load) and two within-subjects factors (part of speech and type of error). Subjects was the random factor. We also report on a parallel set of analyses with sentences as the random factor (for accuracy, our item analyses are carried out with percent correct as the dependent variable, rather than  $A'$ , as the logic of  $A'$  is difficult to apply over items<sup>6</sup>).

### Accuracy ( $A'$ and percent correct)

**EARLY ERRORS** : the following main effects were significant: digit load ( $F(3,108) = 8.01, p < 0.0001$ ), part of speech ( $F(1,108) = 108.19, p < 0.0001$ ), and type of error ( $F(2,216) = 280.61, p < 0.0001$ ). Auxiliaries ( $A' = 98.0$ ) were detected with significantly greater accuracy than determiners ( $A' = 95.2$ ). A Newman-Keuls post-hoc test on the significant effect of digit load over subjects is summarized in Figure 4. A Newman-Keuls post-hoc test on the significant effect of type showed transposition = omission > agreement. Unsurprisingly, both the part of speech and

6. The  $A'$  score is an index of accuracy designed to correct for response bias. On both psychological and mathematical grounds, this means that the  $A'$  score can only be analyzed over subjects (i.e. treating subjects as a random variable and items as a fixed effect), and not over items (i.e. treating items as a random variable and subjects as a fixed effect).

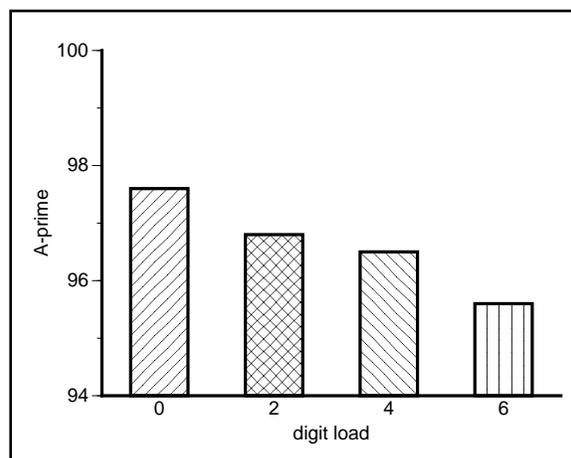


Figure 4:  $A'$  to Early Errors Only. By digit load. Bars sharing internal lines of the same orientation are not significantly different. The pattern is  $0 > 4 > 6$ , with  $2 > 6$  but not significantly different from 0 or 4.

type main effects mirror the pattern of results in the 0-digit-load condition. The type and part of speech effects were also significant over items, with identical profiles.

The one significant interaction was part of speech  $\times$  type ( $F(2,216) = 42.72, p < 0.0001$ ; this interaction approached significance over items,  $p < 0.09$ ). To investigate the part of speech  $\times$  type interaction, Newman-Keuls post-hocs were performed on type at each level of part of speech. The pattern of transposition = omission > agreement holds at both levels of part of speech, but is significantly sharper for determiner errors than for auxiliary errors (both over subjects and over items). In addition, the auxiliaries were detected with significantly greater accuracy than determiners at each level of type.

Because of our particular hypothesis, we conducted 3 sets of further analyses, one at each level of type. Each analysis consisted of a one-factor analysis of variance with digit load as the factor, followed by a Newman-Keuls post-hoc test in the case of a significant

result. As Figure 5a shows, the result was significant for all three error types, with agreement errors affected by the digit load at only 2 digits, omission errors affected at 6 digits, and transposition errors less affected, but still affected by 6 digits.<sup>7</sup> In order to emphasize that it is not just the type means that are different, but that agreement is more affected by the digit load task than omission, and omission more than transposition, we have redrawn this graph with the baseline (no digits) value at each type subtracted out. Figure 5b shows the decrement in performance as a function of digit load, starting at zero as baseline. Agreement shows the sharpest, earliest drop, omission shows a less steep decline, and transposition shows the least decline of all three error types.

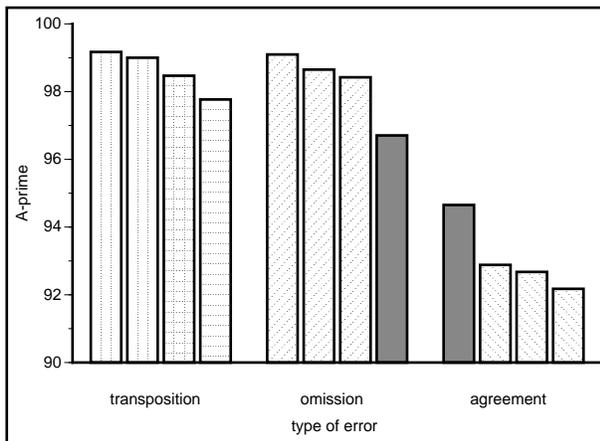


Figure 5a: A' to Early Errors. By type and digit load. Bars sharing internal lines of the same orientation are not significantly different; these comparisons are within-type only. The patterns are:

transposition: 0 = 2 > 6; 4 is not significantly different from any condition

omission: 0 = 2 = 4 > 6

agreement: 0 > 2 = 4 = 6

7. The patterns were: agreement, 0 > 2 = 4 = 6; omission, 0 = 2 = 4 > 6, and transposition, 0 = 2 > 6, with 4 not significantly different from either 2 or 6.

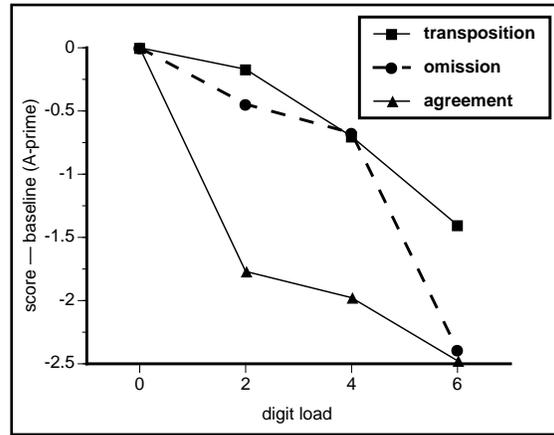


Figure 5b: A' Difference Scores to Early Errors. By type and digit load. Graph depicts score at each level of digit load with baseline for that error type subtracted out.

A more fine-grained analysis showed that the agreement effect was strongly (though not entirely) contributed by the 4 early agreement items that were more than 2 standard deviations from the item mean.<sup>8</sup> Three of these were the three determiner agreement items of the form “A [plural noun] were/are...”, while the fourth was item #5.1, which began, “The guests was \* ...” At least for the three determiner agreement items it is clear that the relatively low-salience indefinite article is particularly vulnerable. However, a re-analysis with these four particularly vulnerable items removed still showed agreement vulnerable at 2 digits, compared to the baseline condition, although of course overall performance in this cell was much better with those items removed. Thus, although these 4 particular (and legitimately termed) early agreement items are helping to push accuracy down, even without them agreement still showed vulnera-

8. Averaged over all 4 digits conditions, they were #5.1 at 28.5%, #7.3 at 50.4%, #7.7 at 18.3% and #7.9 at 20.0%.

bility at fewer digits than the other two error types.

We may summarize these findings for accuracy to early errors as follows:

- the effect of digit load is agreement > omission > transposition
- auxiliary > determiner (over subjects and items)
- transposition = omission > agreement (over subjects and items)
- selective vulnerability of agreement interacts with the greater vulnerability of determiners—i.e., the type effect (transposition = omission > agreement) held at both levels of part of speech, but was sharper for determiner errors (over subjects and items)

**LATE ERRORS:** The following main effects were significant: part of speech ( $F(1,108) = 16.23, p < 0.0001$ ) and type ( $F(2,216) = 31.29, p < 0.0001$ ). Auxiliary errors ( $A' = 98.8$ ) were detected with significantly greater accuracy than determiner errors ( $A' = 97.7$ ). These results also held over items. A Newman-Keuls post-hoc test on the significant effect of type of error showed that, as for early errors, transposition = omission > agreement, over both subjects and items.

As with early errors, only part of speech  $\times$  type ( $F(2,216) = 27.64, p < 0.0001$ ; also significant over items) was significant. We examined the part of speech  $\times$  type interaction by performing Newman-Keuls post-hoc tests on type at each level of part of speech. The type

effect was significant for determiner errors, but not for auxiliary errors (over subjects and over items), with the pattern transposition = omission > agreement.

In summary, as for early errors, for late errors *auxiliary* errors were detected with significantly greater accuracy than determiner errors, and transposition and omission errors were detected with significantly greater accuracy than agreement errors. The type effect was contributed by determiner errors (type was not a significant effect for auxiliary errors).

**DIGIT LOAD  $\times$  LOCATION:** The effect of digit load on accuracy is concentrated entirely in the early errors. Our previous work in the visual domain demonstrated strong differences between early and late error types, including greater certainty and “wrap-up” effects for late errors. Thus, this “selective sparing” of late errors should come as little surprise -- but its implications for the interpretation of clinical symptoms is still interesting.

**OVERALL SUMMARY FOR ACCURACY:**

We may come away from this wide array of analyses with the following five basic findings:

1. The effect of digit load is restricted to early errors; later errors appear to be “buffered” by information build-up.
2. Subjects are more accurate in general at detecting transposition errors and omission errors than agreement errors, both over subjects and over items.
3. The effect of digit load is greater for

agreement errors than omission errors, and greater for omission errors than transposition errors, although these data demonstrate a point at which omission and agreement errors converge (this point depends upon on the mix of items used).

4. Subjects are more accurate at detecting auxiliary errors than determiner errors, both over subjects and over items.
5. For late errors, the selective vulnerability of agreement was contributed by determiner errors, not auxiliary errors.

**Reaction time :** Earlier work with very similar stimuli (Blackwell et al., 1993) indicated to us that reaction time points between the three error types might vary widely, as the reaction time depends not just upon the time that subjects take to make their decision, but also the point at which subjects perceive the error to begin. As we pointed out in our discussion of the stimuli (see Method section) certain error types—most notably agreement errors—tend to be associated with very crisp and definite error points, while other error types—omissions, for example—tend to show much less distinct error points, both over subjects and over items. Thus, it is impossible to anchor all reaction times to a single error point. For this reason, we analyzed reaction times in 6 separate analyses, divided by location and type of error. There was one between-subjects factor (digit load) and one within-subjects factor (part of speech). Subjects was the random factor. An effect of digit load was

found for all 3 early error analyses, and for late auxiliary errors. The analyses may be summarized as follows (including a parallel set of analyses over items):

**Early errors:** For early transposition errors, both part of speech ( $F(1,108) = 25.80$ ,  $p < 0.0001$ ; also significant over items) and digit load ( $F(3,108) = 7.67$ ,  $p < 0.0001$ ; also significant over items) were significant. The pattern for part of speech was auxiliary > determiner, while a Newman-Keuls post-hoc test for digit load showed that 4 digits was the lowest digit condition that was significantly different from the 0-digit condition (see Figure 6). For early omission errors, both part of speech ( $F(1,108) = 13.47$ ,  $p < 0.0004$ ; also significant over items) and digit load ( $F(3,108) = 4.53$ ,  $p < 0.005$ ; also significant over items) were significant. The pattern for part of speech was auxiliary > determiner, while a Newman-Keuls post-hoc test for digit load showed that 6 digits was the lowest digit condition that was significantly different from the 0-digit condition (see Figure 6). For early agreement errors, both digit load ( $F(3,108) = 7.16$ ,  $p < 0.0002$ ; significant over subjects only) and part of speech  $\times$  digit load ( $F(3,108) = 3.59$ ,  $p < 0.02$ ) were significant. A Newman-Keuls post-hoc test for digit load showed that 6 digits was the lowest digit condition that was significantly different from the 0-digit condition. This profile was sharper for auxiliary errors, though type was significant at both levels of part of speech (see Figure 7).

