

## Agrammatic Production of Subject–Verb Agreement: The Effect of Conceptual Number

Robert J. Hartsuiker, Herman H. J. Kolk, and Wendy J. Huinck

*Nijmegen Institute for Cognition and Information, University of Nijmegen,  
Nijmegen, The Netherlands*

Three experiments tested the hypothesis that the deficit underlying agrammatic sentence production difficulties can be characterized as a limitation of computational resources and that these resources are not restricted to syntactic processing. This hypothesis was tested by eliciting subject–verb agreement errors in a sentence fragment completion paradigm. Sentence fragments were complex noun phrases, containing a subject (head) noun and a modifying prepositional phrase, containing a “local” noun. We varied the number of “tokens” a singular head noun referred to. Therefore, in one condition, grammatical and conceptual number of the head noun mismatched, whereas these numbers were the same in another condition. In Experiments 1 and 2, we observed an effect of this variable (i.e., more agreement errors when conceptual number was plural and grammatical number singular) in normal controls. Broca’s aphasics, on the other hand, showed no effect. Experiment 3 consisted of a sentence/picture matching test. This test showed that the lack of effect with Broca’s aphasics cannot be attributed to a comprehension deficit. We argue that these results are incompatible with the notion of a limitation in resources specific for syntactic processing. Instead, we interpret this as the result of a trade-off: Broca’s aphasics lack computational resources to take into account both grammatical and conceptual information in morphosyntactic processing and rely on grammatical information only. © 1999 Academic Press

In the last decade, an increasing number of researchers have proposed that agrammatics’ difficulties in the comprehension and production of sentences

This research was funded by Grant 575-56-074 from the Dutch Organization for Scientific Research (NWO). Experiment 1 was also reported in the first author’s dissertation. Some of the present results were reported at the Academy of Aphasia meeting, Santa Fe, 1998. We thank Gerdien Kikstra for assistance with scoring and Rogier Donders for his help with the analysis of Experiments 2 and 3. We thank Roelien Bastiaanse, Kay Bock, David Caplan, Gerard Kempen, Pim Levelt, Patrick Oor, Herbert Schriefers, Gabriella Vigliocco, and two anonymous reviewers for valuable comments on previous presentations of parts of this study or previous versions of this article.

All correspondence should be addressed to Robert J. Hartsuiker, NICI, University of Nijmegen, PO Box 9104, 6500 HE Nijmegen, The Netherlands. E-mail: [hartsuiker@nici.kun.nl](mailto:hartsuiker@nici.kun.nl). Fax: +31 24 3616066.



can be characterized as the result of a limitation in computational resources (e.g., Caplan, Baker, & Dehaut, 1985; Caplan & Hildebrandt, 1988; Carpenter, Miyake, & Just, 1994; Frazier & Friederici, 1991; Hagiwara, 1995; Hartsuiker & Kolk, 1998; Kolk, 1995; Kolk & Van Grunsven, 1985; Linebarger, Schwartz, & Saffran, 1983; Martin & Romani, 1994; Miyaki, Carpenter, & Just, 1994, 1995). However, there is considerable debate about the specificity of these resources. For instance, Just and Carpenter (1992) presented a theory of a general verbal working memory which is used for different levels of processing within sentence comprehension (however, they do suggest that there is a separation between resources applying to production and comprehension). On the other hand, Martin and colleagues (Martin, 1995; Martin & Romani, 1994; Martin, Shelton, & Yaffee, 1994) conceive of resources as specific for each level of processing. For instance, Martin and Romani (1994) report upon one aphasic patient (A.B.) who shows decreased performance with memory load in a sentence anomaly judgment task, but not in a grammaticality judgment task, and another aphasic patient (M.W.) who shows the reverse pattern of results. Caplan and colleagues (e.g., Caplan & Waters, 1995; Waters & Caplan, 1996) favor a different fractionating of resources, i.e., a fractionating into one pool that is used for online computation of sentence structure and another pool that is involved in conscious, verbally mediated tasks. However, they leave open the question of whether these putative resource pools should be further divided (Waters & Caplan, 1996).

In the present article we focus on the question of how specific the resource limitation is in sentence *production*. We test Broca's aphasics in a constrained production task, which requires morphosyntactic processing: constructing agreement between subject and verb. Getting agreement right is necessary in almost every sentence a speaker produces (in most of the world's languages). It is usually unproblematic, even in patients who produce many paragrammatisms in spontaneous speech (Butterworth, Panzeri, Semenza, & Ferreri, 1990). We present evidence for the claim that Broca's aphasics suffer from an impairment in verbal working memory resources and that these limited resources are not specific to syntax, but are involved both in syntactic and semantic processes.

One of the earliest papers in which the notion of a resource limitation was proposed was a study by Linebarger et al. (1983). In this study, four agrammatic patients were required to judge syntactically anomalous sentences. All of these patients showed agrammatic comprehension in picture-pointing tests. Quite surprisingly, these patients were able to correctly judge grammaticality of a wide variety of constructions. This finding clearly disagrees with many theories of agrammatic comprehension which seem to locate the agrammatic deficit in part of the parsing device itself (Berndt & Caramazza, 1980; Grodzinsky, 1986). Linebarger et al. suggested two possible explanations of the discrepancy between grammaticality judgment and comprehension. First of all, the deficit could reside in the mapping of syntactic represen-

tations on semantic objects such as thematic roles, rather than in parsing. Second, because of a lack of computational resources, there could be a trade-off between syntactic and semantic processing. A task requiring both syntactic and semantic analysis would overload the processing capacity, but a task requiring mainly syntactic analysis or mainly semantic analysis would not. These two hypotheses have received attention in a number of subsequent studies (Kolk & Weijts, 1996; Schwartz, Linebarger, Saffran, & Pate, 1987). Schwartz et al. (1987) rejected the trade-off hypothesis because they failed to find an effect of syntactic complexity ("padding" sentences) in a semantic anomaly judgment task. However, in line with the mapping hypothesis, they did obtain effects of *argument movement*. They reasoned that argument movement leads to a more difficult mapping and that if agrammatic comprehenders have problems in mapping syntactic structure to semantic roles, sentences with argument movement should indeed be more difficult. Recently, Kolk and Weijts (1996) presented evidence for the opposite view, that the trade-off hypothesis is in fact more plausible than the mapping hypothesis. They argued that Schwartz et al.'s complexity manipulation (conjoining sentences) was insufficient. They replicated Schwartz et al.'s experiment, but added a new condition: center-embedded sentences. In the latter condition, there were complexity effects, as predicted by the trade-off hypothesis. Furthermore, Kolk and Weijts (1996) proposed that the trade-off hypothesis may even be a better alternative than the mapping hypothesis, because the mapping hypothesis makes no predictions concerning phrase-structure complexity, whereas the trade-off hypothesis predicts effects of phrase-structure complexity as well as effects of movement. Indeed, a computational model of parsing, based on Just and Carpenter's (1992) theory of sentence comprehension, successfully simulates effects of phrase-structure complexity as well as effects of movement (Haarmann, Just, & Carpenter, 1997).

Because the trade-off hypothesis thus seems a reasonable alternative to the mapping hypothesis as an account of agrammatic comprehension, we consider it an important question whether evidence for a similar kind of trade-off can be obtained in agrammatic sentence *production*. In order to test the trade-off hypothesis in production, we need to show that there are situations in which normal speakers simultaneously apply information from different components. Therefore, in the remainder of this section, we first discuss some examples of production tasks in which speakers rely on syntactic as well as semantic information. Then we discuss an example from sentence comprehension in which a trade-off between syntactic and semantic processing occurs when resources can be assumed to be scarce. Finally, we present our specific hypothesis and show how we test it.

The first example of usage of multiple information components deals with computation of the correct form of a reflexive pronoun. Bock and colleagues (Bock, 1995; Bock, Eberhard, & Cutting, 1992) reported evidence for *semantic* control of constructing agreement between reflexive pronouns or the pro-

nouns in tag questions with their antecedents in production. In their experiment, subjects had to repeat sentence fragments and complete them using reflexive pronouns or tag questions. Sentence fragments consisted of a complex Noun Phrase (NP) and a main verb not marked for number (1a and 1b). The critical manipulation concerned the conceptual number of the head noun. Sometimes the head noun was collective, i.e., conceptually plural (1a) and sometimes it was conceptually singular (1b). The dependent variable was the proportion of pronoun number agreement errors. The results were clear-cut: Following a conceptually plural head noun, participants were more likely to make errors in tag questions and reflexive pronouns, both within and across clauses.

(1a) The gang with the dangerous rivals armed . . .

(1b) The gangster with the dangerous rivals armed . . .

These results support the notion that the process of constructing pronoun agreement accesses the discourse model. However, it cannot be the case that this process is entirely driven by semantics. Consider the following examples, taken from Pollard and Sag (1994).

(2a) The faculty is voting itself a raise

(2b) The faculty are voting themselves a raise

(2c) \*The faculty are voting itself a raise

(2d) \*The faculty is voting themselves a raise

In British English, a collective noun such as *faculty* allows either a singular verb (2a) or a plural verb (2b). This does not mean that the number of the subject NP is unspecified, because once the speaker has dedicated herself to singularity or plurality, the inflection of the pronoun has to agree with that of the verb. Thus, (2c and 2d) are ungrammatical. A purely semantic theory of pronoun agreement cannot explain these data (for instance, if one has a noncollective notion of “the faculty voting,” but a collective notion of “the faculty receiving a raise,” (2d) should be acceptable). Therefore, we conclude that construction of pronoun agreement relies upon both semantic and syntactic information.

It is important to note here that agrammatic comprehension of sentences with reflexive pronouns and tag questions has received scrutiny in Linebarger et al.’s (1983) study. The agrammatics, while generally being quite capable of performing grammaticality judgments, performed significantly worse on sentences containing tag questions and reflexive pronouns. According to Linebarger et al. (1983), this pattern of results can be expected if the trade-off hypothesis is correct. Problems with reflexives and tag questions follow from the assumption that (lexical) semantic encoding is shallow in a task requiring syntactic analysis. The agrammatic participant has insufficient resources for both parsing the sentence and incorporating information on number and gender in a semantic representation. Therefore, information neces-

sary to evaluate the acceptability of the tag question or reflexive pronoun is missing when such an element is subsequently encountered.<sup>1</sup>

Importantly, evidence for a process that requires the speaker to simultaneously apply semantic and syntactic information has also been obtained with a slightly different task. Analogous to the subject-pronoun agreement experiments described above, semantic influences on the production of subject-*verb* agreement have been studied. Bock and Miller (1991) presented participants with sentence fragments, which had to be repeated and completed. Sentence fragments consisted of complex NPs which contained a head (subject) noun, and a "local noun," embedded within a modifier. For instance, a sentence fragment could be (3).

(3) The king of the colonies

The dependent variable was the proportion of subject-verb agreement errors in the various conditions of these experiments. One manipulation concerned the conceptual number of the head noun. This could be singular, as is its grammatical number, as in (4a). However, conceptual number could also be plural, as in (4b).

(4a) The baby on the blankets

(4b) The label on the bottles

In (4a), there is only one baby. Such a sentence fragment is called a "single token" fragment. In (4b), the head noun *label* has to refer to a multitude of labels (one on each bottle) in order to be in line with our world knowledge. Such items are called "multiple token" or "distributive" items. If, during the process of subject-verb agreement construction, the speaker takes semantic factors such as the conceptual number of a head noun into account, more errors should be observed following a multiple token item than following a single token item. Although such an effect was absent in Bock and Miller's (1991) study with English speaking subjects, Vigliocco and colleagues (Vigliocco, Butterworth, & Semenza, 1995; Vigliocco, Butterworth, & Garrett, 1996a; Vigliocco, Hartsuiker, Jarema, & Kolk, 1996b) have consistently found conceptual effects in Italian, Spanish, Dutch, and French. Although it is at present unclear why effects with conceptual number have not been firmly established in English, the evidence suggests that speakers take both information from the discourse model and grammatical information (number) into account when performing a morphosyntactic process such as construction of subject-verb agreement.

There are also examples from parsing, showing that the process of constituent structure generation takes into account syntactic as well as semantic information (Trueswell, Tanenhaus, & Garnsey, 1994). In these experiments,

<sup>1</sup> Notice, however, that Schwartz et al. (1987) provided a different account of the deviating results with coindexed elements, in terms of a (revised) mapping hypothesis.

sentences that were locally ambiguous with respect to a main clause or reduced relative clause reading were presented in an eye-tracking paradigm. For instance, (5a and 5b) are locally ambiguous upon encountering the verb *examined*.

(5a) The evidence examined by the lawyer . . .

(5b) The defendant examined by the lawyer . . .

These sentence fragments differ in an important respect. The subject noun in (5a), *evidence*, is not a good agent of the verb *to examine*, because it is inanimate. On the other hand, the subject noun of (5b) can be the agent of *to examine*, for instance in (6).

(6) The defendant examined the courtroom.

