ACADEMIC PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Brain and Language 84 (2003) 222–249

Brain and Language

www.elsevier.com/locate/b&l

On-line syntactic processing in aphasia: Studies with auditory moving window presentation☆

David Caplan[a,*] and Gloria Waters[b]

[a] Neuropsychology Laboratory, Department of Neurology, Vincent Burnham 827, Massachusetts General Hospital, Fruit Street, Boston, MA 02114, USA

[b] Department of Communication Disorders, Boston University, USA

Abstract

Twenty-eight aphasic patients with left hemisphere strokes and matched control subjects were tested on an auditory moving windows task in which successive phrases of a sentence were presented in response to subjects' self-paced button presses and subjects made timed judgments regarding the plausibility of each sentence. Pairs of sentences were presented that differed in syntactic complexity. Patients made more errors and/or took longer in making the plausibility judgments than controls, and were more affected than controls by the syntactic complexity of a sentence in these judgments. Normal subjects showed effects of syntactic structure in self-paced listening. On-line syntactic effects differed in patients as a function of their comprehension level. High-performing patients showed the same effects as normal control subjects; low performing patients did not show the same effects of syntactic structure. On-line syntactic effects also differed in patients as a function of their clinical diagnosis. Broca's aphasic patients' on-line performances suggested that they were not processing complex syntactic structures on-line, while fluent aphasics' performances suggested that their comprehension impairment occurred after on-line processing was accomplished. The results indicate that many aphasic patients retain their ability to process syntactic structure on-line, and that different groups of patients with syntactic comprehension disorders show different patterns of on-line syntactic processing.

1. Introduction

Aphasic patients often have difficulties performing tasks such as sentence–picture matching, enacting thematic roles in a sentence, or making judgments about whether a sentence is plausible when they have to construct a relatively complex syntactic structure in order to arrive at the meaning of a sentence (for reviews and representative studies, see Berndt, Mitchum, & Haendiges, 1996; Caplan, Baker, & Dehaut, 1985; Caplan, Hildebrandt, & Makris, 1996; Caplan, Waters, & Hildebrandt,

☆ The research reported here was supported by a grant from NIDCD (DC00942).

* Corresponding author. Fax: 1-617-724-7836.

E-mail address: dcaplan@partners.org (D. Caplan).

0093-934X/02/$ - see front matter © 2002 Elsevier Science (USA). All rights reserved.
PII: S0093-934X(02)00514-X

1997; Grodzinsky, 2000). Most researchers have assumed that these abnormal performances result from a primary inability on the part of a patient to assign the syntactic structure of a sentence (Caplan & Hildebrandt, 1988; Grodzinsky, 2000) or to use that structure to determine aspects of the meaning of a sentence, such as its thematic roles (Linebarger, 1995; Linebarger, Schwartz, & Saffran, 1983a, 1983b). By "primary inability," we mean that the parsing or interpretive process is abnormal; that a patient does not assign the correct structure and/or meaning of a sentence.

If this view is correct, there should be evidence of such a dysfunction in patients' performances on on-line tests of syntactic processing. However, studies of on-line syntactic processing by aphasic patients show complex relationships between on-line and off-line performances. Several studies have been taken to show comparable on-line and off-line syntactic processing in aphasics, but there are a variety of problems that arise in interpreting these studies. Other studies have shown a dissociation between on-line and off-line tasks, with good-to-normal on-line and impaired off-line performances. We briefly review these studies.

Three studies have been taken to suggest that some aphasic patients have comparable on-line and off-line syntactic processing deficits. The first of these are reports by Zurif, Swinney, Prather, Solomon, and Bushell (1993), Swinney and Zurif (1995), and Swinney, Zurif, Prather, and Love (1996) regarding eight Broca's aphasics. These patients' performances were below normal on sentence–picture matching for reversible passive sentences. Swinney and his colleagues tested these patients for on-line syntactic processing using a cross-modal lexical priming (CMLP) task. In this task, normal subjects showed priming for words related to the head noun of relative clauses at the verb of the relative clause but not before it; i.e., they showed priming for *doctor* at point 2 but not at the control point 1 in sentences such as (1) and (2).