**Late errors:** Part of speech was significant for late transposition errors ( $F(1,108) = 205.58$ ,  $p < 0.0001$ ; also significant over items), for late omission errors ( $F(1,108) = 379.13$ ,  $p < 0.0001$ ; also significant over items), and for late agreement errors ( $F(1,108)$

= 126.77,  $p < 0.0001$ ; also significant over items). In all three cases, the profile was auxiliary > determiner. In addition, for late agreement errors, digit load ( $F(3,108) = 4.64$ ,  $p < 0.005$ ; also significant over items) and part of speech  $\times$  digit load ( $F(3,108) = 2.90$ ,  $p < 0.04$ ; significant over subjects only) were significant. A Newman-Keuls post-hoc test for digit load showed that 4 digits was the lowest digit condition that was significantly different from the 0-digit condition. This profile was sharper for auxiliary errors, though type was significant at both levels of part of speech (see Figure 7).

In summary, determiner errors were detected significantly more quickly than auxiliary errors, both over subjects and over items, in five of the six analyses—all but the early agreement analysis. In that analysis, while there was no main effect of part of speech, digit load and part of speech did interact in such a way that auxiliary agreement errors were more

affected by the digit load than were determiner agreement errors. Digit load significantly affected all three *early* error types, transposition, omission, and agreement, but for *late* errors, digit load only had a significant effect on agreement errors. Auxiliary errors seem to take longer to detect (in general) than do determiner errors, early errors seem to be much more affected by the digit load manipulation than do late errors, and agreement errors again seem to show a selective vulnerability in that only agreement errors are affected by the digit manipulation in the late condition.

**Percent correct to digits:** For this dependent variable, we performed an analysis of variance with one between-subjects factor (digit load, with 3 levels: 2, 4, and 6 digits) and three within-subjects factors (location of error; type of error, and part of speech). Subjects was the random factor. Part of speech ( $F(1,81) = 17.85$ ,  $p < 0.0001$ ), digit load ( $F(2,81) = 55.77$ ,

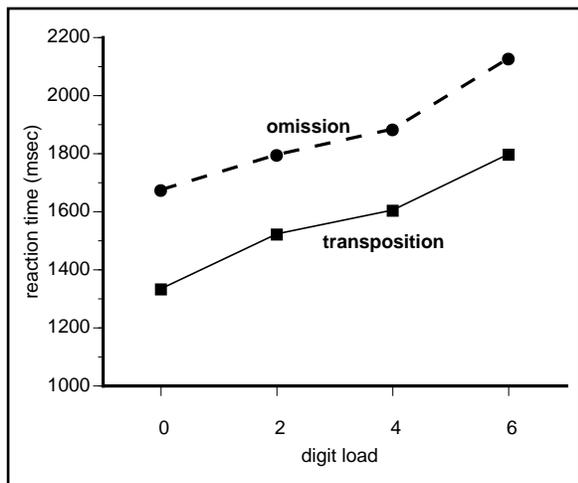


Figure 6: Reaction Times for Early Errors. For transposition and omission errors.

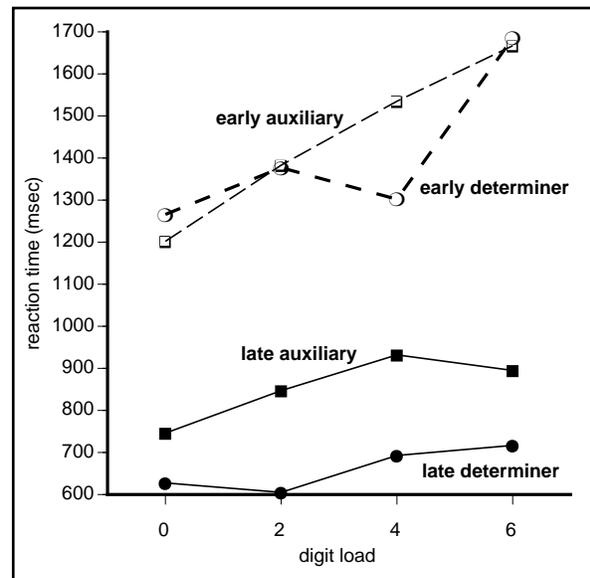


Figure 7: Reaction Times for Agreement Errors. By location and part of speech.

$p < 0.0001$ ) and location  $\times$  part of speech ( $F(1,81) = 7.32, p < 0.009$ ) were significant (also over items). For part of speech, the profile was auxiliary  $>$  determiner. Newman-Keuls post-hoc tests for digits showed the 2-digit condition = 4-digit condition  $>$  6-digit condition. Newman-Keuls post-hocs investigating the significant interaction showed the part of speech effect (auxiliary  $>$  determiner; see Figure 8) held for late errors only (over subjects and over items), and the location effect (early  $>$  late) held for determiner errors—over subjects only.

The results are partially suggestive of a correlation between grammaticality judgment and the digits task, such that the same items that are hard for grammaticality judgment are also hard for the digits task. Subjects are overall more accurate at detecting grammatical errors, and at the digits task, for auxiliary errors compared to determiner errors, in both the early and late conditions. The pattern early  $<$  late is seen for auxiliary errors for both error detection and the digits task. However, for determiner errors, there does appear to be a trade-off between error detection and the digits task, with subjects more accurate at error detection but less accurate at the digits task for late errors, although this pattern for digits of early  $>$  late was significant over subjects but not over items: the effect, such as it was, was small (see Figure 8). In addition, there was a significant positive correlation (over items) between

grammaticality judgment and the digits task (Pearson correlation coefficient = 0.20,  $p < 0.002$ ). Because we were not interested in the digits task *per se*, but only as a way to diminish performance on grammaticality judgment, we had no strong theoretical predictions about the particulars of this task and will have no more to say about it here.

In summary, subjects are in general more accurate with the digits task when the errors that they are monitoring are auxiliary errors, rather than determiner errors, an effect that appears to come from the late errors. The digit load manipulation, not surprisingly, did have an effect on subject performance (i.e., the more digits subjects had to remember, the less well they performed at remembering digits).

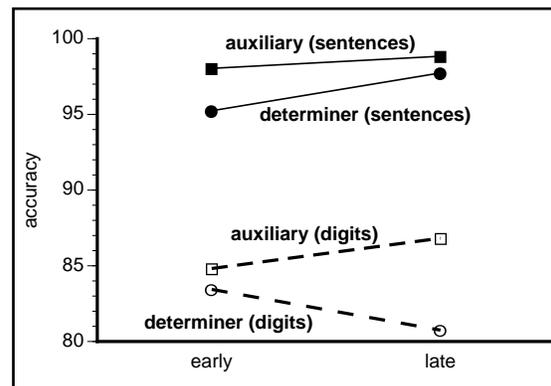


Figure 8: Percent Correct to Digits and A' for Sentences. Location  $\times$  part of speech.

## Summary and Discussion

The final discussion is divided into three parts. First, we present a brief summary of these complex results. Second, we will consider an account for all these findings within a psycholinguistic performance theory called the Competition Model (Bates & MacWhinney, 1989). This part of the discussion will emphasize issues surrounding (1) the “cue validity” or information value of linguistic structures, (2) the “cost” or saliency of linguistic cues, (including some parallels between our findings and results obtained with simple recurrent neural networks), (3) whether omission errors are a separate error type in their own right, or merely a special type of substitution error, (4) the relation between expressive and receptive deficits. Third and last, we will consider some alternative proposals within a different framework, a linguistic competence approach based on a variant of generative grammar called Government and Binding Theory, with recommendations for future studies that could help to decide among these alternatives.

### Summary of Findings

**Baseline performance profile (no dual task):** Subject sensitivity to error type was transposition = omission > agreement. For the location  $\times$  error type interaction, the pattern was transposition = omission > late agreement > early agreement (see Figure 2). For the part of speech  $\times$  error type interaction, the pattern was transposition = omission > auxiliary agreement > determiner agreement (see Figure 3). Thus, even under normal conditions we found agreement errors to be particularly vulnerable, especially early agreement and determiner agreement.

### Stressed performance profile (dual task):

Profiles of sensitivity under cognitive stress were much as predicted: transposition > omission > agreement, seen in early errors only. Early agreement errors showed an immediate accuracy drop at 2 digits, early omission errors at 6 digits, and early transposition errors showed a much smaller drop than omission errors at 6 digits (see Figures 5a and 5b). In other words, the three error types differ in their degree of vulnerability, a conclusion that is supported by performance in the baseline condition and by the clear differences that we observe in the “breaking point” when each item type is subjected to stress. This is the most important finding in the current experiment, demonstrating a clear link between the error profiles shown by agrammatic aphasics (in both comprehension and production) and the vulnerability profile displayed by normals under different degrees of cognitive overload. However, we must also note that the digit task had no effect upon late errors. It seems that the build-up of information across the course of the sentence is sufficient to “buffer” normal listeners against the effects of stress.

There is no evidence for a direct trade-off between the judgment and digit tasks. That is, subjects were *not* poorer at agreement and omission errors because (for some reason) they chose to do better at the *digits* task for those particular item types. This fact bolsters our confidence in use of the digits task to simulate the effects of a cognitive resource reduction.

All of our effects interacted with part of speech, which means that it is misleading to talk about “error types” as homogeneous categories. In both the baseline and dual task conditions, auxiliary errors were detected with

greater accuracy than determiner errors; this difference interacted with the selective vulnerability of agreement errors (i.e., while both auxiliary agreement and determiner agreement errors were selectively vulnerable, determiner agreement errors showed a greater vulnerability than auxiliary agreement). This part of speech effect was also reflected in performance on the digit load task, where subjects were *better* at holding the digits in memory when the stimuli were *auxiliaries* than when they were determiners. The same superiority of auxiliaries over determiners was found in the visual domain for grammaticality judgment (Blackwell et al., 1993). Wulfeck and Bates (1991) and Wulfeck et al., (1991) have also reported superiority of auxiliary over determiner errors in auditory grammaticality judgment in normals. Those authors suggest this difference may be because sentence-level errors (i.e., auxiliary verb agreement) may be more structurally important than phrase-level errors (i.e., agreement of the determiner with its noun).

### **The Competition Model: a performance account of selective vulnerability in normals and aphasics**

This research adds to the growing body of evidence for the selective vulnerability of morphology, both in patient populations and in normals under conditions of diminished resources (see also Kilborn, 1991). Restricting ourselves to the current experiment for the moment, what makes agreement errors—and, to a lesser extent, omission errors—comparatively vulnerable?

One possible explanation arises from the framework of the Competition Model (Bates & MacWhinney, 1989) designed to explain

language acquisition and performance under real-time conditions and in a manner compatible with connectionist models of language processing (e.g., 1991; Elman, 1990). Constructs in this model include cue validity—the usefulness or informational value of a cue—and cue cost, the amount of resource (cognitive or otherwise) needed to use the cue (e.g., the more salient a cue the less costly).

**(1) The role of cue validity.** The Competition Model predicts that in normal acquisition the processor comes to tune itself more to those cues which have higher validity. In a language such as English, word order is a nearly deterministic cue to agenthood (i.e., in a noun-verb-noun sentence subjects almost invariably choose the first noun as agent). In Italian, word order is more free to vary, while verbs carry a greater load of morphological marking, and so agent choice is much more driven by verbal-agreement marking (Bates & MacWhinney, 1989).

The Competition Model predicts that in situations of resource diminution—such as in our experiment, and as in, we claim, receptive agrammatism—the vulnerability of a cue is proportional to its validity or information content. In a strict word order language such as English, sensitivity will show the pattern transposition > agreement, as we found here. In a language such as Italian, where agreement morphology is much more important than word order in determining agenthood (see Figure 1), we would expect the opposite pattern of agreement > transposition. As we have already discussed in the introduction, while this profile is consistent with findings in Italian, it is overlaid with the crosslinguistic finding that morphology is selectively vulnerable compared to

word order. A purely information-based explanation of our results, then, may not be in order. This brings us to the other major construct of the Competition Model mentioned above, cue cost.