Do comprehenders take this semantic information into account? If they do, fixation times on the part of the sentence in which the ambiguity is resolved (the by-phrase) should be shorter in those sentences in which semantic information constrains interpretation. Ferreira and Clifton (1986), who were the first to address this question, found no such effect: Fixation times on the by-phrases were equally long. However, this may have been due to a too-weak manipulation of the conceptual variable. Some verbs, similar to *examine* may taken an inanimate entity as an instrument, as in *the car towed*. . . , or may have an ergative reading, as in *the thrash smelled*. . . (examples taken from Trueswell et al., 1994, p. 289). Indeed, Trueswell et al. (1994) did obtain longer reading times for the “defendant” sentences, with different materials.

Just and Carpenter (1992) also found effects in this task. They too eliminated from the materials sentences such as *the car towed*. . . . Moreover, they showed that the occurrence of an effect depended on the verbal working memory capacity of their subjects. They divided the participants into groups with low and high spans of verbal working memory, as determined with the reading span task (Daneman & Carpenter, 1980). The reading span task tests recall of sentence-final words in a paradigm that taxes both hypothesized components of verbal working memory: storage and processing. It turned out that participants with high reading span scores would take into account semantic and syntactic information, but that participants with low spans did not.<sup>2</sup> Just and Carpenter (1992) concluded that verbal working memory can be conceived of as a general pool of resources, used for syntactic as well as

<sup>2</sup> Notice that Waters and Caplan (1996) pointed out that when one compares the reduced relative clauses with the control condition, unreduced relative clauses, one can still maintain that both groups of participants were garden-pathed in both types of sentence. Although this is certainly true, it is still the case that the two groups make differential use of the animacy cue: High-span speakers use the animacy cue, but it is insufficient to completely avoid the garden path. Low-span speakers, on the other hand, do not use the animacy cue. Such a notion of an animacy cue as a “gentle force,” influencing, but not fully determining, garden-pathing, seems compatible with modern constraint-satisfaction approaches (e.g., Trueswell et al., 1994).

semantic processing. When there is enough capacity, one can use both types of information simultaneously. However, when there is not enough capacity, one will rely on only one source of information, the source of information most valid for the task at hand: syntactic information.

Given the assumptions that in principle speakers and comprehenders can take both syntactic and semantic information into account, that the degree to which they are able to do that depends on computational resources, and that Broca's aphasics have a severe limitation in the resources necessary to draw upon different information components, we predict that Broca's aphasics, as opposed to normals, should be unable to apply both semantic and syntactic information in morphosyntactic processes of sentence production. In order to test this hypothesis, we presented a group of agrammatic speakers and a control group of healthy adults with sentence fragments similar to the ones used by Vigliocco et al. (1996b),<sup>3</sup> that is, complex NPs with either a conceptually plural reading or a conceptually singular reading. We will refer to the second (immediately preverbal) noun in the preambles as the 'local' noun. Experimental sentence fragments all had grammatically singular head nouns and grammatically plural local nouns. Control sentence fragments were variants of the experimental sentence fragments, but had singular local nouns. The main motivation for including control conditions is that single and multiple token items are bound to differ with respect to lexical content and perhaps with respect to many other variables. Such idiosyncratic factors may be responsible for differential rates of randomly occurring errors (cf. Bock, 1995). In order to control for such idiosyncratic factors one needs to evaluate any differences between the multiple and single token items with respect to control items that have grammatically singular local nouns. Examples of experimental materials are provided in Table 1.

It is essential to notice that the local noun needs to be plural in order to obtain a conceptually plural reading of the subject. In the remainder of this paper, we will often refer to an effect of conceptual number as a *distributivity* effect. Such an effect thus entails that there are more errors in the multiple token condition than in the single token condition, but no such difference exists between the multiple token control and single token control conditions. Notice that we will talk about a 'multiple token' control condition, even though there is only one label in *the label on the bottle*.

We will also talk about an *attraction* effect. That refers to a main effect of number match between head and local noun, so that there are more agreement errors in the number mismatch conditions (the single token and multiple token conditions) than in the number match (control) conditions.

From the hypothesis of an aspecific, semanticosyntactic resource limitation in Broca's aphasia, it follows that agrammatics are unable to rely on both cues. Instead, when constructing subject-verb agreement they will use

<sup>3</sup> In fact, with one exception, these were the same items used in their Experiment 1.

TABLE 1  
Examples of Experimental Sentence Preambles

Single token, number match (ST, control)	De eigenaar van de koffer (The owner of the suitcase) De discussie over de wet (The discussion about the law)
Single token, number mismatch (ST)	De eigenaar van de koffers (The owner of the suitcases) De discussie over de wetten (The discussion about the laws)
Multiple token, number match (MT, control)	De reclame op de bus (The advertisement on the bus) De bon in de folder (The coupon in the flyer)
Multiple token, number mismatch (MT)	De reclame op de bussen (the advertisement on the busses) De bon in de folders (The coupon in the flyers)

the cue most relevant for the task at hand: the grammatical number of the subject. Notice that it might also be argued that in the face of an aspecific resource limitation the speaker would more often rely on the semantic number instead of the grammatical number. Although we admit this possibility, we do not think it is likely, because in Dutch, the grammatical number always specifies correct agreement whereas the conceptual number does not.<sup>4</sup> In terms of the competition model, proposed by Bates and MacWhinney (1982), grammatical number has a higher cue validity. We assume that in the face of a resource limitation, Broca's aphasics will rely on the more valid cue, grammatical number. Given that assumption, we predict that Broca's aphasics, as opposed to normal controls, are not sensitive to a mismatching conceptual number. In other words, there should not be an increased number of errors in the multiple token condition for Broca's aphasics.

What would follow from an alternative hypothesis, the hypothesis that agrammatics suffer from a limitation in resources specific for syntactic processing or that they have problems with specific syntactic operations (cf. Caplan & Waters, 1995; Martin, 1995)? From such hypotheses, different predictions could be derived. If specific syntactic resources were limited, agrammatics may be unable to maintain the (grammatical) number of the subject. Because number of the subject noun would tend to be forgotten or would be weakly represented relative to the number of the local noun, such a specific resource limitation could result in relatively strong interference from the local noun. A syntactic resource limitation would then lead to a stronger than normal attraction effect. Alternatively, if syntactic operations

<sup>4</sup> Notice for instance that in Dutch, collective nouns like "army" or "faculty" require agreement with grammatical number and not with conceptual number.

were specifically impaired, agrammatics may overrely on semantic information, as they do in comprehension (e.g., Caramazza & Zurif, 1976). That would result in a stronger than normal distributivity effect.

A similar experiment with a single agrammatic patient and a conduction aphasic was reported in Vigliocco, Butterworth, Semenza, and Fossella (1994). For the agrammatic patient, these authors obtained an effect of distributivity: There were more errors in the multiple token condition than in the single token condition. Interestingly, the conduction aphasic showed no effect of distributivity. However, the result is limited to a single agrammatic patient and, according to the test results mentioned by the authors, a patient with a relatively mild impairment. It seems desirable to test a *group* of patients in order to examine whether any distributivity effect occurs consistently across individuals.

## EXPERIMENT 1

### Method

#### *Participants*

Participants were 8 aphasic patients, 6 males and 2 females, diagnosed as Broca's Aphasic on the Dutch version of the Aachen Aphasia Test (AAT) and 12 healthy control participants, 8 males and 4 females, matched in age and educational level to the group of Broca's aphasics. The mean age was 52, both for Broca's aphasics and for normal controls. Demographic information on the group of Broca's aphasics is supplied in Table 2. Furthermore, Table 3 lists some relevant AAT scores. Table 3 shows that the group of Broca's aphasics had relatively spared comprehension and repetition skills. This is important, as the present task required them to (1) comprehend a complex NP and (2) repeat and complete the NP.

#### *Materials*

Experimental materials were sentential fragments consisting of subject NPs followed by Prepositional Phrases (PPs). All experimental preambles had a singular head noun. The number of the local noun was varied. Thus, there were two versions of each item: one with a singular local noun (the number match or control condition) and one with a plural local noun (the number mismatch or experimental condition). Each participant received only one version of each preamble.

Notice that distributivity applied only to singular head noun and plural local noun sentential preambles, where it is possible to contrast a singular to a plural reading of the sentential subject (e.g., one single baby on several blankets or a label for each of several bottles); for items with a singular head noun followed by a singular local noun the preferred reading is congruent with the syntactic characteristics of the subject (e.g., one road to the island or one label on the bottle). A complete listing of materials is provided in Appendix A.

The preferred semantic reading of a larger set of preambles with plural local nouns (i.e., single or multiple token) was evaluated by 10 independent judges, all native speakers of Dutch. Only those items classified as single or multiple token by at least 7 judges were included in the experiment. This test left us with 24 items: 12 judged as single token and 12 judged as multiple token.

Each preamble consisted of the same number of words (five). The number of syllables was on average higher in the single token set than in the multiple token set (7.6 vs. 7.0). This

TABLE 2  
Demographic Information on the Group of Broca's Aphasics

Subject	Experiment	Age	Sex	Years since onset	(Former) Profession	Etiology
NU	1-3	65	m	12	Municipal worker	Infarction of the area of the a. cerebri media
FR	1-3	28	m	5	Truck driver	Large hypodense area in almost complete flow area of a. cerebri media left, with left frontal extensions
HA	1-3	47	m	7	Bricklayer	Contusio cerebri; cerebral edema; cortical atrophy left hemisphere
KL	1-3	64	f	>3	Hairdresser	Infarct, left hemisphere
LO	1-3	44	m	7	Accountant	Intracerebral hematoma, left
TH	1-3	36	m	3	Mover	Infarct in the flow area of the a. cerebri media, left hemisphere
BE	1	70	f	10 months	Cashier	Not available
HO	1	61	m	1	Planner	Not available
BO	2,3	41	f	5	Housewife	Large infarct in the area of the a. cerebri media, left hemisphere
HOE	2,3	62	m	3	Director	Infarct in the capsula interna, left hemisphere
CO	2,3	55	f	2	Housewife	Hemorrhage in the capsula interna, left hemisphere
LA	2,3	71	m	7	Manager	Infarct, left hemisphere, recidive
KO	2,3	66	m	1	Pipefitter	Infarct in the area of the a. cerebri media
GE	2,3	69	m	19	Mechanic	Infarct, left hemisphere

*Note.* The column labeled Experiment shows in which experiment(s) these participants were tested.

difference was not significant ( $t(22) = 1.47; p = .157$ ). Experimental items all contained head and local nouns of the nonneuter gender. Nonneuter nouns require the determiner "de," which is unmarked for number and can therefore be considered as parallel to the English "the."

There were also 36 filler items, 24 of which had the same syntactic structure as the experimental items, but had a plural head noun. Half of these had a singular local noun, the other half had a plural local noun. The remaining 12 fillers were simple NPs, consisting of a determiner, one or more adjectives, and a noun. Simple NP fillers were singular in six cases, and plural in the remaining six cases. In this way, the set of items a participant received (experimental items and filler items) was completely balanced for number of head noun and local noun. Filler items all had head nouns of the nonneuter gender.

We also coupled an adjective to each experimental and filler item. In the sentence types tested here, adjectives are not marked for number. Adjectives were selected by presenting preambles to eight normal subjects (graduate students from Nijmegen University) and requir-

TABLE 3  
Scores on a Subset of AAT Scales

Subject	Experiment	Spontaneous speech (0-5)	Token test (50-0)	Repetition (0-150)	Repetition sentences (0-30)	Auditory comprehension (0-60)
NU	1-3	1	5	?	?	48
FR	1-3	2	13	122	24	46
HA	1-3	2	24	115	21	40
KL	1-3	2	11	120	24	53
LO	1-3	2	33	112	17	50
TH	1-3	3	10	129	19	55
BE	1	2	14	138	23	49
HO	1	2	32	118	23	51
BO	2,3	2	19	97	11	53
HOE	2,3	2	32	118	23	51
CO	2,3	2	30	141	24	51
LA	2,3	2	33	150	30	48
KO	2,3	3	25	126	21	47
GE	2,3	3	14	101	25	21

*Note.* The score on Spontaneous speech indicates, on a scale from 0 (very severe disturbance) to 5 (no or minimal disturbance), the syntactic structure of utterances. Scores of 1 and 2 indicate that no complex sentence structures were observed and that function words and inflections were missing.

ing them to complete them with an adjective. The most frequently supplied adjective was selected. This ranged from at least once to maximally four times of eight. If there were any ties, we chose the adjective we preferred (this happened eight times). Adjectives in the multiple and single token conditions did not differ with respect to number of syllables, number of characters, imageability, frequency, and age of acquisition (norms for the latter three variables were taken from Van Loon-Vervoorn, 1985).

Four 60-item lists were created. Each list consisted of 24 experimental items (6 single token items, 6 single token control items, 6 multiple token items, and 6 multiple token control items) and 36 fillers. In each list the experimental and filler items were organized in a pseudo-random order, with the constraint that the list started with four fillers and that no more than two experimental items followed each other. Across the four lists, each item occurred twice in the match condition and twice in the mismatch condition.