**(2) Cue cost and the structural qualities of the three error types:** Perhaps agreement errors are hardest to detect because they represent the only errors of commission in our data set—there are no elements actually missing or out of place. Omission errors may be easier to detect than agreement errors because the structure is actually missing an element. Transposition errors may be the easiest to detect as they are in essence an omission error plus a moved element, providing two cues to ungrammaticality. The omission error represents only one violation of the relative ordering of elements, while the transposition error represents two such violations. Elman (personal communication) reports a similar profile of sensitivity for transpositions and omissions (transposition > omission) in simple recurrent nets (SRNs) trained to anticipate temporally ordered stimuli.<sup>9</sup> When such networks have learned a simple grammar, they are more able to recover (i.e., continue successfully with the prediction task) when the error is an omission rather than a transposition error. As Elman points out, this particular profile indicates a sensitivity to relative rather than absolute order. As 3.1 shows, omission errors have three elements in the wrong absolute position,

9. Although he cautions that these findings may not be intrinsic to SRNs but may be dependent upon the particular tasks upon which he has trained them.

while as 3.2 shows, transposition errors have only two elements in the wrong position.

If these networks (and, by extension, our subjects) were sensitive to absolute order, one would expect the opposite profile of sensitivity: omission > transposition. Thus, to the extent that we continue to find parallel performance profiles between humans and networks on this type of diachronic detection task, we may find support for the idea that the manner in which these networks come to organize information reflects similar processes in human acquisition.

Finally, if we are to invoke cue cost as responsible for these differences in detectability, we must ascertain what cue cost properties the system is sensitive to; in other words, a psychophysics of cues. In Kilborn (1991) the stimuli were subjected to a noise mask, so that what was being manipulated was the physical salience of the cue. However, there is no reason to believe that there is only one dimension of cue cost, and the effects that our cognitive stress manipulation had on cue cost were almost certainly linearly independent to this noise effect. This is another reason why this line of research in tandem with a compatible line of neural network research would be helpful.

		<b>position</b>			
		1	2	3	4
<b>3.1. omission error</b>	A	C	D	...	
<b>3.2. transposition error</b>	A	C	B	D	

**(3) What are omission errors?** As we pointed out in the introduction, transposition errors are very rare in aphasic patients. Errors of substitution and omission are both quite common, although they are distributed differently across fluent and non-fluent aphasics (i.e. substitution errors are more common than omissions in fluent, paragrammatic patients; omission errors are more common than substitutions in non-fluent, agrammatic patients (Bates, Friederici & Wulfeck, 1987b; Goodglass, 1993). Our data for normals under stress clearly suggest that substitutions constitute the most vulnerable form. That is, substitution errors are difficult to detect and hence more likely to slip by some putative output-monitoring device. Nevertheless, at the highest levels of stress (six digits), omission errors also start to elude the monitoring device. In the literature on speech errors in normals and aphasic patients, there is some controversy about the status of omission errors. Indeed, some investigators have suggested that omissions constitute a form of substitution. Let us consider that literature briefly here, and see what it offers for our interpretation of these results.

Some researchers have characterized omission errors as a sub-type of agreement errors, while others claim omissions to be a separate error type in their own right (Grodzinsky, 1982; Grodzinsky, Swinney & Zurif, 1985; Heeschen, 1985; Menn & Obler, 1990; Stemberger, 1982; Stemberger, 1985; Stemberger & MacWhinney, 1986). To some extent, this controversy is tied to a language-specific confound: In a language in which zero-morphs are common (such as English), if a subject produces an incorrect form such as “move” for the target “moves,” are they omitting the “-s” in-

flexion or simply substituting the more frequent form?

Some insight into this question comes from the study of agrammatism in Hebrew, a language where verbs cannot be phonologically realized without some form of inflection (inflections are formed by inserting vowels into a consonantal-trigram base morpheme). For a verb to be produced, omission errors are not possible. What *is* possible is for speakers to insert vowels that do not correspond to legal inflections in any context. This would be equivalent to an English-speaking aphasic producing some neologistic inflection on the verb such as “move-et.” Hebrew agrammatics do not make these types of production errors; when they produce a wrongly inflected verb form it is a legal form in the wrong context (Grodzinsky, 1982; Grodzinsky et al., 1985). This finding would tend to support the “omission as substitution” hypothesis. Subjects have intact access to the set of proper inflections, but have difficulty in selecting the context-appropriate one.

Of course, so far we have been discussing omissions in production; failure to detect omission errors in receptive processing may have at least partially non-overlapping origins. Even if productive profiles reflect receptive sensitivity (through some sort of self-monitoring process), not all omission errors in production necessarily reflect a lack of sensitivity to that error type in receptive processing. As has been pointed out (e.g., Heeschen, 1985), some omission errors may reflect an intentional move on the speaker's part to a short, elliptical and economical mode of discourse. Given the special status of omission errors (i.e., there are many possible causes for an element's not being produced) and that subjects do display

very real difficulties in lexical selection in a grammatical context, these subjects may in some cases be sensitive to their errors and simply displaying a more conscious strategy; e.g., “I know that I can't get the right form; I'll just omit it and push on.” Thus, especially with this error type, it seems unrealistic to characterize all errors as having the same origin, or to characterize productive and receptive vulnerability as necessarily always coming from the same underlying deficit.

The current experiment may offer some insight into this issue, as we have demonstrated that in normal receptive processing omission errors are different from agreement errors. In the baseline task, omission errors were detected with the same high accuracy as transposition errors, while in the dual task condition they fell between transposition and agreement errors in detectability. Of course, all of our errors, including omission errors, involved *free-standing* morphemes for which an argument of “omissions *qua* omissions” is more reasonable. Thus, it may still be more parsimonious to classify the omission of *bound* morphology as a sub-type of agreement error in which a zero-marked form appropriate in other contexts is substituted (in those cases where zero-morphs exist, of course). Yet even here we should point out that many of our free-standing morphemes have, in a sense, legal zero-forms. For example, in many contexts a determiner before the noun is optional (e.g., “Boys are playing” vs. \* “Boy is playing”), which is not the case in all languages. If the current stimuli (which use free-standing morphemes whose omission might be interpreted, in some cases, as a form of substitution error) show different profiles for omission and agreement errors, it is possible that for bound morphemes

with legal zero-forms we may still need to distinguish omission and agreement errors. Thus, we conclude that in receptive processing there is evidence for classifying omission errors as different from agreement errors, at least for free-standing morphemes, while the causes of omission errors in production may be somewhat more diverse.

**(4) What is the relation between expressive and receptive deficits?** First, let us emphasize that we are not arguing that aphasia is nothing more than an overall reduction in working-memory capacity. Certainly the well-documented differences in performance profiles between Broca's and Wernicke's alone are sufficient to counter that claim.<sup>10</sup> Second, dissociations between receptive and expressive agrammatism have been demonstrated, as we have already discussed, so we cannot offer an account of vulnerability that is completely modality free. Yet there is also no reason to assume that the modalities are totally independent; there are many reasons to postulate the existence of a monitoring device that imposes a receptive look at the subject's own output before or after it happens. For example, research on speech errors in normals suggests that speech production is accompanied by a covert process of error monitoring (Levelt, 1989; Dell & Reich, 1981; Dell, 1985; Stemberger, 1982; 1985; 1989). Some such process is needed to explain the rapid self-corrections that often occur after er-

---

10. Although we do note that there are many more similarities between these two populations than has generally been realized, e.g., Heeschen, (1985); and, as the results of Miyake et al., (in press) show, global resource limitations can have selectively different effects on different normal subjects due presumably to individual differences.

rors are produced by normals, and may also be involved in emending errors before they are produced. If this is the same process being measured by our cognitive stress task, then possibly at least some errors in agrammatic production are due to the failure of this mechanism. As we have already discussed, however, the paths to productive errors—especially omission errors—are undoubtedly varied, and will present a continuing challenge to any attempt at a unifying theory.

Obviously we are convinced that performance-oriented theories like the Competition Model offer a better account of the facts. Such models can provide a principled account of probabilistic data, i.e. all those “in-between”, 80% results that are difficult to explain in models that explain deficits in terms of presence or absence of structural types. At the same time, such models can explain results for aphasic patients and normal controls, by assuming that these groups occupy different points along one or more continua of performance.

### **Linguistic competence models: an alternative view**

We began our introduction with an historical overview, recounting how the doctrine of central agrammatism first arose within the field of aphasiology, and why that doctrine fell on hard times across the 1980's. One reason, of course, is that agrammatic symptoms can be induced in normal adults by placing those adults under some form of stress (i.e. results like those that we have presented here). Lest we give the impression that linguistic approaches to agrammatism have disappeared from view (or continue to exist in a discredited form), it behooves us to consider some recent

variants of linguistic aphasiology that are more compatible with our data.

The most recent variant of the closed-class theory of agrammatism can be found in the papers contained within Grodzinsky (Ed., 1993), a special issue of the journal *Brain & Language* devoted to aphasia research within the framework of modern generative grammar. Although there are some detailed theoretical and empirical differences among the papers in this volume, all have eschewed the notion of a “closed-class processor”. That is, what appears at first glance to be a morphological deficit (i.e. difficulty with bound inflections and free-standing function words) is now attributed to a specific problem with syntax (where “syntax” is defined as the set of rules that govern hierarchical relations and linear ordering among sentence constituents). In particular, it is proposed that aphasic patients suffer from a deficit in the coindexation of traces, a proposal that is referred to as the “trace-deletion hypothesis” (TDH), first proposed by Grodzinsky (1986) and later revised by Hickok (1992) as the “revised trace-deletion hypothesis” (RT-DH).

To understand this proposal, we have to consider how morphological values are fixed within the grammar. According to the theory of Government and Binding, or “GB” (Chomsky, 1981), certain elements within the sentence undergo an abstract form of movement, from their canonical position in an underlying tree structure (now called “D-structure”, a descendant of what was once called “deep structure”) to their position in the ordered string that is ultimately realized in sound (now called “S-structure”, a descendant of what was once called “surface structure”). The critical point is that mapping from D-structure to S-structure

ture requires a form of coindexation, i.e. a notation by which the speaker/listener can track the identity of elements at each level of analysis. In current versions of the theory, a moved element leaves a uniquely-marked “trace” that indicates where it used to be within the string (a kind of “Kilroy was here”). Syntactic theory contains a number of constraints that indicate who is allowed to move where, so the ability to interpret a sentence and the ability to recognize a sentence as grammatical or ungrammatical both depend (at least in part) on the ability to match moved elements with the positions that they once occupied. It is argued that agrammatic aphasics still possess the ability to recognize whether a string of elements is properly ordered (i.e. whether certain types of words are in the right position relative to their neighbors, e.g. the difference between “the dog” and “dog the”), and they can comprehend sentences in which recovery of meaning depends on these ordered relations. However, these patients have lost the ability to coindex traces, i.e. to match moved elements against their original positions along the road from D-structure to S-structure.

To see how this hypothesis works, consider the contrast between object- and subject relative sentences. In an object relative sentence such as 4.1, GB posits a “trace” or abstract, empty marker (indicated here by  $(t_i)$ ) after the verb “chasing”, in the canonical position for the direct object. The lexicalization of the direct object, “lion”, is “coindexed” with the trace, as indicated by the subscripts. As a simplified explanation, according to GB a listener must successfully apprehend this chain in order to correctly capture the structure and meaning of the utterance; particularly, that it is the lion being chased.