### Procedure

The experimenter (R.H.) read each item aloud. First, he read the adjective and then he read the sentence fragment. The participant was requested to repeat the sentence fragment and complete it to a full sentence using a form of *to be* and the adjective. The main reason for presenting adjectives was to facilitate ease of sentence completion.<sup>5</sup> In those cases where the participant failed to repeat the sentence fragment correctly or when the participant explicitly asked for a repetition, the experimenter read the item again (including the adjective). A second

<sup>5</sup> In the first patient tested, NU, no adjectives were presented. Because he showed considerable difficulty in rapidly coming up with a plausible continuation of the sentence fragments, we decided to present adjectives to the remaining participants.

repetition was only supplied on specific request of the participant. If the participant repeated the preamble correctly, supplied a verb, but had difficulties in recalling the adjective, the experimenter supplied the adjective.

Before the experiment started, at least three example items were presented. If the participant did not understand the task, more examples were provided until he did. If the participant was unable to repeat and complete five example items or more, the experiment was terminated.<sup>6</sup>

### Scoring

The tape-recorded sessions were first transcribed and responses were assigned to one of the following scoring categories: *Correct responses* were scored when the participant said the preamble correctly and produced a correctly inflected verb form in the completion. We allowed for correct sentence completions containing a different adjective, as well as a correct sentence completion with a verb other than *to be*. We also allowed for repetitions of part of the preamble. *Agreement errors* were scored when the utterance adhered to all criteria for a correct response, with one exception: The verb form was wrongly inflected for number. *Number repetition errors* were scored when the participant incorrectly reported the number of the sentential subject and then produced a verb form inflected correctly with the produced number of the subject. *Repetition plus agreement errors* were scored when the participant changed the number of the head noun but then the verb form mismatched in number with the produced subject. Finally, utterances were scored as *miscellaneous responses* when even after two or more presentations of the preamble the participant failed to report the preamble, omitted the subject noun or local noun, changed the preposition, changed the number of the local noun, produced a sentence with a postverbal subject, or produced an incomplete sentence. If two different utterances were produced in succession, only the first was scored. This included cases in which an agreement error was produced and immediately corrected.

### Design

We conducted a global analysis of variance, with group (Broca's aphasics, normal controls) as a between-subjects variable. We also carried out separate analyses for each group. The number of agreement errors and the number of miscellaneous responses constituted the dependent variables for the statistical tests. The experimental factors orthogonally combined were (1) item type (single token, multiple token) and (2) number match (match, mismatch). The combination of these two factors yielded four conditions. Every participant received six items in each condition.

## Results

For the group of normal controls ( $N = 12$ ), application of our scoring criteria yielded 230 (79.9%) correct responses, 50 agreement errors (17.4%), 7 miscellaneous responses<sup>7</sup> (2.4%), and 1 number repetition error (0.3%). For the group of Broca's aphasics ( $N = 8$ ), there were 83 correct responses

<sup>6</sup> This happened with four patients, who were therefore excluded from further participation.

<sup>7</sup> Notice that there are far fewer miscellaneous responses and repetition errors here than in Vigliocco et al. (1996b), where the same materials were presented to college students. There are two reasons for this difference: first, in order to elicit more errors, they presented their subjects with a deadline. Many of their miscellaneous responses were cases in which the speaker failed to produce a verb before the deadline. Second, in the current experiment a wrongly repeated preamble was presented another time by the experimenter.

TABLE 4  
Distribution of the Responses in the Scoring Categories: Experiment 1

	<i>n</i>	Correct	Agreement errors	Miscellaneous responses	Number of repetition errors	Agreement/repetition errors
Normal controls						
ST control	72	71	0	1	0	0
ST	72	51	16	4	1	0
MT control	72	70	1	1	0	0
MT	72	38	33	1	0	0
Broca's aphasics						
ST control	48	27	3	18	0	0
ST	48	18	17	11	1	1
MT control	48	26	13	9	0	0
MT	48	12	24	11	1	0

*Note.* The experimental conditions were ST control (single token, number match); ST (single token, number mismatch); MT control (multiple token, number match); MT (multiple token, number mismatch).

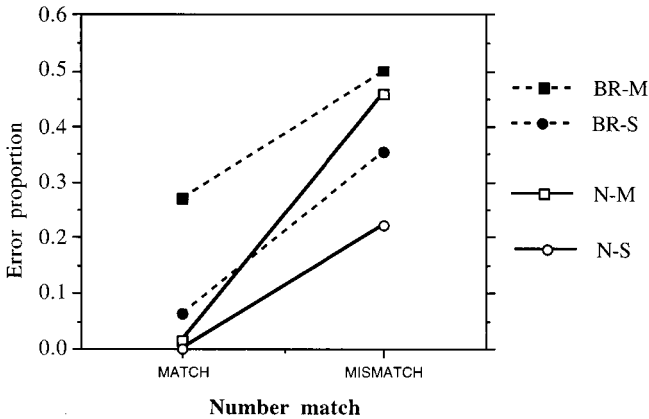
(43.2%), 57 agreement errors (29.7%), 49 miscellaneous responses (25.5%), 2 number repetition errors (1.0%), and 1 (0.5 %) repetition/agreement error.

The distribution of responses in the different scoring categories for the experimental conditions is listed in Table 4.

Let us first consider the control participants. Table 4 shows that agreement errors were most common (33) for multiple token preambles, while we found only 16 errors in the single token condition and respectively 1 error and no errors in the control conditions. This pattern of results is similar to the data Vigliocco et al. (1996b) obtained with the same items, with university students as participants (multiple token 52, single token 6, both control conditions 0). Now consider the distribution of responses for Broca's aphasics. Broca's aphasics also made more errors in the multiple token condition (24) than in the single token condition (17). However, there was a difference of similar magnitude between the two control conditions: Errors were more frequent in the multiple token control condition (13) than in the single token control condition (3). In other words, there does not seem to be an interaction between item type (MT, ST) and number match (match, mismatch).

The proportions of agreement errors in the different conditions are depicted in Fig. 1.

Figure 1 shows the proportion of agreement errors in the multiple and single token match and mismatch conditions. Proportions were calculated relative to the total number of observations in each cell. First consider the filled lines, which depict the error proportions for normal controls. The figure suggests that there is an interaction, so that there is a difference between the multiple and single token items in the mismatch condition, but not in the



**FIG. 1.** Proportions of agreement errors (computed as percentage of all responses). Interrupted lines: BR-M (Broca's aphasics, multiple token); BR-S (Broca's aphasics, single token); filled lines: N-M (normal controls, multiple token); N-S (normal controls, single token).

match condition. Now consider the interrupted lines, which depict the error proportions for the Broca's aphasics. These lines almost run parallel to each other. There appear to be main effects of item type and of number match, but no interaction.

These observations were confirmed in the statistical analyses. Because the normal controls produced zero agreement errors in the single token control condition, and only one agreement error in the multiple token control condition, an ANOVA that included those cells would not be appropriate. Therefore, in the global analyses, we analyzed the *difference* scores (number of errors in the mismatch condition minus number of errors in the match condition). It is essential to note that with the data cast in this format, an effect of conceptual number would be reflected as a main effect of item type. Further, if there is a conceptual number effect for the normal controls, but not for the Broca's aphasics than we expect a significant Group  $\times$  Item type interaction.

Thus, we conducted a  $2 \times 2$  repeated measures ANOVA on the difference scores, with group as a between-subjects variable with two levels (Broca's aphasics,  $N = 8$ ; Controls,  $N = 12$ ) and item type as a within-subjects variable with two levels (MT, ST). This analysis yielded no effect of group ( $F < 1$ ), no effect of item type [ $F(1, 18) = 2.06$ ;  $MSE = 2.06$ ;  $p = .168$ ] but a significant Group  $\times$  Item type interaction [ $F(1, 18) = 6.54$ ;  $MSE = 2.06$ ;  $p = .02$ ]. This interaction shows that item type has a different effect for the two groups.

We also conducted separate analyses *within* each group. First, we conducted an analysis within the group of normal controls on the difference scores between the mismatch and match conditions. The sum of the differ-

ence scores for normal controls was 32 in the multiple token condition and 16 in the single token condition. A repeated measures ANOVA with item type as a within-subjects variable with two levels (ST, MT) showed that the effect of item type was significant [ $F(1, 11) = 16.00$ ;  $MSE = .67$ ;  $p < .002$ ]. Thus, normal controls show an effect of conceptual number. Second, we conducted a similar analysis on the difference scores for the Broca's aphasics. The sum of the difference scores was 14 in the single token condition and 11 in the multiple token condition. That difference was, however, far from significant [ $F(1, 7) < 1$ ;  $MSE = 1.71$ ;  $p = .584$ ].

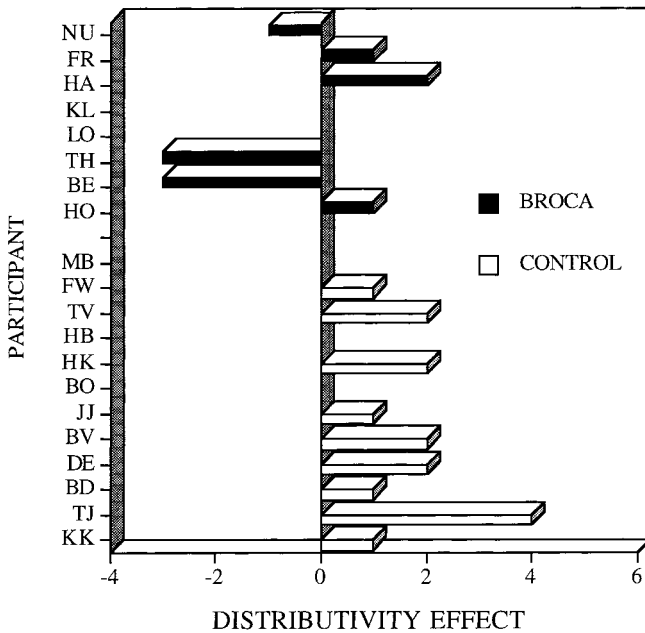
In sum, the overall analysis on difference scores show a significant interaction between group of participants and item type. Tests within each group confirm that there is an effect of item type for normal controls, but not for Broca's aphasics. Thus, normal controls show an effect of conceptual number, but normal controls do not.

Because there were sufficient numbers of agreement errors in each cell for the group of Broca's aphasics, we also performed an analysis in that group on the *raw numbers* of agreement errors for each participant in each condition, instead of with difference scores. We conducted a  $2 \times 2$  repeated measured ANOVA, with item type (ST, MT) and number match (match, mismatch) as within-subjects variables. With the data cast in this format, an effect of conceptual number is reflected by an *interaction* between these two variables, because there is conceptual plurality only in the multiple token mismatch condition.

The analysis on the raw numbers of errors showed a significant effect of item type [ $F(1, 7) = 14.99$ ;  $MSE = .60$ ,  $p < .01$ ]. There were significantly more agreement errors in the multiple token and its control condition than in the single token and single token control conditions. The effect of number match was also significant [ $F(1, 7) = 14.07$ ;  $MSE = 1.39$ ;  $p < .01$ ]. There were significantly more errors in the experimental conditions than in the control conditions. However, the Item type  $\times$  Number match interaction was far from significant [ $F(1, 7) = .33$ ;  $MSE = .85$ ;  $p = .584$ ]. This analysis thus confirms the findings from the analysis on difference scores: There is no effect of conceptual number for Broca's aphasics.

Somewhat surprisingly, there was a significant effect of item type for Broca's aphasics. To further explore the differences between the conditions, we conducted separate analyses comparing the number of errors for multiple token and single token items within each number match condition. In the condition with plural local nouns, there was no effect of item type [ $F(1, 7) = 3.33$ ;  $MSE = .92$ ;  $p = .111$ ]. In the condition with singular local nouns, however, the effect was highly significant [ $F(1, 7) = 11.67$ ;  $MSE = .54$ ,  $p = .01$ ].

We also conducted an analysis of variance on the number of miscellaneous responses for Broca's aphasics. There were more responses of that type in the single token conditions than in the multiple token conditions. This test

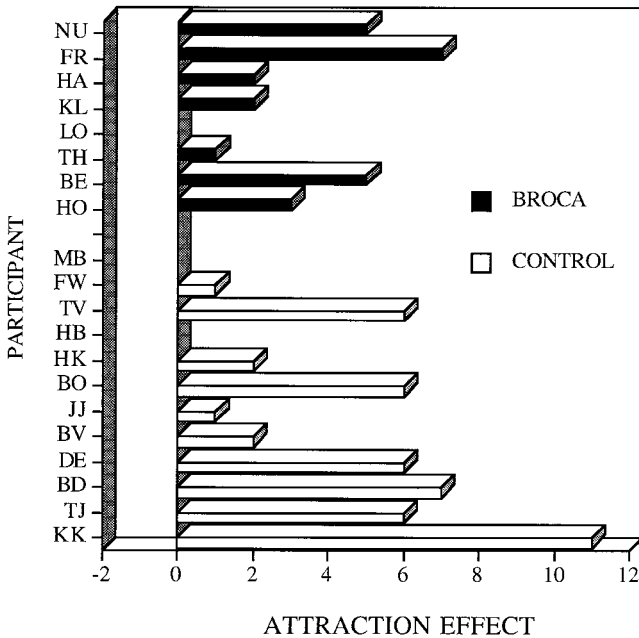


**FIG. 2.** Individual scores of distributivity effect in Experiment 1, in absolute numbers of errors: Multiple token effect (multiple token condition minus multiple token control condition) minus single token effect (single token condition minus single token control condition). The effect is zero where no bar is plotted.

revealed a marginally significant effect of item type [ $F(1, 7) = 5.50$ ;  $MSE = .46$ ;  $p = .051$ ]. There was no effect of local noun number nor was the interaction between these variables significant. We did not perform a comparable analysis with normal controls, because there were too few miscellaneous responses in that group.

### *Individual Data*

The global analyses reported above established that conceptual number has a different effect within the two groups of participants. Normal controls show an effect of this variable, but Broca's aphasics do not. How consistent are these findings over individuals? Figure 2 shows the effect of distributivity for each individual. 'Effect' refers here to the difference in error frequency between the multiple and single token conditions, with error frequencies in the respective control condition subtracted from each experimental condition. In other words, these are differences between the difference scores that were analyzed in the global ANOVAs. A positive effect thus indicates that items with a distributed reading elicited more errors than nondistributed items, when idiosyncratic effects of item type are controlled for. As is clear



**FIG. 3.** Individual scores of attraction effect in Experiment 1, in absolute numbers of errors: Number of mismatch conditions (multiple token condition plus single token condition) minus number match condition (multiple token control condition plus single token control condition). The effect is zero where no bar is plotted.

from the figure, 9 of 12 normal controls show a positive effect and 3 normal controls show no effect. The data are quite different for the Broca's aphasics: 3 of them show a positive effect, 3 show a negative effect, and 2 show no effect at all. In other words, there is a rather consistent positive effect with normal controls, but the effect for Broca's aphasics varies around zero. There is certainly not a stronger distributivity effect for Broca's aphasics, as might be predicted by the hypothesis that Broca's aphasics overrely on semantic information.

What about the effect of number mismatch (the "attraction effect") for each individual? Figure 3 shows the number of errors in the number mismatch conditions (single token and multiple token conditions) minus the number of errors in the number match conditions (the control conditions). As is clear from the figure, these effects are very similar in the two groups. All of the participants produced more errors in the mismatch conditions than in the match conditions, except for one Broca's aphasic and two normal controls, who had effects of zero. Furthermore, the size of the effect for Broca's aphasics lies completely in the range of the normals. It is certainly not the case that Broca's aphasics have an exaggerated attraction effect, as

could be predicted from the hypothesis of a limitation in a dedicated syntactic resource.

In the experimental conditions, Broca's aphasics made subject-verb agreement errors on about 40% of all trials (ranging from 35% in the single token control condition to 50% in the multiple token condition). Are they simply guessing? That is not likely, because the group of Broca's aphasics performed rather well on the filler items, with plural head nouns (fewer than 8% agreement errors). Interestingly, there was no difference between filler items with number matching local nouns (7.1%) and those with number mismatching local nouns (7.4%). In other words, for items with plural head nouns there seems to be no effect of attraction. This finding, however, stands in good agreement with results obtained *inter alia* by Bock and Miller (1991) and Bock and Cutting (1992), who found no attraction effect for items with plural head nouns.

### Discussion

The present results can be summarized as follows: Both Broca's aphasics and normal controls are sensitive to the number marking of head and local noun in complex noun phrases. Both show an attraction effect with singular head nouns, that is, more errors when the local noun is plural than when it is singular. Both groups also show an effect of item type, that is, more errors with multiple token and multiple token control items than with single token and single token control items. However, there is a crucial difference: Normal controls show a significant interaction between the variables, so that most errors occur in the multiple token condition, the only condition in which conceptual number is plural, but grammatical number is singular. They take both syntactic and semantic information into account when constructing subject-verb agreement. Broca's aphasics on the other hand have additive effects of item type and number. In other words, Broca's aphasics are not sensitive to a mismatch between conceptual and grammatical number. They do not take semantic information into account when constructing subject-verb agreement. This confirms our prediction derived from the theory that Broca's aphasics suffer from a pathological reduction in resources which are applied both for syntactic and semantic processing.

However, there are a number of objections one can raise against this interpretation of the results. First, what remains unclear is whether Broca's aphasics have a limitation in their ability to *comprehend* distributed noun phrases. Obviously, if the aphasics did not understand the implied plurality of the grammatically singular head noun in multiple token items, there would be no conceptual plurality to drive agreement errors. This consideration will be dealt with in Experiment 3. Second, the data in Experiment 1 elicited a number of concerns that need to be dealt with. One of these concerns is the main effect of item type with Broca's aphasics (more agreement errors in the multiple token condition and multiple token control condition than in

the single token and single token control condition). What can explain this finding? Furthermore, why is there a (marginally significant) main effect of item type on the number of miscellaneous responses with Broca's aphasics (more miscellaneous responses in the single token than in the multiple token condition)? It is conceivable that these two effects are related to some idiosyncratic characteristics of the particular items used in the experiment, such as the plausibility of each item. We know from a previous study with (almost) the same materials (Vigliocco et al., 1996b, Experiment 1) that some of the conditions differed with respect to plausibility. Given these concerns and the relatively small number of observations we felt it important to replicate the experiment with a new set of materials, that were more balanced with respect to that variable, so that we can establish the robustness of our findings.

## EXPERIMENT 2

Experiment 2 served to replicate Experiment 1 with a larger group of participants and with different materials. The materials were the same as the ones used in Vigliocco et al. (1996b; Experiment 2), with a small number of improvements. Just as in Experiment 1, we presented a number of sentence fragments that had to be repeated and completed. Again, we varied item type (multiple token, single token) and number match of the head and local nouns (match, mismatch). Only in the multiple token, number mismatch condition did items have a conceptually plural reading.

### Method

#### *Participants*

Sixteen healthy normal controls, 7 males and 9 females, and 12 aphasic patients, 9 males and 3 females, participated in this experiment. Six of the aphasic patients (NU, FR, HA, LO, and TH) also participated in the previous experiment. However 1 patient (KL) was later excluded from the present experiment, because she appeared rather distracted during the session and did not produce any verb in the singular in the experimental items of the production experiment. We considered that a strategy.

The interval between the two experiments was more than 1 year. The aphasic patients were all diagnosed as Broca's aphasic on the Dutch version of the AAT. As can be seen from the AAT scores, supplied in Table 3, the group of Broca's aphasics has relatively spared comprehension and repetition skills. Demographic information on the group of Broca's aphasics is supplied in Table 2. Time postonset varied from 1 to 19 years, with an average of 9 years. The aphasic patients were on average 55 years old (range: 31–71). Normal controls were on average 57 years old (range 25–72). Normal controls were matched with respect to educational background to the group of Broca's aphasics.

#### *Materials*

Experimental materials were sentential fragments consisting of subject NPs followed by PPs. All experimental preambles had a singular head noun. The same items were used as in Experiment 2 of Vigliocco et al. (1996b), with some slight adaptations; in adapting the items, we avoided long, multisyllabic words and words with complex consonant clusters, because

agrammatic patients often have problems in repeating such words. The preferred semantic reading of the preambles (single token or multiple token) was evaluated by 11 independent judges, all normal adults and native speakers of Dutch. We included the items only if they were classified as single or multiple token by at least nine judges.<sup>8</sup> All of the items satisfied this restriction. There were 24 experimental items: 12 judged as single token and 12 judged as multiple token. The number of the local noun was varied, so that there corresponded a control item with a singular local noun to each experimental item. A complete listing of the items used is provided in Appendix B.

Each preamble consisted of the same number of words (five). The two sets of items (single token and multiple token) did not significantly differ from each other with respect to the average number of syllables [ $t(22) = .54$ ;  $p = .596$ ]. Experimental items all contained head and local nouns of the nonneuter gender.

There were 36 filler items, 24 of which had the same syntactic structure as the experimental items, but had a plural head noun. Half of these had a singular local noun, the other half had a plural local noun. The remaining 12 fillers were simple NPs, consisting of a determiner, one or more adjectives, and a noun. Simple NP fillers were singular in 6 cases and plural in the remaining 6 cases. In this way, the set of items a participant received (experimental items and filler items) was completely balanced for number of head noun and local noun. Filler items all had head nouns of the nonneuter gender.

We also coupled an adjective to each experimental and filler item. In the sentence types tested here, adjectives are not marked for number. Adjectives in the multiple and single token conditions did not differ with respect to number of syllables, number of characters, imageability, and age of acquisition (norms for the latter two variables were taken from Van Loon-Vervoom, 1985).

Four 60-item lists were created. Each list consisted of 24 experimental items (6 single token items, 6 single token control items, 6 multiple token items, and 6 multiple token control items) and 36 fillers. In each list the experimental and filler items were organized in a pseudo-random order, with the constraint that the list started with four fillers and that no more than two experimental items followed each other. Across the four lists, each item occurred twice in the number match condition and twice in the number mismatch condition.

### *Procedure*

The procedure was the same as in Experiment 1, with the exception that the preambles were read by another experimenter (W.H.).

### *Scoring*

Tape recordings were transcribed and assigned to the same scoring categories as in the previous experiment. Two raters scored transcriptions independently and then discussed any cases in which their scores were different until consensus was reached.

### *Design*

The design was identical to that of the previous experiment.

## **Results**

For the group of normal controls ( $N = 16$ ), application of our scoring criteria yielded 296 (77.1%) correct responses, 65 (16.9%) agreement errors,

<sup>8</sup> Using the normal approximation to the binomial distribution it was established that the critical number of judges for rejecting the null hypothesis of chance judgment, with  $p < .05$ , was 8.2.

TABLE 5  
Distribution of Responses: Experiment 2

	<i>n</i>	Correct	Agreement errors	Miscellaneous responses	Number of repetition errors	Agreement/repetition errors
Normal controls						
ST control	96	83	7	2	4	0
ST	96	68	17	7	4	0
MT control	96	85	8	1	2	0
MT	96	60	33	2	1	0
Broca's aphasics						
ST control	66	43	16	7	0	0
ST	66	30	27	7	1	1
MT control	66	42	14	7	2	1
MT	66	30	19	11	6	0

*Note.* The experimental conditions were ST control (single token, number match); ST (single token, number mismatch); MT control (multiple token, number match); MT (multiple token, number mismatch).

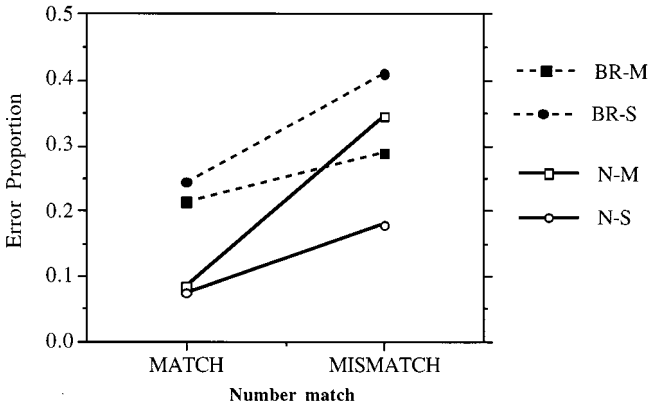
11 (2.9%) repetition errors, 0 repetition/agreement errors, and 12 (3.1%) miscellaneous responses. The Broca's aphasics ( $N = 11$ ) produced 145 (54.9%) correct responses, 76 (28.8%) agreement errors, 9 (3.4%) repetition errors, 2 (0.8%) repetition/agreement errors, and 32 (12.1%) miscellaneous responses. The distribution of responses is listed in Table 5.

Let us first consider the results obtained with the normal controls. As is clear from the table, agreement errors were most common (33) in the multiple token condition, whereas only 17 errors were found in the single token condition. In the multiple token control condition, 7 errors were found and in the single token control condition we obtained 8 errors. The data pattern for the Broca's aphasics is quite different. Agreement errors were *less* common in the multiple token condition (19) than in the single token condition (27). In the multiple token control condition 14 agreement errors were found and in the single token control condition we obtained 16 errors.

In Fig. 4, we plot the error proportions for both groups of participants.

The figure clearly shows an interaction between number match and sentence type for normal controls: This suggests an effect of distributivity. The Broca's aphasics show no distributivity effect. They made more agreement errors than the normal controls, but interestingly they made more errors in the single token conditions than in the multiple token conditions, a finding that reverses the direction of the effect in Experiment 1.

Because in this experiment, as opposed to the previous one, there were agreement errors in each cell of the design, we conducted analyses on the *number* of errors for each participant. We conducted a  $2 \times 2 \times 2$  repeated measures ANOVA with group as a between-subjects variable with two levels



**FIG. 4** Proportion of agreement errors in Experiment 2 (computed as a percentage of all responses). Interrupted lines: BR-M (Broca's aphasics, multiple token); BR-S (Broca's aphasics, single token); filled lines: N-M (normal controls, multiple token); N-S (normal controls, single token).