**4.1. The lion<sub>i</sub> whom the tiger<sub>j</sub> is ( $t_lj$ ) chasing ( $t2j$ ) is running fast.**

**4.2. <[lion],  $t2j$ >, <[tiger],  $t1j$ >**

**4.3a.<[lion]<sub>i</sub>,  $t2j$ > <[tiger]<sub>j</sub>,  $t1j$ >**

---

The essential claim of the TDH and the RTDH is that in agrammatism, it is the processing of this chain that is disrupted. Although variants of the theory differ on details such as what the exact nature of the resulting representation is (e.g., Mauner, Fromkin & Cornell, 1993; Shapiro et al., 1993; Zurif et al., 1993), the basic notion is that except for trace deletion, syntactic processing is intact. The resulting somewhat underspecified structure provides ambiguous information to the role-mapping module of the processor, explaining why subjects are above chance on grammaticality judgments for sentences which they have difficulty completely comprehending. For example, Mauner et al. (1993) propose that agrammatics correctly form a chain—or, in their version, an “R-dependency” (see 4.2)—without being able to properly coindex the elements (see 4.3). In the normal processor, the “coindexation condition” holds, that “If  $\alpha$  is R-dependent on  $\beta$ , then they must bear the same R-index” (op. cit., p. 357), while in the agrammatic case subjects cannot make use of this condition. While a normal would always process 4.1 such that the R-dependencies coincided, as in 4.3a, an agrammatic would only process 4.1 as far as 4.2, leaving either 4.3a or 4.3b as a possible interpretation. Mauner, et al. use this to explain (e.g.) Linebarger et al.’s (1983) and Linebarger’s (1989) results. Agrammatics successfully

judge the grammaticality of, e.g., passive sentences such as 5.1 and 5.2 because computing the R-indices is not necessary for correct grammaticality judgment. However, agrammatics have difficulty with violations of, e.g., Wh-head agreement (as tested by sentences 5.3 and 5.4) because this *does* involve R-indices.

Notice that the claim by Maurer et al. is restricted to a comparison between Broca's aphasics and normal controls. They bring no data to bear suggesting that Broca's aphasics are uniquely impaired in their ability to recognize agreement errors like the one in 5.4. And indeed, studies comparing sensitivity to agreement in Broca's and Wernicke's aphasics have shown that both groups find structures like 5.4 difficult to resolve (for a review, see Bates, MacWhinney and Wulfeck, 1991). The results presented in the current study show that selective deficits in the processes required to recognize agreement errors can be induced in normal adults subjected to a global form of stress. In other words, deficits in agreement are not uniquely associated with Broca's aphasia or (for that matter) with any other form of focal brain injury.

On the other hand, Zurif et al., 1993 have shown that Broca's aphasics have difficulty with a different aspect of coindexation called "gap filling", which refers to the use of traces to recover the referents of pronouns and other

anaphoric expressions. Consider the following sample sentence from Zurif et al.:

**6.1. The gymnast loved the professor<sub>i</sub> from the northwestern city<sup>1</sup> who<sub>i</sub> <sup>2</sup> (t<sub>i</sub>) complained about the bad coffee.**

Normal and aphasic subjects are given sentences like these in the auditory modality. At the same time, they are told that words and non-words will appear periodically on the screen. Their task is to make a word/non-word decision about the target word as soon as it appears. Targets include words that are related to the meaning of the antecedent word (e.g. "teacher") and control words that are unrelated to the antecedent (e.g. "address"). These targets appear in one of two probe positions: before the referring expression "who" (i.e. at 1) or after the referring expression "who" (i.e. at 2). Several studies have shown that normal listeners are faster at making a lexical decision at point 2 if the probe word is semantically related to the antecedent, but they show no such priming effects at point 1, before the referring expression was presented. This suggests that pronouns and other anaphors prime their antecedents, an effect which proponents of GB theory attribute to the operation of coindexation between the trace (t<sub>i</sub>) and its referent (professor<sub>i</sub>). The critical point for present purposes is this: the Broca's aphasics tested by Zurif et al. failed to show priming at the gap, whereas the Wernicke's aphasics tested in this study showed the normal pattern of gap-filling. This single dissociation contrasts markedly with results obtained in a different study (Shapiro et al., 1993) showing that Broca's aphasics automatically activate all the arguments associated with a verb (i.e. they "fill out

---

**5.1. The boy was followed by the girl.**

**5.2. \*The boy was followed the girl.**

**5.3. The pencil which he bought was nice.**

**5.4. \*The pencil who he bought was nice.**

the thematic grid”), while Wernicke’s aphasics do not. Based on findings like these, Zurif et al. offer the following speculations concerning the localization of syntactic functions:

The conclusions we have drawn clearly do not provide a characterization of the role actually played by the cortical tissue implicated in Wernicke’s aphasia. Still, we do provide a lower boundary on its functional commitment: whatever its role, it is not crucially involved in the real-time structural analysis required for the recognition and filling of gaps left by constituent movement. By contrast, *left anterior cortex*—the cortical region usually implicated in Broca’s aphasia—*does appear to be necessary for the operation of gap filling*.<sup>11</sup> [italics added]

They go on to suggest that such findings argue against diminished-capacity hypotheses like the one that we have presented in the present study:

These capacity limitation theories, although different in many important respects, commonly imply that a single “pool” of resources is shared by various sentence processing devices [footnote omitted]. We suggest that there is not a single resource, but multiple resources, each dedicated to a particular process. Of course, perhaps the process of activating thematic and other lexical properties does not fall within the domain of the capacity limitation hypothesis. But if the hypothesis allowed such a restriction on its sphere of influence, other, similar, restrictions could also apply, and the account would not be falsifiable.<sup>12</sup>

In response, we are certainly not saying that *all* language deficits in the aphasias are due to global resource diminution, as clearly the data do show different sub-types that cannot all be linked to a unitary resource. Nor do we believe that our more limited claim, simply because it allows “such a restriction on its sphere of influence,” is therefore unfalsifiable. While the

data show that there are certainly some impairments that appear to be unique to either Broca’s or Wernicke’s, there are also others that have been traditionally ascribed to these syndromes alone, and often to the specific tissue sites involved, which in fact may be explicable by more general mechanisms. Therefore, one of the goals of current research should be to distinguish those aspects of agrammatism which really *are* unique to agrammatism with those aspects that might be found in a variety of disorders or even, as we have seen, in normals. This, in turn, has obvious implications for the organization of language in the normal brain, but the nature of this organization cannot be determined until we know what is at the core of these various disorders and what is more peripheral to their description. We speculate that it is *not* that the areas of brain implicated *directly* and *explicitly* contain the syntactic representations or operations implicated. Rather, we propose that—for those deficits which can be correlated to damage to a specific brain site—the site helps to support types of processing that are crucial for those representations or operations.<sup>13</sup> Further, it would seem that much of the field appears to be moving towards this position as well, even those who have traditionally held very localist interpretations of aphasic data:

Our data indicate that these grammatical limitations are rooted to fairly elementary processing disruptions—specifically, to disruptions of automatic lexical reactivation (access) at the gap. In this view, the brain region implicated is *not* the locus of syntactic representations per se. Rather, we suggest that this region

11. Zurif et al., 1993, p. 461.

12. Shapiro et al., 1993, p. 444.

13. see, e.g., Farah & McClelland, 1991 for an example of a highly interconnected model that nonetheless demonstrates what appear to be quite specific dissociations.

provides processing resources that sustain one or more of the fixed operating characteristics of the lexical processing system—characteristics that are, in turn, necessary for building syntactic representations in real time.<sup>14</sup>

### In conclusion

What makes these particular error types—agreement errors and, to a lesser extent, omission errors—vulnerable? Does it have to do with the form of the information they carry (cue cost) or the amount or type of information that they carry—both are possibilities we have suggested—or are other factors at work, such as frequency of appearance during learning, neighborhood density, or memory constraints? There are a variety of angles of attack with which to pursue these questions, some of which we are currently engaged in within our laboratory. For example, one might carry out similar investigations of stressed grammaticality judgment across different languages, where these factors are different (e.g., languages in which agreement carries less information than in English, or where agreement cues are more salient). Another approach is to train subjects in small artificial-language experiments where these factors can be manipulated directly (e.g., salience of a particular cue, information carried by that cue).

---

14. Zurif et al., 1993, p. 461.

Each approach has both strengths and weaknesses. Using the artificial-language approach, each parameter may be manipulated by the experimenter with precision, an option not available when using natural language, where the experimenter is essentially bound to whatever language types are available in the world. In addition, the artificial-language experiment is closer to what a neural network simulation is presented with, if one is interested in exploring psycholinguistics through that venue. But of course, such artificial-languages are toy languages only, and it is not clear how much of what can be learned through such precise, parametric manipulations with pseudo-linguistic stimuli will tell us about *real* natural-language processing—or indeed how much of the subject's native language will influence the outcome of the experiment. We believe, however, that both approaches taken together should provide sizable insight into the issue.

Thus, although the underlying etiology of this selective vulnerability remains to be fully mapped, we have nonetheless demonstrated a definite problem with the transparency assumption: Selective dissociations needn't always reflect selective disruptions of mental/neural architecture. *Apparent* damage to a "syntax module" may be due to the particular vulnerability of those aspects of syntax implicated.

## Method

### Subjects

Subjects were University of California, San Diego undergraduates participating for either course credit or money. All subjects were assessed as right-handed non-multilingual native speakers of English. Of the 112 subjects used (those that made it past the “two standard deviations” criteria discussed below), 59 were female and 53 were male.

### Grammaticality Judgment Stimuli

Stimuli for the grammaticality judgment task include a total of 168 sentences: 84 ungrammatical target sentences, 40 grammatical control sentences matched for length and grammatical structure, and 44 distractors (see below). The design of the experiment is focused on the ungrammatical targets, which vary in the part of speech involved in the error (auxiliary vs. determiner), the position of the error within the sentence (early vs. late), and the kind of violation created from a common pool of grammatical types (i.e. errors of omission, agreement and transposition). The ungrammatical target sentences fall within a  $2 \times 2 \times 3$  design (with error type, error location, and part of speech as within-subjects variables; see Appendix I). Each of the twelve cells within this design contains seven ungrammatical sentences. For each of these ungrammatical sentences, subjects also hear a grammatical control sentence matched for length and grammatical structure. To keep the length of the experiment within reasonable bounds, some of these grammatical sentences are used as controls for more than one particular ungrammatical sentence. There are also 44 distractor sentences (22 grammatical and 22

ungrammatical) that reflect a wider array of lengths (from 3 to 17 words) and structural types. The distractors were included to prevent subjects from detecting the error and length regularities in the target sentences. A complete list of stimuli is provided in Appendix I.

To develop the 84 ungrammatical targets and 40 grammatical controls, we began with a pool of grammatical sentences, varying in length from 8 to 12 words. Half of these sentences contained at least one auxiliary verb that would serve as the target of all auxiliary violations. On half of these items, the auxiliary was located early in the sentence (e.g., “They were reading several large maps while waiting for the next train.”), while on the other half, the auxiliary was located near the end of the sentence (e.g., “In a big, old, red boat, two girls were rowing slowly.”). The remaining target items contained at least one determiner (including numerals and demonstrative adjectives) that would serve as the target of all determiner violations. On half of these items, the target determiner was located early in the sentence (e.g., “The girl was eating some dark chocolate ice cream.”), while on the other half, the target determiner was located near the end of the sentence (e.g., “My new blue and green silk ball gown was costing a fortune.”). Although there were no passives or embedded sentences in this stimulus set, the sentences represent a range of different structural types, varying in factors such as the presence and location of prepositional phrases, presence or absence of relative clauses or subordinate clauses, and the number of adjectives modifying the subject and object. A list of these structure types is provided in Appendix IV. Approximately twenty different sentence tokens were constructed for each of these struc-

tural types, and randomly assigned to the appropriate grammatical or ungrammatical conditions.

Early errors occurred within the first 1200 msec of the sentence, while late errors occurred after the first 1500 msec of the sentence. All errors in the experiment were “local,” i.e. the licensing word and the error were always adjacent. Because omission, agreement and transposition errors were all created from the same basic sentence types, it can be argued that these stimuli represent a set of minimal contrasts. Nevertheless, even within a well-controlled stimulus set of this kind, there are a number of complicating factors that bear on our interpretation and thus require some explanation before we proceed.