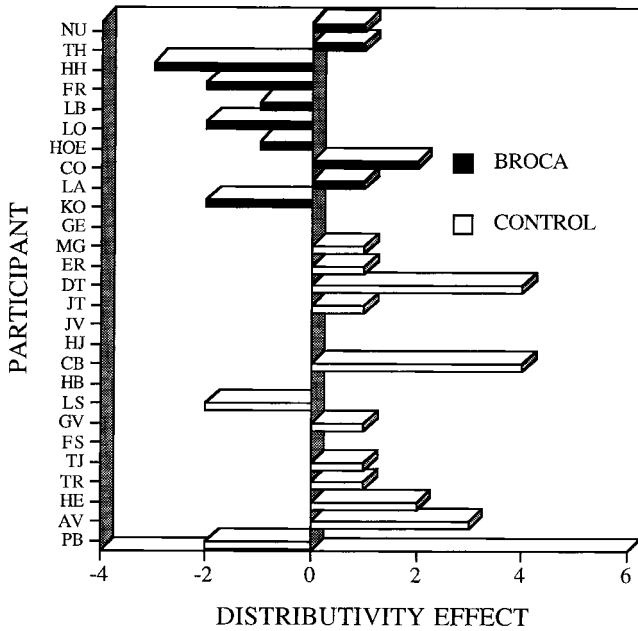
(normal controls, Broca's aphasic) and number match (match, mismatch) and item type (single token, multiple token) as within-subjects variables. This analysis showed a significant three-way interaction [ $F(1, 25) = 5.00$ ;  $MSE = .72$ ;  $p = .034$ ]. This significant three-way interaction implies that the interaction pattern between number match and item type is different for the two groups of participants.

We also conducted separate analyses for each participant group. These were  $2 \times 2$  repeated measures ANOVAs, with number match and item type as within-subjects variables. For normal controls, the effect of item type was close to significance [ $F(1, 15) = 3.93$ ;  $MSE = 1.15$ ;  $p = .066$ ]. The effect of number match was highly significant [ $F(1, 15) = 27.06$ ;  $MSE = .71$ ;  $p < .001$ ]. Crucially, the interaction between item type and number match was significant [ $F(1, 15) = 4.69$ ;  $MSE = .75$ ;  $p = .047$ ].

For the Broca's aphasics, there were actually more agreement errors in the single token and single token control condition (43) than in the multiple token and multiple token control conditions (33). This is the reverse of the pattern obtained in Experiment 1. However, the effect of item type was not significant [ $F(1, 10) = 2.76$ ;  $MSE = .82$ ;  $p = .127$ ]. Broca's aphasics did show an effect of number mismatch [ $F(1, 10) = 4.98$ ;  $MSE = 1.17$ ;  $p = .05$ ]. Finally, the interaction between item type and number match was not significant [ $F(1, 10) = 1.22$ ;  $MSE = .67$ ;  $p = .294$ ].

What about results at the individual level? Figures 5 and 6 show individual plots for the distributivity effect (Fig. 5) and for the attraction effect (Fig. 6), analogous to Figs. 2 and 3.

Figure 5 shows that a majority of normal controls have a positive effect of distributivity. Ten of 16 normal controls have a positive effect, 4 have a null effect, and the remaining 2 have a negative effect. The Broca's aphasics,



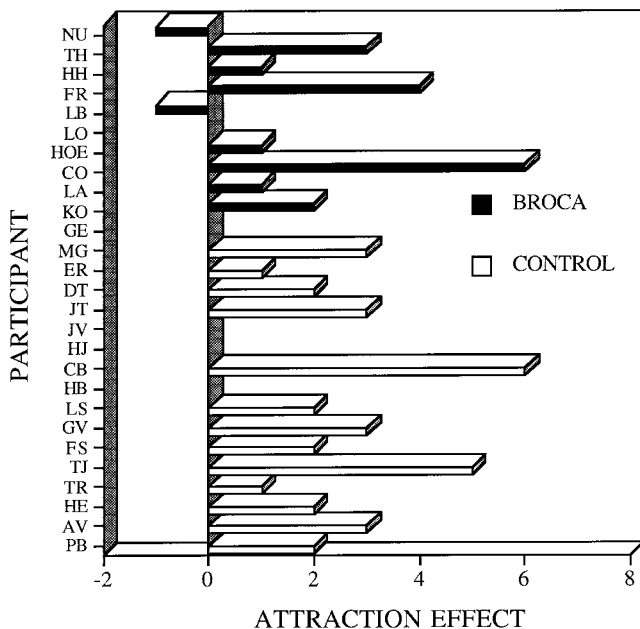
**FIG. 5.** Individual scores of distributivity effect in Experiment 2, in absolute numbers of errors: Multiple token effect (multiple token condition minus multiple token control condition) minus single token effect (single token condition minus single token control condition). The effect is zero where no bar is plotted.

on the other hand, show individual effects much more centered around zero: Four of them have a positive effect, 1 has a null effect, and the remaining 6 have a negative effect. It is interesting to notice that if we compare Fig. 5 with Fig. 2 and focus on the five aphasic patients that participated in both experiments, we see that four of them (NU, FR, HA, TH) have an effect in the second experiment in an opposite direction from that in the first experiment. This strengthens our conviction that any individual distributivity effects (positive or negative bars) obtained for the Broca's aphasics reflects sampling noise rather than an individual trait. By considering the data at the group level, the sampling noise is averaged out.

What about the effect of attraction? Figure 6 shows that the majority of Broca's aphasics (7/11) and of normal controls (13/16) show a positive attraction effect. Two Broca's aphasics actually produce more errors in the match conditions. The remaining participants (1 Broca's aphasic, 3 normal controls) show no attraction effect. These results are very similar to the results obtained in Experiment 1.

### Discussion

The second experiment replicated the first with respect to the most important finding: There was a significant effect of distributivity for normal con-



**FIG. 6.** Individual scores of attraction effect in Experiment 2, in absolute numbers of errors: Number of mismatch conditions (multiple token condition plus single token condition) minus number match condition (multiple token control condition plus single token control condition). The effect is zero where no bar is plotted.

trols but not for Broca's aphasics. This replication constitutes a clear confirmation of our hypothesis. Once again, the data do not provide support for the predictions that could be derived from the hypothesis of a resource limitation in a dedicated syntactic resource (a relatively strong attraction effect). Further, the data do not support the prediction that could be derived from the hypothesis of impairments to specific syntactic operations (a relatively strong distributivity effect).

There are a number of aspects in the present data that deserve some attention. First, the present experiment shows that the effect of item type on the number of miscellaneous responses with Broca's aphasics (more miscellaneous responses in the single token conditions), which was obtained in Experiment 1, is not robust. In fact, in the present experiment the numbers of miscellaneous responses were approximately equally distributed across the conditions.

Experiment 1 raised another concern: There was a main effect of item type on the number of agreement errors with Broca's aphasics. In fact, there were more errors in the multiple token condition and the multiple token control condition than in the corresponding single token conditions for the group of Broca's aphasics. Are multiple token items inherently more difficult? The

present experiment had a different outcome: The single token items elicited on average more errors than the multiple token items, but this difference was statistically not reliable. Thus, multiple token items are not inherently more difficult. In our opinion, the different pattern of results in the two experiments with respect to the direction of the main effect of item type in the Broca's aphasics should be attributed to idiosyncratic characteristics of the different item sets (notice, once again, that one cannot closely match the single token items to the multiple token items with respect to variables such as the particular lexical items. That is, in fact, the purpose of including control conditions).

A further concern that might be raised is that the effects of distributivity are so weak that they are not detectable in Broca's aphasics. However, both of our experiments had sufficient power to detect significant interactions between group and item type. In other words, our conclusions are not based on null results but on positive evidence. Furthermore, if we compare the size of the effect with that in several published papers testing the effect of distributivity in various languages (Vigliocco et al., 1996a,b), we notice that the effect of distributivity is of a similar magnitude. For instance, Vigliocco et al. (1996a) tested the effect of distributivity in Spanish. In Experiment 1, they found error percentages of 7% (single token) and 16% (multiple token). In an experiment which added a manipulation explicitly designed to increase the number of errors (presenting number mismatching adjectives) they found 13% agreement errors in the single token condition and 24% in the multiple token condition. In experiments in Dutch, conducted by Vigliocco et al. (1996b), which also had manipulations explicitly designed to increase error rate, there were error percentages of 3% (single token) and 27% (multiple token) and 3% (single token) and 15% (multiple token). Now compare the normal controls in the present experiment, which did *not* have a manipulation to increase error rate. In Experiment 1 the percentage of errors was 22% (ST) and 46% (MT) and in Experiment 2 the percentage was 18% (ST) and 34% (MT). Thus, the number of errors per condition is about twice as high in the MT condition as in the ST condition for normal controls. This is comparable to the Spanish data. Further, the increase in absolute terms (24 and 16%) is comparable to the increase in the Dutch data. Thus, the effect is of at least a similar size as that in comparable published experiments.

One final concern, the concern that Broca's aphasics might have problems in comprehending multiple token preambles, is addressed in the next experiment.

### EXPERIMENT 3: SENTENCE/PICTURE MATCHING TEST

The previous experiments showed that the normal controls have a strong effect of distributivity (more agreement errors when the conceptual number is plural and the grammatical number singular), but that the Broca's aphasics

do not. We interpret this as the consequence of a resource limitation. However, there is an important alternative interpretation with which we have to deal with before we accept that conclusion. Could it be the case that Broca's aphasics have problems in *comprehending* distributed noun phrases? Suppose Broca's aphasics do not understand that in a sentence such as *the label on the bottles*, the subject noun refers to multiple instances of *label*. Then, our manipulation should fail to induce a conceptually plural interpretation for the distributive items. Thus, we should not expect to find an effect of distributivity.

To our knowledge, agrammatics' comprehension of complex NPs such as those used in the present study has never before been the topic of systematic investigation. However, Kolk and Friederici (1985) tested comprehension of sentences containing prepositional phrases, such as (7a and 7b)

- (7a) De pop is op het bed.  
The doll is on the bed.  
(7b) Op het bed is de pop.  
On the bed is the doll.

Inspection of Kolk and Friederici's data (their Table V, p.55) reveals that in the conditions with nonreversible sentences such as (7a and 7b), very few errors were made by Dutch speaking Broca's aphasics, regardless of whether word order was canonical (7a) or not (7b). Because most of our experimental items are similar to the sentences in their nonreversible condition (for instance, the corresponding preamble in our experiment would be (7c)) we expect no comprehension difficulties with these items.

- (7c) De pop op het bed  
The doll on the bed

However, Kolk and Friederici (1985) tested only sentences with singular subject and prepositional object nouns. Suppose Broca's aphasics do not understand the meaning of number inflections? That is unlikely. Haarmann and Kolk (1994) showed in an online word monitoring task that agrammatics are sensitive to violations of subject-verb agreement. If they encountered a subject-verb agreement violation, and the next word was the target for which they had to monitor, their reaction times were delayed.

Thus, it seems we can rule out the possibility that Broca's aphasics have problems with the comprehension of prepositions and number inflections. However, it is still possible that comprehension of a multiple token phrase poses difficulties. Therefore, we tested comprehension of single token and multiple token sentence fragments in a sentence/picture matching test. In this test, we requested the participants to point to one out of four pictures that corresponded to an auditorily presented sentence fragment. We presented three types of items: multiple token, single token, and fillers. Sentence fragments were varied with respect to number of the head and local nouns

and pictures were varied with respect to the number of tokens of head and local noun referents.

## Method

### *Participants*

Participants were 16 normal control subjects and 11 Broca's aphasic patients. They were the same participants as in Experiment 2 (with the exception of aphasic participant KL, who was excluded from Experiment 2).

### *Materials*

The test contained 108 items that each consisted of a sentence fragment and a combination of four pictures. The four pictures were arranged in a square grid (see Fig. 7). There were 36 multiple token items, 36 single token items, and 36 filler items. The single and multiple token sentence fragments were complex noun phrases, consisting of a subject NP and a modifying PP. Fillers were simple NPs consisting of a determiner, one adjective, and a noun.

The 36 multiple token items were based on nine sentence fragments that occurred in four different conditions, as exemplified in Table 6. The four versions of each item were (1) singular head noun, plural local noun (SP); (2) plural head noun, plural local noun (PP); (3) singular head noun, singular indefinite local noun (SSindef); (4) singular head noun, singular definite local noun (SSdef). Notice that the SP sentence fragments are comparable to the multiple token sentence fragments in Experiments 1 and 2 and that the SSdef sentence fragments are comparable to the multiple token control fragments of those experiments.

An example of a multiple token *picture combination* is provided in Fig. 7. Each picture combination in the multiple token conditions consisted of four pictures, that showed (1) multiple tokens of a single type (MTST), e.g., two bottles, each having the same label; (2) multiple tokens of multiple types (MTMT), e.g., two bottles, each with a different label; (3) a single token (figure) and multiple "grounds" that carry the figure (SFMG), e.g., two bottles, but only one with a label; (4) a single token (figure) and one ground (SFSG), e.g., one label on one bottle.

*Single token* sentence fragments occurred in four conditions, as exemplified in Table 6. The conditions were (1) singular head noun, plural local noun (SP); (2) plural head noun, plural local noun (PP); (3) singular head noun, singular indefinite local noun (SSindef); (4) plural head noun, singular local noun (PS). The picture combinations consisted of four pictures in which the number of tokens of the referents of head and local noun was systematically varied. Thus, the four pictures showed (1) singular referent of head noun, plural referent of local noun (SP), e.g., one person with two cats; (2) singular referent of head and local noun (SS), e.g., one person with one cat; (3) plural referents of head and local nouns (PP), e.g., two persons with two cats; (4) plural referent of head noun, singular referent of local noun (PS), e.g., two persons with one cat.

The 36 filler sentences were simple NPs consisting of a determiner, one adjective, and a noun. Eighteen of these were singular and 18 were plural. Filler picture combinations occurred twice, once with a singular and once with a plural sentence fragment. Two of the pictures corresponded to the singular and plural versions of the sentence fragments and two were unrelated distractors. For instance, a filler item was *the laughing snowman*. The four pictures showed (1) one snowman; (2) two snowmen; (3) one pencil; (4) two pencils.

The picture combinations and sentence fragments were designed so that, conceivably, there is one sentence matching with each picture. However, it is very important to note that in the case of a multiple token item there is no unique "correct" response. For the most critical sentence fragments, the ones with singular head nouns and plural local nouns (SP) both the

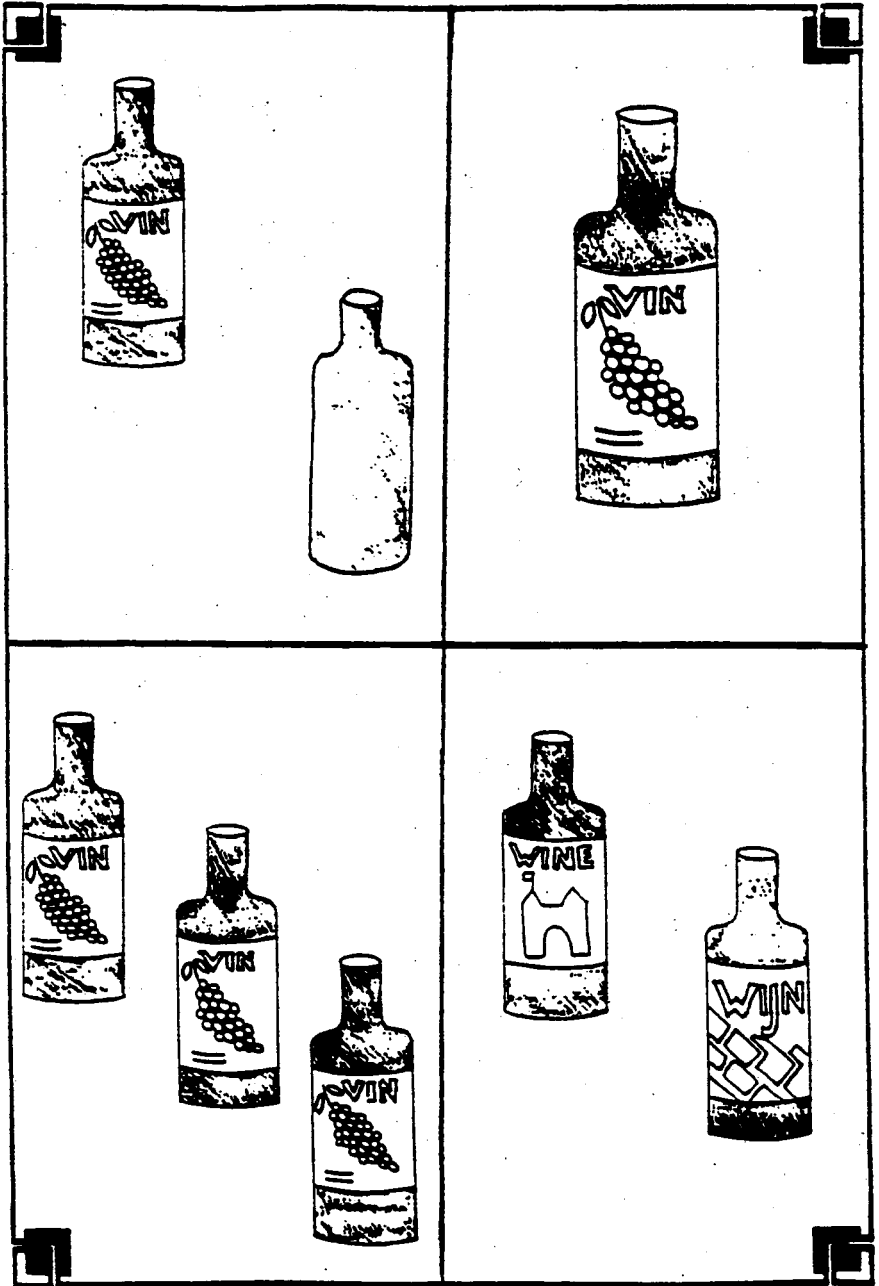


FIG. 7. Example of experimental picture combination in Experiment 3. Such a picture combination was presented four times in total, once with each of the following sentence fragments (translated to English): (1) the label on the bottles; (2) the labels on the bottles; (3) the label on the bottle; (4) the label on a bottle.

TABLE 6  
 Example Set of Experimental Sentence Fragments  
 in Experiment 3

Sentence type	Example
Multiple token	
SP	Het etiket op de flessen (The label on the bottles)
PP	De etiketten op de flessen (The labels on the bottles)
SSind	Het etiket op een fles (The label on a bottle)
SSdef	Het etiket op de fles (The label on the bottle)
Single token	
SP	Het baasje van de katten (The owner of the cats)
PP	De baasjes van de katten (The owners of the cats)
SSind	Het baasje van een kat (The owner of a cat)
PS	De baasjes van de kat (The owners of the cat)

MTST condition and the MTMT condition may be considered valid interpretations.<sup>9</sup> The main goal of the test, at least with respect to the multiple token items, is to establish whether the two groups of participants have similar interpretations of single token and multiple token items.

The experiment had two important features: First, it tested the preference for the *kind* of multiple token interpretation that is built up when hearing a multiple token fragment. In fact, does our mental representation contain multiple tokens of the same type or multiple tokens of a different type? Second, this test took into account the possibility that the participants, in particular the aphasic participants, comprehended the number of the head noun and the local noun, but were unable to integrate those elements into a satisfactory interpretation. In that case they might have been tempted to match a sentence in the SP condition ('the label on the bottles') with a SFMG picture (e.g., Fig. 7a).<sup>10</sup>

The sentence fragments with singular head nouns and plural local nouns were judged by 11 raters with respect to distributivity. The fragments were classified as single token or multiple token if at least 9 of 11 judges agreed on this judgment. Nine items were thus classified as single token and nine as multiple token. None of these fragments was used in Experiments 1 or 2. The single and multiple token preambles are listed in Appendix C.

Each picture combination was presented four times, one with each of the four sentence fragments. The four presentations were presented in a random order. The position of the different picture types in each picture combination was also chosen at random.

<sup>9</sup> This was established in a pretest, reported in Huinck (1998). In that pretest, participants were requested to judge whether a sentence fragment was a correct description of a target picture. It turned out that a combination of a SP sentence fragment and a MTMT picture was judged 'correct' in 41% of the cases.

<sup>10</sup> Notice, however, that there is still a perfectly grammatical way of denoting that situation with a multiple token sentence: 'the label on the bottles was not always present.'

## Procedure

The sentence/picture matching task was administered after the participants had completed the production experiment (Experiment 2) with a short break between the experiments. The experimenter read each sentence aloud, and the participant was requested to point to the corresponding picture out of four alternatives. Each combination of four alternatives was presented on a separate sheet. The participants were instructed that sometimes more than one picture could be correct, but that only one alternative should be chosen, the one that fit the sentence best. If a participant first pointed to one picture and then to another, the final decision counted. Responses were circled by the experimenter on a score form.

## Design and Analysis

We presented nine items in each sentence condition for each item type. Separate analyses were conducted for each item type: multiple token, single token, and filler items. We calculated a measure of agreement between the majority judgments of each group, Cohen's  $\kappa$  (Rietveld & Van Hout, 1993). The  $\kappa$  statistic varies between 0 (no correspondence at all) and 1 (perfect agreement). This statistic is based on whether the majority of each group chose a picture that was "consistent" or "inconsistent" for each sentence fragment. It is important to note that the scoring categories were not the same for each of the item types. In fact, for each *multiple token* sentence fragment, two of the pictures can be considered consistent interpretations. Thus, we subdivided the responses for these items in the following manner. First, for sentence fragments with a plural local noun (SP and PP), both pictures with multiple tokens of one type (MTST), and pictures with multiple tokens of multiple types (MTMT) were considered "consistent" responses. Second, for sentence fragments with a singular local noun (SSdef and SSindef), the choice for each of the other two pictures (SFMG, SFSG) was considered "consistent" (see Fig. 7).

For each type of *single* token sentence fragment, only one picture of four was considered "consistent": the one that showed the correct number of exemplars of the referents of the two nouns. For the filler items, also only one picture of four was considered consistent. Thus, chance probability of choosing a consistent response was 0.5 for each item in the multiple token condition and 0.25 in the single token and in the filler conditions.

## Results

In Table 7, we list the percentages with which each multiple token sentence type was matched to a particular picture type, for both groups of participants.

As the table shows, the overall pattern is very similar for the two groups. For each sentence fragment condition, there were clear preferences for certain picture types, and the two groups have the same preferences. The most relevant condition for the purpose of determining comprehension of distributive phrases is the condition with singular head nouns and plural local nouns (i.e., these are the only "real" multiple token items). We have plotted the percentages responses for those items in Fig. 8

For multiple token items, both groups of participants most often chose a picture that showed multiple tokens of a single kind (e.g., the same label on all bottles). Somewhat surprisingly, the second most popular choice was that for the picture that depicted multiple backgrounds, but only one with a figure (e.g., two bottles, but only one with a label). This choice was made in 20.8%

TABLE 7  
 Proportions of Responses for Each Picture Type, in All Four Multiple  
 Token Sentence Conditions

Condition	Type of response (%)			
	MTST	MTMT	SFMG	SFSG
Normal controls				
SP	61.8	13.2	20.8	4.2
PP	49.3	47.9	1.4	1.4
SSind	0.7	0.0	9.0	90.3
SSdef	0.7	0.0	1.4	97.9
Broca's aphasics				
SP	47.5	16.2	25.3	11.1
PP	61.6	34.3	3.0	1.0
SSindef	1.0	4.0	6.1	88.9
SSdef	0.0	3.0	5.1	91.9

*Note.* The conditions were SP (singular head noun, plural local noun); PP (plural head noun, plural local noun); SSindef (singular head noun, singular local noun, with indefinite determiner); SSdef (singular head noun, singular local noun, with definite determiner). Response types were MTST (multiple token, single type); MTMT (multiple token, multiple type); SFMG (single figure, multiple grounds); SFSG (single figure, single ground).

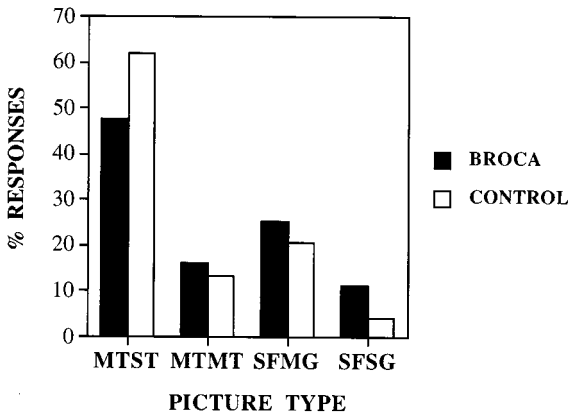


FIG. 8. Results for the multiple token, mismatch sentence condition in Experiment 3 (e.g., “the label on the bottles”), expressed as proportions picture selections in each group. Picture types were: (MTST), e.g., a number of bottles, each with a label of the same type; (MTMT) e.g., a number of bottles, each with a label of a different type; (SFMG), e.g., a number of bottles, only one with a label; (SFSG) e.g., one bottle with one label.

of the cases by normal controls and in 25.3% of the cases by Broca's aphasics. That choice definitely constitutes a different interpretation than we intended. However, this choice is made about equally often in each group. Both groups also sometimes chose a picture showing multiple tokens of multiple kinds (e.g., a number of bottles, with different labels). Again, the groups showed similar percentages of choosing this option. Finally, in fewer than 10% of the cases, a picture was selected showing only one element (e.g., one bottle, with one label). This clear misinterpretation of the sentence fragment sometimes occurred in both groups of participants. In sum, distributive preambles were most often matched with a picture showing multiple tokens, and this held for normals (in 75%) as well as for Broca's aphasics (in 63.7%). It does seem that Broca's aphasics have a somewhat more noisy data pattern.