**Stimulus design: Ungrammatical targets.** Stimuli were 168 sentences: 84 ungrammatical target sentences, 40 grammatical control sentences matched for length and grammatical structure, and 44 distractors (see below). Experimental design focused on the ungrammatical targets, which varied in:

- part of speech of the error (auxiliary vs. determiner)
- the position of the error (early or late in the sentence)
- type of violation (i.e., errors of omission, agreement and transposition)

Thus, the ungrammatical target sentences formed a  $2 \times 2 \times 3$  design, with part of speech, location, and error type as within-subject variables.

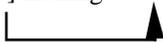
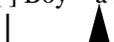
**Grammatical controls.** Each of the twelve cells in the design had seven ungrammatical

sentences. For each of these ungrammatical sentences, there was a grammatical control sentence matched for length and grammatical structure. To keep the experiment reasonably short, some grammatical sentences were used as controls for more than one particular ungrammatical sentence. There were also 44 distractor sentences (22 grammatical and 22 ungrammatical) from 3 to 17 words long, and of various structures. Distractors were to prevent subjects from detecting regularities in the target sentences (see Appendix I for all stimuli and how they were generated).

**Creation of ungrammatical targets.** The 84 ungrammatical targets and 40 grammatical controls come from a pool of grammatical sentences from 8 to 12 words long. This pool of sentences represents a range of seven structural types, varying in presence and location of prepositional phrases, presence or absence of relative clauses or subordinate clauses, and the number of adjectives modifying the subject and object (see Appendix IV). Approximately twenty different sentence tokens were constructed for each of these seven structural types, and randomly assigned to the appropriate ungrammatical target cell or grammatical control condition. Half of the sentences in this pool had at least one auxiliary verb to be the target of an auxiliary violation, while the other half of sentences had at least one determiner (including numerals and demonstrative adjectives) to be the target of a determiner violation:

**Auxiliary verb sentences:** On half of these items, the auxiliary was located early in the sentence (e.g., “They were reading several large maps while waiting for the next train.”), while on the other half, the auxiliary was located near the end of the sentence (e.g., “In a

**Table 4: Creation of Grammaticality Judgment Stimuli**

auxiliary verbs	omission agreement transposition	Joan [was] making * several big and tasty ice cream drinks. Joan <u>were</u> * making several big and tasty ice cream drinks. Joan [ ] making * was several big and tasty ice cream drinks. 
determiners	omission agreement transposition	[A] Boy * is driving a large van that the artist has painted. Those <u>boy</u> * is driving a large van that the artist has painted. [ ] Boy * a is driving a large van that the artist has painted. 

big, old, red boat, two girls were rowing slowly.”).

*Determiner sentences:* On half of these items, the target determiner was located early in the sentence (e.g., “The girl was eating some dark chocolate ice cream.”), while on the other half, the target determiner was located near the end of the sentence (e.g., “My new blue and green silk ball gown was costing a fortune.”).

**Location of error.** Early errors occurred within the first 1200 msec of the sentence, while late errors occurred after this point. The licensing word and the error were always adjacent (i.e., all local errors). Because omission, agreement and transposition errors were created from the same basic sentence types, it can be argued that these stimuli represent a set of minimal contrasts. Nevertheless, even within a well-controlled stimulus set, there are complicating factors affecting our interpretation. We now discuss this.

**Stimulus design considerations:** At what point does a sentence first become ungrammatical? For a particular error type (e.g., transposition errors on auxiliary verbs early in

the sentence) the ungrammaticality does not necessarily begin at the same structural point, yet it is this structural point that the sentences have in common. One remedy: Give subjects several variants on the same sentence; however, these results would have questionable generalizability.<sup>15</sup> As a nod to this dilemma, we have locked sentences of a particular error type to a “Common Structural Point” (CSP; see Table 4).

**Rules for creating the three error types**

**Omission errors:** remove the relevant word (auxiliary or determiner) from the sentence (see Table 4). The asterisk refers to the aforementioned CSP.<sup>16</sup>

**Agreement errors:** replace the target word with an item that doesn’t agree in number. **NB:** violations of determiner agreement within a subject noun phrase provide *two* cues to the agreement violation. Cue one is the mismatch in number between determiner and noun (i.e.,

15. I.e., a conclusion about “transposition errors,” as a group, is methodologically questionable if based only upon transposition errors which are immediately identifiable at the point where the element is missing (e.g., “My mother visiting \* was her aunt”) and not at the point where the element is moved (e.g., “The boy visiting was \* his aunt”).

“a girls \*”); cue two is the mismatch between the auxiliary verb and the determiner.<sup>17</sup> This situation can be symbolized as, “A girls \* were \* working quietly near the small, red house.”

**Transposition errors:** move the relevant word one word downstream from where it belonged. We have placed the asterisk after the word opening, the matching structural point for omissions, and the first point at which the subject might notice that a potential element is missing (although see the note above about this). This suspicion will, of course, be confirmed when the subject encounters the displaced element. Hence transposition errors constitute another instance in which there are really two cues to the existence of an error, one at the first point at which a subject might notice that there is a hole (similar to omission errors) and another at the point further downstream where the displaced element occurs.

There is one further difference between late omission errors and the other two late-error

---

16. This is a good place to note in passing that an additional complication comes from the contrast between early and late omission errors. Because all of our sentences are marked with normal English punctuation, omission errors often involve a double cue. For example, given a late auxiliary-omission error such as, “While sitting on the red sofa, her older friend eating\* some cake,” the subject actually has two cues to help him decide whether an error has occurred. First, after reading the word “eating,” the subject knows that a verb that should have been preceded by an auxiliary was not. Second, because the word “eating” is soon followed by a period (visible at the end of every sentence stimulus), the subject may conclude that no further items will come along to salvage the sentence (e.g. the sentence will not turn into something such as, “While sitting on the red sofa, her older friend eating some cake, was watching TV.”). Hence we might argue that the above examples each provide the subject with two distinct error cues, illustrated as follows: “While sitting on the red sofa, her older friend eating \* some cake. 9)

17. The auxiliary verb can only agree with one of the two elements within the subject noun phrase, either locally with the preceding noun, or globally with the determiner.

types: on transposition errors, the sentence will necessarily last approximately one word longer after the zero point than it does with errors of omission or agreement. Because the three error types share the same CSP (i.e. they start to go wrong at exactly the same point in the sentence), this need not constitute a problem. However, if subjects cannot make up their minds at the CSP and want to wait for more information before they decide, then we are faced with an artifact: subjects are forced to make up their minds at the zero point on many late agreement and omission errors, because the sentence is already over; by contrast, they are able to delay their decisions for a while on the transposition errors. Hence any differences that we may observe in the size of the decision region for late errors may be a by-product of unavoidable structural differences among the three late violation types. For this reason, all analyses of timing and decision points will be conducted separately for early vs. late errors.

We first presented the caveat about using discrete error points with such stimuli in Blackwell et al., 1993 using the sentence-level gating paradigm, in which subjects read a sentence word-by-word and indicate at each word if they believe the sentence to be grammatical. In that paper, we demonstrated that certain error types, most notably omission errors, may provide a wide window of variability about where the “error point” lies, both over subjects and over items. Figure 9 demonstrates this. It shows the decision function (i.e., percentage of subjects who judge the sentence to be grammatically correct) for the three types of early auxiliary errors as a function of distance to the CSP, or the “0” point on the graph. The graph shows that for agreement errors, subjects make a decision directly at the CSP, for transposition

errors their decision tends to be between the CSP and the moved element, and for omission errors the decision region is a quite protracted area following the CSP. Thus, we found that some items, while ungrammatical, may be close enough to grammatically correct items to give subjects pause (as noted above). We also found that some subjects may use quite strict decision criteria, pressing the button immediately upon detecting a potential error, while others may accept strings which could conceivably pan out as legal continuations. The lesson is that for any interactive activation model of sentence processing, reaction time is generally taken to be indicative of the strength of activation of the item being responded to. We must also consider that reaction time mea-

asures are also a reflection of where we decide to start the clock—a decision that is not always clear-cut.

### Digits stimuli

In conditions where subjects also had to carry a digit load, they saw either two, four, or six different digits before the sentence, which we call the “target” digits. After the sentence was played, the same number of digits again appeared on the screen. Half of the time, the “test” digits were identical to the target digits (and in the same order); half of the time one and only one digit was substituted with another digit *not* in the target set.

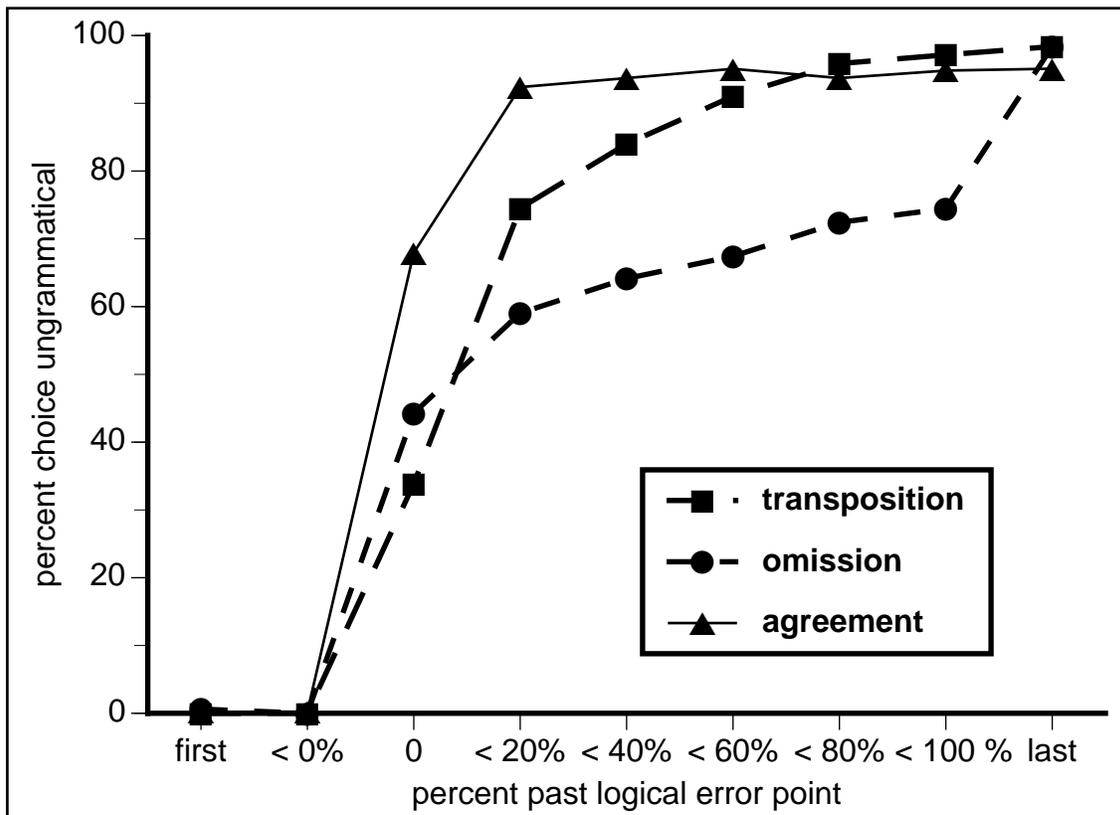


Figure 9: Gating experiment, for early auxiliary errors. Graph shows percentage of subjects who judge the sentence to be ungrammatical at each gate. The zero point corresponds to the “logical error point” (see text for details).

## Equipment

The experiment was run on a Zenith Data Systems 12 MHz PC/AT, with a Carnegie-Mellon button box. Sound files were digitized and presented using the Atlas sound program, developed at Carnegie-Mellon University, with 12-bit sound.