What kinds of pictures were selected in the other versions of the multiple token items? Table 7 lists the percentage of responses in each of the categories for the multiple token items. Clearly, in the conditions labeled PP, SSdef, and SSindef, the two groups of participant show very similar behavior. First, the distinction definite/indefinite seems to be relatively unimportant for both groups. In the case of SSdef, both groups almost always chose the picture with one figure on one background. In the case of SSindef, both groups also have a clear preference for the picture with a single figure on a single ground, but sometimes chose the picture with a single figure and multiple grounds. In the case of sentences with plural head and local nouns, both groups of participants tended to prefer multiple token pictures. However, whereas the normal controls had no clear preference for each of the two multiple token pictures, Broca's aphasics more often preferred multiple tokens of a single type.

The overall measure of agreement,  $\kappa$ , was 0.88 for the multiple token items, an excellent agreement between the (majority) judgments of the two groups. There were only two items in which the two groups differed: The majority of Broca's aphasics rated *the stamp on the envelopes* and *the candle in the candlesticks* as single token, whereas the normal controls rated them as multiple token.

The ratings of single token items also showed a strong overlap between the two groups of raters, as shown in Table 8. The Broca's aphasics did rather poorly with the PS items (only 62.6% of the responses were "consistent"). However, performance was much better for the SP and SSdef items, which are similar to the items tested in Experiments 1 and 2 (SP, 82.8%; SSdef, 91.9% correct). The  $\kappa$  statistic of agreement between the majority judgments was 0.89 for the single token items. On the filler items, normal controls chose the picture showing a singular object in 99.3% of singular sentences and chose the picture showing multiple objects in 99.3% of the plural sentences. These figures were 97.5 and 98.5%, respectively, for the Broca's aphasics.  $\kappa$  was 1 for the filler items.

TABLE 8  
Proportions of Responses for Each Picture Type, in All Four Single  
Token Sentence Conditions

Condition	Type of response (%)			
	SP	PP	SS	PS
Normal controls				
SP	100	0	0	0
PP	1.4	98.6	0	0
SSind	0	0	95.8	4.2
PS	1.4	4.2	0.7	93.8
Broca's aphasics				
SP	82.8	9.1	2.0	6.1
PP	4.0	91.9	0	4.0
SSind	7.1	0	89.9	3.0
PS	10.1	16.2	11.1	62.6

*Note.* Conditions were SP (singular head noun, plural local noun); PP (plural head noun, plural local noun); SSind (singular head noun, singular local noun, with indefinite determiner); SSdef (singular head noun, singular local noun, with definite determiner). Response types were SP (head noun singular referent, local noun plural referent); PP (head and local noun plural referent); SS (head and local noun singular referent); PS (head noun plural referent, local noun singular referent).

## Discussion

The sentence-picture matching task tested a number of aspects of comprehension of sentence fragments consisting of a subject NP and a modifying PP. First of all, it is clear that the Broca's aphasics have no problems in distinguishing between singular and plural, as witnessed by the near perfect behavior on the filler items (containing one noun inflected for number). Furthermore, the excellent performance on the ST and MT items shows that when the sentence fragments are made slightly more complex (i.e., containing two nouns, both inflected for number), Broca's aphasics can distinguish which number goes with which noun. Although Broca's aphasics have a somewhat more noisy pattern of results than the control subjects, they still choose the "consistent" alternative in the majority of cases. The measure of agreement, Cohen's  $\kappa$ , indicates good agreement between the majority judgments of both groups.

What about the subtle semantics of multiple token items? The comprehension profile for (number mismatching) multiple token sentences was very similar for normals and Broca's aphasics, with the Broca's aphasics being slightly more "noisy." However, *all* participants, except one control participant, sometimes chose an inconsistent picture. It is especially interesting to see that quite often the picture with a single figure on multiple grounds is

chosen. This suggests that multiple token items are sometimes not understood, but this happened for normal controls about as often as for Broca's aphasics.

In sum, the Broca's aphasics show the same basic pattern of comprehension for distributive sentences as normal controls. In fact, both groups have a preference for a multiple token interpretation, especially for multiple tokens of a single type. Although the Broca's aphasics sometimes chose the distractor with one token and multiple backgrounds, the normal controls often did so too. Furthermore, the two groups of participants show similar comprehension profiles for single token items (with the exception of the PS items). From this experiment we conclude that Broca's aphasics have a largely spared ability to comprehend complex noun phrases with two nouns that may or may not differ in number and in which the relationship between head noun and local noun may be distributive or not.

### GENERAL DISCUSSION

Our experiments can be summarized as follows. In Experiment 1, we showed that healthy adults, when presented with a sentence fragment with a singular head noun and a plural local noun, make more agreement errors upon completing that fragment when the fragment has a distributive reading. This effect of distributivity confirms earlier studies in Dutch with university students (Vigliocco 1996b). Broca's aphasics on the other hand showed no effect of distributivity. Although it is true that the aphasic group made more errors in the condition with a distributive reading (the multiple token condition) than in the single token condition, a similar difference was observed in the corresponding control conditions. In Experiment 2, the main finding from Experiment 1 was replicated with a larger group of Broca's aphasics. Again, the normal controls showed an effect of distributivity, but the Broca's aphasics did not. Finally, in Experiment 3, we tested the auditory comprehension of complex noun phrases, such as the ones used in the first two experiments, in a picture/sentence matching task. This experiment clearly showed that (1) Broca's aphasics are sensitive to morphophonological marking of number; (2) Broca's aphasics are still sensitive to number marking when there are two nouns, a head noun and a local noun, that may carry different numbers; (3) Broca's aphasics and normal controls have similar comprehension profiles, both for single token and for multiple token items. In fact, given a multiple token preamble with singular head noun and plural local noun, both groups most often prefer a picture showing multiple tokens of the referent of the head noun. This comprehension test importantly contributes to a validation of our sentence production results, because it provides evidence against the possibility that the difference between Broca's aphasics and normal controls on the production task results from a difference in comprehension.

We will now discuss a number of concerns about the interpretation of the results of the production experiments. Then in the next section we discuss the implications of our findings. To begin with, in Experiment 1, there was a relatively large number of miscellaneous responses in the single token conditions for the Broca's aphasics. Could it be that because of this unbalance, a possible effect of distributivity is disguised? We do not think so, because in Experiment 2 that unbalance was no longer a problem. In fact, the miscellaneous responses occurred about equally often in all the experimental conditions of Experiment 2, but there was still not a trace of a distributivity effect with Broca's aphasics.

Another concern was that for the Broca's aphasics we obtained a main effect of item type in Experiment 1. The Broca's aphasics made more errors in the multiple token and multiple token control conditions than in the corresponding single token conditions. Are the multiple token conditions more difficult? That is not the case, because there was no effect of item type in Experiment 2, in which we used different materials. In fact, the means were in the opposite direction in that experiment (more errors in the single token conditions). In our opinion, the significant main effect in Experiment 1 should be attributed to idiosyncrasies in the sentence material. As mentioned before, it is very difficult to match items in the single token condition to items in the multiple token condition on every potentially relevant variable. Out of necessity, these are different items.

It should further be noted that the interpretation of Experiment 1 might be troubled by another concern. This concern follows from the fact that normal controls made almost no agreement errors in the number match conditions (zero in the single token match condition and one in the multiple token match condition). That might make one suspicious that the distributivity effect with normal controls is really the result of a floor effect; because of that floor effect an interaction between item type and number match is obtained instead of a main effect of item type. However, that concern can be discarded on the basis of Experiment 2. In that latter experiment, the normal controls did produce agreement errors in both of the number match conditions but there was no difference between these two conditions. Further, as mentioned before, there was no longer a main effect of sentence type for the Broca's aphasics in Experiment 2.

There is one final hesitation about Experiments 1 and 2 that needs to be discussed. This hesitation entails the fact that Broca's aphasics might not have used number information (grammatical or conceptual) but rather supplied infinitives. In Dutch, the third person plural, present tense verb is morphophonologically indistinguishable from the infinitive. That would explain the low error rate in the filler items with plural head nouns and the high error rate in most of the experimental conditions. However, it would not explain the low error rate in the single token control condition, where re-

sponding with an infinitive would lead to an agreement error. In addition, there is another reason that makes this an implausible suggestion: In constructions with the infinitive form in Dutch, one tends to put the verb clause-finally. Kolk and Heeschen (1992) showed that Broca's aphasics placed almost all nonfinite verbs clause-finally and almost no finite verbs clause-finally. Therefore (8) is a plausible agrammatical utterance.

- (8) De man boos zijn  
The man angry be

If Broca's aphasics were really using infinitive forms, they would have tended to reverse the order of adjective and verb. In fact, this never happened.

In sum, two experiments show that in production both Broca's aphasics and normal controls are sensitive to the number marking of head and local noun in complex noun phrases. Both show an attraction effect with singular head nouns, that is, more errors when the local noun is plural than when it is singular. However, there is a crucial difference: Normal controls show an effect of distributivity, but Broca's aphasics do not. Thus, it can be argued that normal controls take both syntactic and semantic information into account when constructing subject-verb agreement. Broca's aphasics on the other hand are not sensitive to a mismatch between conceptual and grammatical number. They do not take semantic information into account when constructing subject-verb agreement. This confirms our prediction, derived from the theory that Broca's aphasics suffer from a pathological reduced capacity of resources which are applied both for syntactic and semantic processing. The data argue against the predictions derived from the hypotheses of a specific syntactic resource limitation or of an impairment in certain syntactic operations.

### Implications

What are the implications of the present findings? To begin with, we have shown once again that the mechanism that constructs subject-verb agreement is not blind to the referential properties of the subject head. The clear distributivity effect obtained with the healthy adults constitutes a replication of findings reported in Vigliocco et al. (1996b). Further, we have shown that such a distributivity effect does not occur in the group of Broca's aphasics, in support of the hypothesis that taking into account conceptual number is too resource-consuming for the latter group.

The lack of a distributivity effect with Broca's aphasics cannot be attributed to comprehension problems with distributed phrases. Experiment 3 showed that normal controls and Broca's aphasics tend to point to the same pictures when they have to choose the picture best matching with a multiple token sentence fragment. The data further show that the most common interpretation for a multiple token sentence is an interpretation of multiple tokens

of the same kind. In itself, that is an interesting finding, because one might argue that in that sort of interpretation, there may be conceptual singularity. That would be the case if the head noun were considered to refer to the single *type* rather than to the multiple tokens. Given the effect of distributivity obtained in Experiments 1 and 2, the results in comprehension imply that the number of tokens in the discourse model interferes with agreement, not the number of types.

Our findings also have repercussions for the debate on the specificity of computational resources. Our data are compatible with the notion of more general verbal working memory resources (cf. Just & Carpenter, 1992; Miyake, Carpenter, & Just, 1994, 1995), used at least for semantic and for syntactic processing. A resource limitation leads to the inability of conceptual information (plurality of head noun referent) to influence the setting of a syntactic feature. The group of Broca's aphasics, having (*ex hypothesi*) fewer resources, drop conceptual number. Metaphorically, whereas normal controls use all the number information they can, Broca's aphasics use only the number information they need. Why do the Broca's aphasics use grammatical information and not conceptual information? As mentioned earlier, in Dutch, grammatical number *always* specifies correct agreement (as opposed to English, which sometimes allows a plural verb with collective nouns). Thus, syntactic information really provides the most valid cue, and thus it is that syntactic information is used at the expense of semantic information, in the face of a resource limitation. Notice that the assumption that the most valid cue (syntactic information) is used also holds for the reading experiment in which semantic information could help resolve a local syntactic ambiguity (Just & Carpenter, 1992). Low-span readers did not use the semantic information (animacy of the subject), which can be considered a less valid cue than phrase structure for the interpretation.

As mentioned before, the theory of a specific impairment of syntactic resources or of a deficit in specifically syntactic operations (Caplan & Waters, 1995; Martin, 1995; Martin & Romani, 1994) would make different predictions in the current experiment. A severe shortage of syntactic resources might lead to a too-fast decay of the subject noun number. It is known that local nouns with a different number than the subject noun exert an attraction effect. Thus, if the subject noun number is represented weakly, as might be hypothesized in agrammatic aphasia, there would be stronger interference from the local noun, which is produced more recently than the subject noun. That would result in an stronger than normal number attraction effect. An impairment in specific syntactic operations might lead to an overreliance on semantic information, as agrammatics show in comprehension (e.g., Caramazza & Zurif, 1976). That would lead to an extremely strong effect of distributivity. Neither of these data patterns was obtained with Broca's aphasics. Instead, as predicted by the theory of a reduction in specific resources,

subserving syntactic as well as semantic processes, we found an attraction effect within the normal range and no distributivity effect for Broca's aphasics, but a clear distributivity effect for normal controls.