## Procedure

**Without digits:** The actual experiment consisted of 168 trials of the sentence stimuli described above. The sentences were arrayed in a Latin Square design, with the constraint that no two items from the same cell (e.g., “early determiner omission”) could appear consecutively, and no two fillers could appear consecutively, except at the end of the experiment if only fillers were left. Subjects received a break after the first ten trials, and at every forty trials thereafter.

A trial consisted of the following: the sentence was played to the subject, followed by a 3000 msec window for subjects to make a button press. Subjects were instructed to press the proper key “as soon as a decision had been made, even if the sentence was still running.” The button press did not affect the timing of the next sentence’s presentation, which began after the window.

Subjects were given thirty baseline practice trials to acquaint themselves with the button box. In these trials, they heard the either the word “bad” or the word “good,” and were to press the appropriate key. Subjects were also given twenty practice sentences before the actual experiment.

Both choice and reaction time were collected. Reaction time was measured from the off-

set of the word just before the “common structural point” (CSP), defined above (see section “Grammaticality Judgment Stimuli”).

**With digits:** The digits conditions were identical to the without-digits condition with the following modifications: subjects also used the Amdek amber monochrome monitor, directly in front of them. A trial consisted of the following: the word “READY” appeared on the screen for 667 msec, accompanied by a short tone. After the offset of the ready cue there was a 333-msec pause, followed by the presentation of the “target” digits, which appeared one after the other in the center of the screen. Each digit appeared for 350 msec. After the last digit offset, a “#” mask appeared. The sentence began immediately after the 333 msec after the “#” mask onset, and, as in the “without digits” condition, subjects were instructed to press the button in response to the sentence as soon as their decision had been made, even if the sentence was still running. Following the end of the sentence, there was a 3000-msec window to allow for a response if the subject had not already made a response, followed by the “#” offset and a 333-msec delay. The “test” digits then appeared in the same manner as the “target” digits, one every 350 msec, followed again by the “#” mask. Subjects were instructed to make a second button press to these digits as fast as possible without making an error, indicating if the test digits matched the targets or not. They had a 1500-msec window after the offset of the last digit to make a response to the digits. If correct, they were to press the “good” button (corresponding to grammatical in the sentence judgment task); if incorrect, they were to press the “bad” button. If correct, the digits were identical to

the target, including the order. Thus, if the target was “1-2-3-4” the test was “1-2-3-4”. If incorrect, the test digits would have one and only one digit wrong. Thus, if the target were “1-2-3-4” the test might be “1-2-6-4”. Each digit was generated at random, with the constraint that no digit repeat in either the target or the test set. The test digits were correct 50% of the time. Following the 1500-msec window in which to make a response, the “#” mask vanished, and there was a 1500-msec inter-trial interval. If subjects made more than one button press before the digits appeared, only the first response was recorded. If subjects made more than one button press to the digits after the digits appeared, only the first response was recorded.

In addition to the thirty baseline trials described above, subjects received forty practice trials: 20 without digits and 20 with the same number of digits as the actual experiment.

### Scoring

**Without digits:** Both choice and reaction time were recorded for each trial. In the case where subjects made more than one button press, the first was used. For all dependent measures, “no response” trials were dropped.

Sentence accuracy was evaluated using  $A'$  to grammaticals and ungrammaticals combined.  $A'$  is a non-parametric statistic used to correct for response bias (Grier, 1971; Pollack & Norman, 1964).<sup>18</sup> As such, it is similar to  $d'$ . For  $A'$ , *all* subject responses for target stimuli were used (i.e., ungrammaticals and their grammatical controls). For each of the twelve cell conditions and for each subject, a signal-detection matrix was generated, and “hits” (correct judgments of ungrammaticality) and

“false alarms” (incorrect decisions that a sentence is ungrammatical) were calculated. The  $A'$  analysis was conducted over subjects only, since the logic of  $A'$  is difficult to apply in an analysis over items.

Reaction times were measured from the offset of the last word of the CSP (see section above, “Grammaticality Judgment Stimuli”). Reaction times were to correct responses only. All reaction times, including negative ones, were included. Reaction times where there is a second response were kept if the first response is correct. As noted above, for omissions and transpositions the CSP is the “hole”—the place where the moved element should be. For agreement errors, the CSP is just after the word that violates the agreement rule. Note our warnings above about the utility of reaction times in these sorts of experiments.  $A'$  was calculated by developing a signal detection matrix for each subject and for each of the 12 cells.  $A'$  was then averaged over cells and subjects as needed. For all dependent measures, “no responses” were dropped.

---

18. Raw percent correct scores for grammatical and ungrammatical stimuli do not account for the possibility of subject response bias. For example, a score of 100 for ungrammatical stimuli (all ungrammatical stimuli correctly identified) could mean that the subject is perfect at the task—or simply that the subject has an overwhelming tendency to guess that a sentence is ungrammatical. This cannot be determined without looking at both hits and false alarms. The above subject might also have a false-alarm rate of 100, indicating that in fact they are incapable of differentiating grammatical from ungrammatical sentences and judge everything as ungrammatical. Conversely, a false-alarm rate of 0 (with a hit rate of 100) would constitute perfect performance.  $A'$  is a unified statistic that corresponds to the underlying percent correct in a two-option forced-choice task, *correcting* for bias. It can range from 50 (chance performance) to 100 (perfect discrimination). Of course, normal subjects should not show such strong biases, but  $A'$  still permits discrimination of subject accuracy differences that are more subtle; for example, which subject is more accurate, one with 90% hit rate and a 1% false-alarm rate, or one with a 99% hit rate and a 10% false-alarm rate? (see Pollack and Norman, 1964).

**With digits:** In addition to the sentence choice and reaction time data noted above, we also took both percent correct (not A') and reaction time measured from the onset of the first digit for the test digits.

#### **Post-experimental criteria for including subjects**

**Without digits:** An initial group of subjects was run, from which cut-off values were calculated that were two standard deviations from the mean for A', hit, false alarm, reaction time, number of "no response" trials, and number of "double response" trials. Because this was a Latin Square design, new subjects were run to fill those subject numbers that were dropped.

**With digits:** An initial group of subjects was run, from which cut-off values were calculated that were two standard deviations from the mean for the sentence variables A', hit, false alarm, and reaction time, and for the digits variables percent correct digits and reaction time, as well as for the overall trial variable number of "no response" trials. In the "zero digits" condition, reaction time to both ungrammatical and grammatical sentences were (separate) criteria; due to the difficulty of the dual task and our interest only in the ungrammatical sentences' reaction times, only reaction time to *ungrammatical* sentences was a criterion for the "digits" conditions. In addition, because the computer program could not track "double responses" these were not a factor.

# Appendices

## Appendix I: Experiment Design

Part of speech	Type of error	location of error	
		early	late
auxiliary	omission	Mrs. Brown working * quietly in the church kitchen.	She had written that mystery novel that her mother reading. *
	agreement	The writer were * holding a very big party.	While sitting on the couch, Mr. Lane's daughters was * watching a movie.
	transposition	Miss Hope sending * was several green dresses that Lisa had ordered.	While talking to Jane, Joseph knitting * was a sweater.
determiner	omission	Girl * was working quietly near the small, red house.	The small, thin green vine was sprouting flower. *
	agreement	A boys * are driving a large van that the artist has painted.	Larry is saying that his mother was planting that bushes. *
	transposition	Helicopter * a was hovering loudly over the army base.	Those girls were watching the bright lightning while camping in desert * that.

## Appendix II: Core stimuli, organized by cell

Underlined sentences are grammatical controls; bolded sentences are grammatical repeats.

### 1. Early auxiliary omission

1.1) They examining \* several expensive old paintings while walking through the art museum.

1.2) They were reading several large maps while waiting for the next train.

1.3) My mother visiting \* an expensive and famous plastic surgeon.

1.4) The man was playing both old and modern piano pieces.

1.5) Joan making \* several big and tasty ice cream drinks.

1.6) Julie was eating a large, creamy, chocolate and coconut pie.

1.7) My cousin drawing \* three small pictures of his mother's new cats.

1.8) Her mother was reading some old articles on famous Hollywood movie actors.

1.9) The boy taking \* a black feather that the pigeon had dropped.

1.10) The doctor is reading the medical report that the nurse has written.

1.11) Mrs. Brown working \* quietly in the church kitchen.

1.12) A small boy was walking slowly down the beach.

1.13) Tom's mother forgetting \* that he had taken his new car.

1.14) Several people were saying that fishermen had killed those blue dolphins.

### 2. Late auxiliary omission

2.1) While sitting on the red sofa, her older friend eating \* some cake.

2.2) While baby-sitting for their neighbors, Mrs. Johnson's daughters were eating some candy.

2.3) Her older brother's first guest drinking \* a beer.

2.4) My young cousin's very first dinner party guest was making some drinks.

2.5) The two very famous Italian chefs making \* a salad.

2.6) The two famous New York chefs were making a cake.

2.7) In the very big and shady front yard, Bill's mother picking \* flowers.

2.8) Near the big, old summer house, several animals were drinking water.

2.9) She had written that mystery novel that her mother reading. \*

2.10) They are eating the candy bars that Mrs. Morton has brought.

2.11) The young, new president of John's college speaking \* briefly.

2.12) That very old friend of my father's was walking slowly.

2.13) John's boss has said that the new secretary stealing \* office supplies.

2.14) Sam's friend is saying that his two sisters have made some cookies.

### 3. Early determiner omission

3.1) Boy \* was entering a contest while staying at the hotel.

3.2) The girls were eating some fries while waiting for their friends.

3.3) Girl \* was eating some dark chocolate ice cream.

3.4) The woman was having a very big dinner party.

3.5) Clerk \* was reading several very old and important letters.

3.6) The woman was painting several very large, colorful pictures.

3.7) Woman \* was watching some orange butterflies in the small back garden.

3.8) Her mother was reading some old articles on

famous Hollywood movie actors.

3.9) Woman \* is visiting the old dairy farm that her father has bought.

3.10) The clerk is sending several cotton shirts that Dorothy's mother has ordered.

3.11) Girl \* was working quietly near the small, red house.

3.12) The balloon was floating slowly through the air.

3.13) Woman \* was saying that her husband had bought several big tomatoes.

3.14) The man was reading that many people had protested those new taxes.

#### 4. Late determiner omission

4.1) The boy was finding many big sea shells while playing on beach. \*

4.2) They were reading several large maps while waiting for the next train.

4.3) Her new blue and green silk ball gown was costing fortune. \*

4.4) The large and pale gray cruise ship was hitting an iceberg.

4.5) The small, thin green vine was sprouting flower. \*

4.6) Her two favorite great aunts were making some pie.

4.7) Alice was calling her old college friend at hotel. \*

4.8) Martha was bringing several old dance records to the party.

4.9) The maid whom Sally has hired is cleaning bathroom. \*

4.10) The woman whom Anne's father has hired is cleaning the windows.

4.11) Two very famous art critics were speaking briefly at museum. \*

4.12) A plane was flying slowly over the old landing strip.

4.13) The woman was writing that her two daughters had bought car. \*

4.14) The train conductor was saying that some trash had blocked the tracks.

#### 5. Early auxiliary agreement

5.1) The guests was \* drinking some wine while talking about the movie.

5.2) The girls were eating some fries while waiting for their friends.

5.3) The writer were \* holding a very big party.

5.4) The man was playing both old and modern piano pieces.

5.5) The vine were \* growing a few red and yellow flowers.

5.6) Julie was eating a large, creamy, chocolate and coconut pie.

5.7) The bankers was \* reading those papers on the train.

5.8) Martha was bringing several old dance records to the party.

5.9) She were \* seeing the place where her two older sisters had worked.

5.10) They were visiting the house where Nancy's parents and grandparents had lived.

5.11) Soap bubbles was \* floating slowly into the summer sky.

5.12) Honey bees were flying loudly around a large, old oak tree.

5.13) Mike's parents was \* hoping that he had passed the final exam.

5.14) Several people were saying that fishermen had killed those blue dolphins.

#### 6. Late auxiliary agreement

6.1) While sitting on the couch, Mr. Lane's daughters was \* watching a movie.

6.2) While baby-sitting for their neighbors, Mrs. Johnson's daughters were eating some candy.

6.3) Some famous old Hollywood actor were \* having a party.

6.4) Several very young children were watching a play.

6.5) The old, red brick houses was \* blocking the view.

6.6) Her two favorite great aunts were making some pie.

6.7) In Mrs. Hart's small rose garden, the gardener were \* planting bushes.

6.8) Near the big, old summer house, several animals were drinking water.

6.9) John had finished the candy that his mother were \* saving.

6.10) They are eating the candy bars that Mrs. Morton has brought.

6.11) In the bank's very large lobby, the brokers was \* talking quietly.

6.12) In a big, old, red boat, two girls were rowing slowly.

6.13) Susan had said that she are \* cleaning it.

6.14) Chris is saying that his mother has bought a house.

## 7. Early determiner agreement

7.1) Those girl \* was visiting Jack while driving through the town.

7.2) The boy was reading a comic book while standing on the corner.

7.3) A girls \* were watching the Fourth of July fireworks.

7.4) The woman was having a very big dinner party.

7.5) Those designer \* was selling several expensive imported gowns.

7.6) Those models were wearing that new wave hairstyle.

7.7) A boys \* were feeding the small, brown bird in the yard.

7.8) Those girls were petting the small, brown cat in the yard.

7.9) A boys \* are driving a large van that the artist has painted.

7.10) The clerk is sending several cotton shirts that Dorothy's mother has ordered.

7.11) Those house \* was selling quickly, for very little money.

7.12) The balloon was floating slowly through the air.

7.13) Several sailor \* was saying that the man had predicted a storm.

7.14) The man was reading that many people had protested those new taxes.

## 8. Late determiner agreement

8.1) Jim's sisters were watching the ocean waves while sitting on that rocks. \*

8.2) Mrs. Taylor was eating a turkey sandwich while talking on the phone.

8.3) The very famous rock singer was performing several song. \*

8.4) My young cousin's very first dinner party guest was making some drinks.

8.5) Mr. Hall's entire class was watching several cartoon. \*

8.6) The two famous New York chefs were making a cake.

8.7) Arthur's daughters were driving that red sports car over those mountain. \*

8.8) Those girls were petting the small, brown cat in the yard.

8.9) Several workers whom Mr. Stevens has hired are painting those fountain. \*

8.10) The woman whom Anne's father has hired is cleaning the windows.

8.11) The young man was speaking loudly with two waiter. \*

8.12) A small boy was walking slowly down the beach.

8.13) Larry is saying that his mother has planted that bushes. \*

8.14) Chris is saying that his mother has bought a house.

## 9. Early auxiliary transposition

9.1) Jane's friends watching \* were some fireworks while standing on the hill.

- 9.2) Mrs. Taylor was eating a turkey sandwich while talking on the phone.
- 9.3) Those girls seeing \* were some old and famous silent movies.
- 9.4) The artists were selling several small but expensive watercolor paintings.
- 9.5) She signing \* was her newest and biggest story collection.
- 9.6) The woman was painting several very large, colorful pictures.
- 9.7) Students writing \* are several math problems on the blackboard.
- 9.8) Jane's mother is renting a small apartment in New York.
- 9.9) Miss Hope sending \* was several green dresses that Lisa had ordered.
- 9.10) Jan's hairdresser was learning a new look that Jan had wanted.
- 9.11) The boy walking \* was quickly to the store.
- 9.12) The balloon was floating slowly through the air.
- 9.13) That woman saying \* is that her two friends have stolen several things.
- 9.14) Sam's friend is saying that his two sisters have made some cookies.

### 10. Late auxiliary transposition

- 10.1) While talking to Jane, Joseph knitting \* was a sweater.
- 10.2) While baby-sitting for their neighbors, Mrs. Johnson's daughters were eating some candy.
- 10.3) A small and harmless black dog chasing \* was chickens.
- 10.4) The large and pale gray cruise ship was hitting an iceberg.
- 10.5) My old junior high school friend's favorite little cousin watching \* was cartoons.
- 10.6) My old army friend's beautiful, bright red sports car was burning oil.
- 10.7) In music class, two students singing \* were songs.

- 10.8) Near the big, old summer house, several animals were drinking water.
- 10.9) Bugs have eaten the tomato plants that Martin growing \* was.
- 10.10) They are eating the candy bars that Mrs. Morton has brought.
- 10.11) In a large, old, silver car, several boys driving \* were recklessly.
- 10.12) In a big, old, red boat, two girls were rowing slowly.
- 10.13) Those pilots had said that several clouds covering \* were the sky.
- 10.14) The train conductor was saying that some trash had blocked the tracks.

### 11. Early determiner transposition

- 11.1) Man \* that was reading some books while staying at the hotel.
- 11.2) The boy was reading a comic book while standing on the corner.
- 11.3) Guest \* the was eating a cheese and sausage pizza.
- 11.4) The artists were selling several small but expensive watercolor paintings.
- 11.5) Students \* several were buying some cheap French cheese.
- 11.6) Those models were wearing that new wave hairstyle.
- 11.7) Women \* three are opening a small shop in the city.
- 11.8) Jane's mother is renting a small apartment in New York.
- 11.9) President \* the was reading the report that his advisor had written.
- 11.10) The doctor is reading the medical report that her nurse has written.
- 11.11) Helicopter \* a was hovering loudly over the army base.
- 11.12) A plane was flying slowly over the old landing strip.
- 11.13) Announcer \* the is saying that a big accident

has blocked one lane.

11.14) Sam's friend is saying that his two sisters have made some cookies.

## 12. Late determiner transposition

12.1) Those girls were watching the bright lighting while camping in desert \* that.

12.2) Mrs. Taylor was eating a turkey sandwich while talking on the phone.

12.3) The art museum's owner was buying painting \* a.

12.4) Several very young children were watching a play.

12.5) George's last remaining dinner guest was smoking cigar \* a.

12.6) Her two favorite great aunts were making some pie.

12.7) The magazine reporter was donating one hundred dollars to hospitals \* those.

12.8) The police officer was giving a speeding ticket to that guy.

12.9) The man whom Jack's sister has dated is cleaning car \* the.

12.10) The woman whom Anne's father has hired is cleaning the windows.

12.11) Some drunk men were dancing wildly in streets \* the.

12.12) A small boy was walking slowly down the beach.

12.13) Jerry is hoping that his friends have visited lawyer \* a.

12.14) Chris is saying that his mother has bought a house.

## Appendix III: Fillers

### Grammatical fillers:

1. Those teachers were reading.
2. Sherry was eating a pie.
3. They have talked to John.
4. They were watching some movies.
5. Don spoke to her and laughed.
6. Jim's cousin was on Jack's mind.
7. Once again, Rob planned his vacation late.
8. She instructed her secretary to hold all calls.
9. That woman having a drink is my teacher.
10. That man trying to fix your bike broke mine.
11. While the economy appears sluggish, certain parts are improving.
12. The girl walking quickly to the store was Jim's daughter.
13. Sally believed that she had a detailed knowledge of car engines.
14. The man displayed a fuzzy toy that delighted the little boy.
15. The two women examining the Queen's jewelry collection were famous gem collectors.
16. That little boy climbing the oak tree over there is her son.
17. Steve said that he was promoted quickly because he had worked so hard.
18. The most recent of the conferences differed from those others on several important points.
19. The weather a week ago Saturday, rain, and lots more rain to come, was depressing.
20. What began worrying people in town was the opening of a third huge and sprawling shopping mall.
21. By the time Mrs. London was through, the restaurant had become one of the most popular spots in town.
22. Mr. Harrison, the first successful publisher and

editor of the *Times*, would seem to be one of these entrepreneurs.

### Ungrammatical fillers

1. Jane have\* walked.
2. A horse were\* running.
3. Walk to that houses. \*
4. Three cats\* drinking.
5. Jack was fixing car\* a.
6. Will talked has\* to her.
7. Sam appeared to be thinking hardly. \*
8. Other administration officials calls\* the Green Berets hostages.
9. John seemed to be thinking as he walked aloud. \*
10. One in\* my friends is often working quite late.
11. Who the fact\* that Jim was leaving for good surprised?
12. What were John's parents stating the hope\* that he would get?
13. What did his sister accept the story\* that Mike was delayed by?
14. Which team did the people believe the claim would\* win this year?
15. The three businessmen entering the bank lobby were actually robber. \*
16. Holds in the ship is\* so big that you could store a house.
17. Ellen read, while traveling on the train, several large and complicated company technical reports. \*
18. The rumor that the fight was made up is actually true between\* the two stars.
19. As a resulting\* of her flight delay, Sam's mother was staying in New York an extra night.
20. Last weeks, \* Mary and her two brothers saw a bald eagle flying over the Foothills Fashion Mall.
21. Mrs. Jones was claiming that by the age of two her daughter Carol walking and talking was\* in full sentences.
22. Those film director\* was protesting the destruction of the Amazon rain forest, as are many very famous artists and writers.

## Appendix IV

Sentences were drawn at random from a pool of seven different sentence types. Each cell of the design received one of each of the following sentence types. A sentence demonstrating each sentence type is also given.

1. While clause: "John was eating some cake while talking to Mary."
2. SVO with heavy object: "Her husband was picking a few small, white and yellow daisies."
3. SVO with heavy subject: "My little six-year-old cousin was watching cartoons."
4. SVO with prepositional phrase: "My friend was reading the paper on the express bus."
5. Relative clause: "Meg was reading the book that her mother had written."
6. SV-prepositional phrase with adverb: "A balloon was floating slowly to the ground."
7. Subordinate clause: "Jack was saying that the teacher has graded the tests."

## References

- Bates, E., Friederici, A., & Wulfeck, B. (1987a). Comprehension in aphasia: A Cross-Linguistic Study. *B&L*, 32, 19-67.
- Bates, E., Friederici, A., & Wulfeck, B. (1987b). Grammatical morphology in aphasia: Evidence from three languages. *Cortex*, 23, 545-574.
- Bates, E., & MacWhinney, B. (1987). Competition, variation, and language learning. In B. MacWhinney (Ed.), *Mechanisms of language acquisition*, (pp. 157-193). Hillsdale, NJ: Lawrence Erlbaum.
- Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study of sentence processing*, (pp. 3-73): Cambridge University Press.
- Bates, E., McNew, S., MacWhinney, B., Devescovi, A., & Smith, S. (1982). Functional constraints on sentence processing: A cross-linguistic study. *Cognition*, 11, 245-299.
- Bates, E., Wulfeck, B., & MacWhinney, B. (1991). Cross-linguistic research in aphasia: An overview. *Brain and Language*, 41(2), 123-148.
- Blackwell, A. W., Bates, E., & Fisher, D. (1993). *The time course of grammaticality judgment*. Paper presented at the XVth Conference of the Cognitive Science Society, Boulder, CO.
- Bradley, D. C., Garrett, M. F., & Zurif, E. B. (1980). Syntactic deficits in Broca's aphasia. In D. Caplan (Ed.), *Biological studies of mental processes*, (pp. 269-286). Cambridge, MA: MIT Press.
- Bullinaria, J. A., & Chater, N. (1993). *Double dissociations in artificial neural networks: Implications for neuropsychology*. Paper presented at the XVth Conference of the Cognitive Science Society, Boulder, CO.
- Caplan, D. (1981). On the cerebral organization of linguistic functions: Logical and empirical issues surrounding deficit analysis and functional localization. *Brain and Language*, 14, 120-137.
- Caplan, D., Baker, C., & Dehaut, F. (1985). Syntactic determinants of sentence comprehension in aphasia. *Cognition*, 21, 117-175.
- Caramazza, A. (1986). On drawing inferences about the structure of normal cognitive systems from the analysis of patterns of impaired performance: The case for single-patient studies. *Brain and Cognition*, 5, 41-66.
- Caramazza, A., & Berndt, R. (1985). A multicomponent deficit view of Broca's aphasia. In M. L. Kean (Ed.), *Agrammatism*, (pp. 27-63). Orlando: Academic.
- Caramazza, A., Berndt, R. S., Basili, A. G., & Koller, J. J. (1981). An investigation of repetition and language processing in a case of conduction aphasia. *B&L*, 14, 235-271.
- Caramazza, A., & Zurif, E. B. (1976). Dissociation of