## APPENDIX A

### Experimental Items and, in Capitals, Adjectives to Be Used in the Completion in Experiment 1

#### Single token

BRUTAAL	De aanslag op de minister/s
(BOLD	the strike ON the minister/s)
KNAP	De arts van de zieke/n
(CLEVER	the doctor of the sick person/s)
LIEF	De baby op de foto/'s
(CUTE	the baby ON the photo/s)
GROOT	De beker voor de winnaar/s
(LARGE	the cup for the winner/s)
KORT	De boodschap voor de reiziger/s
(SHORT	the message for the traveler/s)
SUCCESSVOL	De diefstal van de diamant/en
(SUCCESSFUL	the theft of the diamond/s)
SAAI	De discussie over de wet/ten)
(BORING	the discussion about the law/s)
SPOORLOOS	De eigenaar van de koffer/s
(GONE	the owner of the suitcase/s)
ONREDELIJK	De eis van de ontvoerder/s
(UNREASONABLE	the demand of the kidnapper/s)
TERECHT	De klacht van de scholier/en
(VALID	the complaint of the student/s)
HANDIG	De monteur van de motorfiets/en
(HANDY	the mechanic of the motorcycle/s)
BEKEND	De regisseur van de film/s
(WELL-KNOWN	the director of the movie/s)

#### Multiple token

MOOI	De afbeelding op de mok/ken
(PRETTY	the picture ON the mug/s)
LUID	De bel op de fiets/en
(LOUD	the bell ON the bicycle/s)
ONDUIDELIJK	De beschrijving in de gids/en
(UNCLEAR	the description IN the guide/s)
ONGELDIG	De bon in de folder/s
(INVALID	the coupon IN the flyer/s)
ONLEESBAAR	De datum op de munt/en
(ILLEGIBLE	the date ON the coin/s)
SMAL	De ingang van de flat/s
(NARROW	the entrance OF the apartment/s)
VUIL	De kraag van de jas/sen
(DIRTY	the collar OF the coat/s)
GAMMEL	De leuning van de stoel/en
(UNSOUND	the back OF the chair/s)
VALS	De paraaf op de declaratie/s

(FAKED	the initials ( <i>singular</i> ) on the declaration/s)
OPVALLEND	De reclame op de bus/sen
(STRIKING	the advertisement on the bus/ses)
KLEIN	De sleutel van de kast/en
(SMALL	the key of the cupboard/s)
GROEN	De stop op de fles/sen
(GREEN	the stopper on the bottle/s)

## APPENDIX B

### Experimental items and, in Capitals, Adjectives to Be Used in the Completion in Experiment 2

#### Single token

*HARD	De actie tegen de wet/ten
(HARSH	the protest against the law/s)
BRUTAAL	De diefstal van de diamant/en
(BOLD	the theft of the diamond/s)
SAAI	De discussie over de startbaan/banen
(BORING	the discussion about the runway/s)
GROOT	De doos met de moorkop/pen
LARGE	the box with the moorkop/pen [Dutch pastry]
(SCHOKKEND	De intocht van de tank/s)
(SHOCKING	the entry of the tank/s)
OUD	De kerk bij de heuvel
(OLD	the church near the hill/s)
(STEVIG	De kooi met de gorilla/'s)
(SOLID	the cage with the gorilla/s)
(HARD	De muur met de decoratie/s)
(HARD	the wall with the decoration/s)
GROOT	De vaas met de zonnebloem/en
(BIG	the vase with sunflower/s)
*LIEF	De baby op de deken/s)
(SWEET	The baby on the blanket/s)
*LEUK	De reis naar de stad/steden
(NICE	The trip to the city/cities)
*NIEUW	De bus voor de toerist/en
(NEW	The bus for the tourist/s)

#### Multiple token

MOOI	De afbeelding op de mok/ken
(PRETTY	the picture on the mug/s)
(KLEIN	De chip in de computer/s)
(SMALL	the chip in the computer/s)
(*ONJUIST	De datum op de cheque/s)
(INCORRECT	the date on the cheque/s)
*ONLEESBAAR	De naam in de ring/en
(ILLEGIBLE	the name in the ring/s)
(BEKEND	De kop op de gulden/s)
(WELL-KNOWN	the head on the guilder/s [Dutch coin])
*VALS	De code op de bankpas/sen
(COUNTERFEIT	the code on the bankcard/s)
(*BELANGRIJK	De medaille voor de hardloper/s)

IMPORTANT	the medal for the runner/s
KORT	De plaatsnaam op de wegwijzer/s
SHORT	the city name on the roadsign/s
MOEILIJK	De puzzel in de krant/en
DIFFICULT	The puzzle in the newspaper/s
OPVALLEND	De reclame op de bus/sen
STRIKING	the advertisement on the bus/ses
*LANG	De titel op de CD/s
LONG	The title on the CD/s
*ZWAAR	De straf voor de boef/boeven
HEAVY	The punishment for the scoundrel/s

\*Deviations from the Vigliocco (1996b; Experiment 2) preambles.

## APPENDIX C

### Experimental Items to Be Used in the Sentence/Picture Matching Task in Experiment 3

#### Single token

Het baasje van de kat/ten	The owner of the cat/s
De dief achter de tralie/s	The thief behind the bar/s
De kast met het boek/de boeken	The cupboard with the book/s
De boerderij met het varken/de varkens	The farm with the pig/s
De kandelaar met de kaars/en	The candlestick with the candle/s
Het kasteel met de toren/s	The castle with the tower/s
De vlinder bij de bloem/en	The butterfly near the flower/s
De tak met het blaadje/de blaadjes	The twig with the leaf/s
De weg naar de berg/en	The road to the mountain/s

#### Multiple token

Het etiket op de fles/sen	The label on the bottle/s
De stok van de dirigent/en	The stick of the conductor/s
Het gezicht op de munt/en	The face on the coin/s
De bloem in de bloempot/ten	The flower in the flowerpot/s
De windrichting op het/de kompas/sen	The wind-direction on the compas/ses
Het blaadje aan de tak/ken	The leaf on the twig/s
De postzegel op de envelop/pen	The stamp on the envelope/s
De kaars in de kandelaar/s	The candle in the candlestick/s
De kaart op de prikborden	The postcard on the board/s

*Note:* Only the conditions with singular head noun and singular definite or plural local noun are listed.

## REFERENCES

- Bates, E., & MacWhinney, B. 1982. Functionalist approaches to grammar. In E. Wanner & L. R. Gleitman (Eds.), *Language acquisition: The state of the art*. Cambridge: Cambridge Univ. Press.
- Berndt, R. S., & Caramazza, A. 1980. A redefinition of the syndrome of Broca's aphasia: Implications for a neuropsychological model of language. *Applied Psycholinguistics*, **1**, 225–278.
- Bock, J. K. 1995. Producing agreement. *Current Directions in Psychological Science*, **4**, 56–61.

- Bock, J. K., & Cutting, J. C. 1992. Regulating mental energy: Performance units in language production. *Journal of Memory and Language*, **31**, 99–127.
- Bock, J. K., Eberhard, K. M., & Cutting, J. C. 1992. *Controlling number agreement on verbs and anaphors*. Paper presented at the Psychonomic Society, St. Louis.
- Bock, J. K., & Miller, C. A. 1991. Broken agreement. *Cognitive Psychology*, **23**, 45–93.
- Butterworth, B., Panzeri, M., Semenza, C., & Ferreri, T. 1990. Paragrammatisms: A longitudinal study of an Italian patient. *Language and Cognitive Processes*, **5**, 115–140.
- Caplan, D., Baker, C. & Dehaut, F. 1985. Syntactic determinants of sentence comprehension in aphasia. *Cognition*, **21**, 117–175.
- Caplan, D., & Hildebrandt, N. 1988. *Disorders of syntactic comprehension*. Cambridge, MA: MIT Press.
- Caplan, D., & Waters, G. S. 1995. Aphasic disorders of syntactic comprehension and working memory capacity. *Cognitive Neuropsychology*, **12**, 637–649.
- Carpenter, P. A., Miyake, A., & Just, M. A. 1994. Working memory constraints in comprehension: Evidence from individual differences, aphasia, and aging. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics*. San Diego: Academic Press. Pp. 1075–1122.
- Caramazza, A., & Zurif, E. B. 1976. Dissociation of algorithmic and heuristic processes in language comprehension: Evidence from aphasia. *Brain and Language*, **3**, 572–582.
- 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. S. 1986. A spreading-activation theory of retrieval in sentence production. *Psychological Review*, **93**, 283–321.
- Ferreira, F., & Clifton, C. 1986. The interdependence of syntactic processing. *Journal of Memory and Language*, **25**, 348–368.
- Frazier, L., & Friederici, A. D. 1991. On deriving properties of agrammatic comprehension. *Brain and Language*, **40**, 51–66.
- Grodzinsky, Y. 1986. Language deficits and the theory of syntax. *Brain and Language*, **27**, 135–159.
- Haarmann, H. J., & Kolk, H. H. J. 1994. On-line sensitivity to subject-verb agreement violations in Broca's aphasics: the role of syntactic complexity and time. *Brain and Language*, **46**, 493–516.
- Haarmann, H. J., Just, M. A., & Carpenter, P. A. 1997. Aphasic sentence comprehension as a resource deficit: A computational approach. *Brain and Language*, **59**, 76–120.
- Hagiwara, H. 1995. The breakdown of functional categories and the economy of derivation. *Brain and Language*, **50**, 92–116.
- Hartsuiker, R. J., & Kolk, H. H. J. 1998. Syntactic facilitation in agrammatic sentence production. *Brain and Language*, **62**, 221–254.
- Huinck, W. J. 1998. *One or more labels on the bottles? Production and comprehension of distributed phrases in Broca's aphasia*. Unpublished masters thesis, NICI, Univ. of Nijmegen.
- Just, M. A., & Carpenter, P. A. 1992. A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, **99**, 122–149.
- Kolk, H. H. J. 1995. A time-based approach to agrammatic production. *Brain and Language*, **50**, 282–303.
- Kolk, H. H. J., & Friederici, A. D. 1985. Strategy and impairment in sentence understanding by Broca's and Wernicke's aphasics. *Cortex*, **21**, 47–67.
- Kolk, H. H. J., & Heeschen, C. 1992. Agrammatism, paragrammatism, and the management of language. *Language and Cognitive Processes*, **7**, 89–129.

- Kolk, H. H. J., & Van Grunsven, M. J. F. 1985. Agrammatism as a variable phenomenon. *Cognitive Neuropsychology*, **2**, 347–384.
- Kolk, H. H. J., & Weijts, M. 1996. Judgments of semantic anomaly in agrammatic patients: argument movement, syntactic complexity, and the use of heuristics. *Brain and Language*, **54**, 86–135.
- Levelt, W. J. M. 1989. *Speaking: from intention to articulation*. Cambridge, MA: MIT Press.
- Linebarger, M. C., Schwartz, M. F., & Saffran, E. M. 1983. Sensitivity to grammatical structure in so-called agrammatic aphasics. *Cognition*, **13**, 361–392.
- Martin, R. 1995. Working memory doesn't work: A critique of Miyake et al.'s capacity theory of aphasic comprehension deficits. *Cognitive Neuropsychology*, **12**, 623–636.
- Martin, R., & Romani, C. 1994. Verbal working memory and sentence comprehension: A multiple components view. *Neuropsychology*, **8**, 506–523.
- Martin, R., Shelton, J. R., & Yaffee, L. S. 1994. Language processing and working memory: Neuropsychological evidence for separate phonological and semantic capacities. *Journal of Memory and Language*, **33**, 83–111.
- Miyake, A., Carpenter, P. A., & Just, M. A. 1994. A capacity approach to syntactic comprehension disorders: Making normal adults perform like aphasic patients. *Cognitive Neuropsychology*, **11**, 671–717.
- Miyake, A., Carpenter, P. A., & Just, M. A. 1995. Reduced resources and specific impairments in normal and aphasic sentence comprehension. *Cognitive Neuropsychology*, **12**, 651–679.
- Pollard, C., & Sag, I. A. 1994. *Head-driven phrase structure grammar*. Chicago: Univ. of Chicago Press.
- Rietveld, T., & Van Hout, R. 1993. *Statistical techniques for the study of language and language behaviour*. Berlin: de Gruyter.
- Schwartz, M., Linebarger, M. C., Saffran, E., & Pate, D. 1987. Syntactic transparency and sentence interpretation in aphasia. *Language and Cognitive Processes*, **2**, 85–113.
- Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. 1994. Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, **33**, 258–318.
- Van Loon-Vervoorn, W. A. 1985. *Voorstelbaarheidswaarden van Nederlandse woorden: 4600 substantieven, 1000 verba en 500 adjectieven*. [Imageability ratings of Dutch words: 4600 nouns, 1000 verbs, and 500 adjectives]. Lisse, The Netherlands: Swets & Zeitlinger.
- Vigliocco, G., Butterworth, B., & Garrett, M. F. 1996a. Subject-verb agreement in Spanish and English: Differences in the role of conceptual constraints. *Cognition*, **61**, 261–298.
- Vigliocco, G., Butterworth, B., & Semenza, C. 1995. Constructing subject-verb agreement in speech: The role of semantic and morphological factors. *Journal of Memory and Language*, **34**, 186–215.
- Vigliocco, G., Butterworth, B., Semenza, C., & Fossella, S. 1994. How two aphasic speakers construct subject-verb agreement. *Journal of Neurolinguistics*, **8**, 19–25.
- Vigliocco, G., Hartsuiker, R. J., Jarema, G., & Kolk, H. H. J. 1996b. One or more labels on the bottles? Notional concord in Dutch and French. *Language and Cognitive Processes*, **11**, 407–442.
- Waters, G. S., & Caplan, D. 1996. The capacity theory of sentence comprehension: A reply to Just and Carpenter. *Psychological Review*, **103**, 761–772.