- algorithmic and heuristic processes in language comprehension: Evidence from aphasia. *Brain and Language*, 3, 572-582.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, Holland.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450-466.
- Dell, G., & Reich, P. (1981). Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior*, 20, 611-629.
- Dell, G. S. (1985). Positive feedback in hierarchical connectionist models: Applications to language production. *Cognitive Science*, 9, 3-23.
- Dronkers, N. F., Shapiro, J. K., Redfern, B., & Knight, R. T. (1992). The third left front convolution and aphasia: On beyond Broca., *Paper presented at the Twentieth Annual Meeting of the International Neuropsychological Society*, San Diego, CA.
- Elman, J. (1991). Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning*, 7(2-3), 195-225.
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14, 179-211.
- Farah, M. J., & McClelland, J. (1991). A computational model of semantic memory impairment: Modality specificity and emergent category specificity. *Journal of Experimental Psychology: General*, 120, 339-357.
- Friederici, A. (1988). Agrammatic comprehension: Picture of a computational mismatch. *Aphasiology*, 2, 279-284.
- Friederici, A. D., & Graetz, R. (1984). Processing passive sentences in aphasia: Deficits and strategies, *Paper presented at Seventh INS European Conference, Aachen*.
- Garrett, M. (1992). Disorders of lexical selection. *Cog*, 42, 143-180.
- Geschwind, N. (1972). Language and the brain. *Scientific American*, 226(4), 76-83.
- Goodglass, H. (1993). *Understanding aphasia*. San Diego, CA: Academic Press.
- Grier, J. B. (1971). Non-parametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin*, 75(6), 424-429.
- Grodzinsky, Y. (1982). *Syntactic representations in agrammatism: Evidence from Hebrew*. Paper presented at the Abstracts of the 20th Meeting of the Academy of Aphasia, Mohonk, NY.
- Grodzinsky, Y. (1986). Language deficits and the theory of syntax. *Brain and Language*, 27, 135-159.
- Grodzinsky, Y. (1990). *Theoretical perspectives on language deficits*. Cambridge, MA: MIT Press.
- Grodzinsky, Y. (1993a). Grammatical investigations of aphasia: Introduction. *Brain and Language*, 45(3), 299-305.
- Grodzinsky, Y., Swinney, D., & Zurif, E. (1985). Agrammatism: Structural deficits and antecedent processing disruptions. In M. L. Kean (Ed.), *Agrammatism*, (pp. 65-81). Orlando: Academic Press, Inc.
- Grodzinsky, Y. E. (1993b). Special issue: Grammatical investigations of aphasia. *Brain and Language*, 45(3), 299-464.
- Harley, T. (1993). Phonological activation of semantic competitors during lexical access in speech production. *Lang & Cog*, 8(3), 291-309.
- Hart, J. J., Berndt, R. S., & Caramazza, A. (1985). Category-specific naming deficit following cerebral infarction. *Nature*, 316, 439-440.
- Heeschen, C. (1985). Agrammatism versus paragrammatism: A fictitious opposition. In M. L. Kean (Ed.), *Agrammatism*, (pp. 207-248). Orlando: Academic Press, Inc.
- Heilman, K. M., & Scholes, R. J. (1976). The nature of comprehension errors in Broca's, Conduction and Wernicke's aphasics. *Cortex*, 12, 258-265.
- Hickok, G. (1992). Agrammatic comprehension and the trace-deletion hypothesis, *Occasional Paper No. 45, MIT Center for Cognitive Science. MIT, Cambridge, MA*.
- Hickok, G., Zurif, E., & Canseco-Gonzales, E. (1993). Structural description of agrammatic comprehension. *Brain and Language*, 45(3), 371-395.
- Hillis, A., & Caramazza, A. (1994). *Lexical organization of the brain* (Technical report Report #43): Johns Hopkins Univ., Cognitive Neuropsychology Laboratory.
- Kean, M. L. (1979). Agrammatism: A phonological deficit? *Cognition*, 7, 69-84.
- Kilborn, K. (1991). Selective impairment of grammatical morphology due to induced stress in normal listeners: Implications for aphasia. *Brain and Language*, 41, 275-288.
- King, J., & Just, M. A. (1991). Individual differences in syntactic processing: the role of working memory. *JML*, 30, 580-602.
- Kolk, H., van Grunsven, M., & Guper, A. (1982). On parallelism in agrammatism: A case study., *Unpublished manuscript*.
- Kolk, H., & van Grunsven, M. J. (1984). Metalinguistic judgments on sentence structure in agrammatism: A matter of task interpretation. *Neuropsychologia*, 22(1), 31-39.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge: MIT Press.
- Linebarger, M. (1989). Neuropsychological evidence for linguistic modularity. In G. N. Carlson & M. K. Tanenhaus (Eds.), *Linguistic structure in language processing*, (pp. 197-238). Dordrecht: Kluwer.

- Linebarger, M., Schwartz, M., & Saffran, E. (1983). Sensitivity to grammatical structure in so-called agrammatic aphasics. *Cognition*, 13, 361-392.
- Lukatela, K., Crain, S., & Shankweiler, D. (1988). Sensitivity to inflectional morphology in agrammatism: Investigation of a highly inflected language. *Brain and Language*, 33, 1-15.
- MacWhinney, B., & Bates, E. (Eds.). (1989). *The Crosslinguistic Study of Sentence Processing*. Cambridge: Cambridge Univ. Press.
- MacWhinney, B., & Osmán-Sági, J. (1991). Inflectional marking in Hungarian aphasics. *Brain and Language*, 41(2), 165-183.
- MacWhinney, B., Osmán-Sági, J., & Slobin, D. (1991). Sentence comprehension in aphasia in two clear case-marking languages. *Brain and Language*, 41(2), 234-249.
- MacWhinney, B., Pléh, C., & Bates, E. (1985). The development of sentence interpretation in Hungarian. *Cognitive Psychology*, 17, 178-209.
- Marchman, V. (1993). Constraints on plasticity in a connectionist model of the English past tense. *Journal of Cognitive Neuroscience*, 5(2), 215-234.
- Mauner, G., Fromkin, V. A., & Cornell, T. L. (1993). Comprehension and acceptability judgments in agrammatism: Disruptions in the syntax of referential dependency. *Brain and Language*, 45(3), 340-370.
- McClelland, J. L., & Rumelhart, D. E. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 2: Psychological and biological models*. (Vol. 2). Cambridge: MIT Press.
- Menn, L., & Obler, L. K. (Eds.). (1990). *Agrammatic aphasia: A cross-linguistic narrative sourcebook*. Amsterdam: Benjamins.
- Miceli, G., Mazzucchi, A., Menn, L., & Goodglass, H. (1983). Contrasting cases of Italian agrammatic aphasia without comprehension disorder. *Brain and Language*, 19, 65-97.
- Miceli, G., Silveri, M., Romani, C., & Caramazza, A. (1989). Variation in the pattern of omissions and substitutions of grammatical morphemes in the spontaneous speech of so-called agrammatic aphasics. *Brain and Language*, 36(447-492).
- Miyake, A., Carpenter, P., & Just, M. A. (in press). A capacity approach to syntactic comprehension disorders: Making normal adults perform like aphasic patients. *Cognitive Neuropsychology*.
- Naeser, M., Haas, G., Auerbach, S., Helm-Estabrooks, N., & Levine, H. (1984). Relationship between lesion extent in "Wernicke's area" on computed tomographic scan and predicting recovery of comprehension in Wernicke's aphasia. *Archives of Neurology*, 44(1), 73-82.
- Nespoulous, J.-L., Dordain, M., Perron, C., Ska, B., Bub, D., Caplan, D., Mehler, J., & Lecours, A. R. (1988). Agrammatism in sentence production without comprehension deficits: reduced availability of syntactic structures and/or of grammatical morphemes? A case study. *B&L*, 33(2), 273-295.
- Norman, D., & Bobrow, D. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, 7(44-64).
- Plaut, D. C. (in press). Double dissociation without modularity: Evidence from connectionist neuropsychology. *Journal of Clinical and Experimental Neuropsychology*.
- Pollack, I., & Norman, D. A. (1964). A non-parametric analysis of signal detection experiments. *Psychonomic Science*, 1, 125-126.
- Prather, P., Shapiro, L., Zurif, E., & Swinney, D. (1991). Real-time examinations of lexical processing in aphasics. *Journal of Psycholinguistic Research*, 20, 271-281.
- Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 1: Foundations*. (Vol. 1). Cambridge: MIT Press.
- Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge [England]  
New York: Cambridge University Press.
- Shankweiler, D., Crain, S., Gorrell, P., & Tuller, B. (1989). Reception of language in Broca's aphasia. *Language and Cognitive Processes*, 4(1), 1-33.
- Shapiro, L. P., Gordon, B., Hack, N., & Killackey, J. (1993). Verb-argument structure processing in complex sentences in Broca's and Wernicke's aphasia. *Brain and Language*, 45(3), 423-447.
- Slobin, D. (1991). Aphasia in Turkish: Speech production in Broca's and Wernicke's patients. *Brain and Language*, 41(2), 149-164.
- Smith, S., & Bates, E. (1987). Accessibility of case and gender contrasts for assignment of agent-object relations in Broca's aphasics and fluent anomics. *Brain and Language*, 30, 8-32.
- Smith, S., & Mimica, I. (1984). Agrammatism in a case-inflected language: Comprehension of agent-object relations. *Brain and Language*, 13, 274-290.
- Stemberger, J. (1982). Syntactic errors in speech. *Journal of Psycholinguistic Research*, 11, 313-345.
- Stemberger, J. P. (1985). Bound morpheme loss errors in normal and agrammatic speech: One mechanism or two? *Brain and Language*, 25(2), 246-256.
- Stemberger, J. P. (1989). Speech errors in early child language production. *Journal of Memory and Language*, 28, 164-188.
- Stemberger, J. P., & MacWhinney, B. (1986). Form-oriented inflectional errors in language processing. *Cognitive Psychology*, 18(3), 329-354.
- Warrington, E. K., & Shallice, T. (1984). Category-specific semantic impairments. *Brain*, 107, 829-854.
- Wulfeck, B., & Bates, E. (1991). Differential sensitivity to errors of agreement and word order in Broca's

- aphasia. *Journal of Cognitive Neuroscience*, 3, 258-272.
- Wulfeck, B., Bates, E., & Capasso, R. (1991). A cross-linguistic study of grammaticality judgments in Broca's aphasia. *Brain and Language*, 41(2), 311-336.
- Wulfeck, B. B. (1987). *Sensitivity to grammaticality in agrammatic aphasia: processing of word order and agreement violations*. Unpublished Doctoral dissertation, University of California, San Diego.
- Zurif, E., Swinney, D., Prather, P., Solomon, J., & Bushell, C. (1993). An on-line analysis of syntactic processing in Broca's and Wernicke's aphasia. *Brain and Language*, 45(3), 448-464.
- Zurif, E. B., & Caramazza, A. (1976). Psycholinguistic structures in aphasia. In H. Whitaker & H. A. Whitaker (Eds.), *Studies in Neurolinguistics*, (Vol. 1, pp. 261-292). New York: Academic.
- Zurif, E. B., & Grodzinsky, Y. (1983). Sensitivity to grammatical structure in agrammatic aphasics: A reply to Linebarger, Schwartz, and Saffran. *Cognition*, 15, 207-221.