1. The nurse serving on the renal transplant unit (1) who administered (2) the injection replaced the vial.
2. The nurse serving on the renal transplant unit (1) who the injection dismayed (2) replaced the vial.

Broca's aphasics failed to show priming effects at the verb of the relative clause. The authors attributed Broca's aphasics' poor performance to an on-line disorder of the syntactic operation that relates the head of a relative clause to the verb of that clause. This operation is often referred to as the "co-indexation of a trace," based upon terminology from Chomsky's theory of syntactic structures (Chomsky, 1981, 1986, 1994).

There are, however, two issues that cloud this interpretation of the Swinney et al. results. One, which recurs in all existing studies, is that the patients were not tested on the same structures in off-line and on-line tasks. Therefore, it is not possible to say that their on-line and off-line performances reflected difficulties with the same structures. A second problem is that Broca's aphasics did not show semantic priming at any probe point. Broca's aphasics are well-known to often not show priming for words in isolation (Blumstein & Milberg, 1982; Milberg & Blumstein, 1981), so it is not clear whether their performance in the CMLP task resulted from a failure to co-index a trace (a syntactic problem) rather than a lexical semantic problem that affected priming in general.

Two studies by Friederici and Kilborn (1989) and Kilborn and Friederici (1994) in German speaking aphasics were also interpreted by the authors as showing abnormal on-line syntactic processing that was related to the off-line comprehension difficulties of Broca's aphasics. In the first study (Friederici & Kilborn, 1989), five agrammatic patients, whose off-line comprehension was impaired as measured on the

Aachen Aphasia Test, performed written lexical decision tasks for words in isolation and following an auditory sentence fragment. Unlike age-matched normal controls, who showed no effect of the sentence fragment, the patients' reactions times for lexical decisions were slower in the sentential condition. Moreover, the patients' RTs were slower when the lexical decision target was presented immediately after the auditory sentence fragment than when its presentation was delayed by 200 ms. The authors suggested that this result reflected a slowing down of lexical access due to interference from ongoing syntactic processing at the point that the sentence fragment ended, and that it was consistent with slowed syntactic processing in Broca's aphasics. However, Broca's aphasics showed sensitivity to two aspects of the sentence fragment in the 0 ms ISI condition: RTs were longer to words that created grammatically correct sentences than to words that did not, and RTs were longer to words following an auxiliary that imposed constraints on the argument structure of the subsequent verb than to words following auxiliaries that did not. This indicates that some syntactic processing was completed and that some aspects of syntactic structure were available immediately after the auditory sentence fragment. In addition, the sentences on which on-line and off-line performances were assessed were not of the same types, making it hard to know what the relationship was between any on-line disturbance seen on this task and the off-line comprehension impairment documented in these patients.

In a second study, Kilborn and Friederici (1994) repeated the LD task with four of the five Broca's patients, with the instruction that the auditory sentence fragment should be ignored. The age-matched subjects showed grammaticality effects at both 0 and 200 ms ISIs, and auxiliary type effects at 200 ms. The authors interpreted this pattern of results as indicating that syntactic processing is automatic and cannot be consciously overridden in normal subjects. The grammaticality effect was no longer present in Broca's aphasics in the instruction-to-disregard condition. The authors interpreted this as an indication that on-line syntactic processing in Broca's aphasics is controlled, not automatic. However, the finding of persistant effects of grammatical context in normals but not aphasics told to ignore the context does not necessarily indicate that on-line processing is a controlled process in Broca's aphasics. It could be that the patients have difficulty focusing attention in general and tend to ignore a spoken sentence more readily than normals. Especially in a complex cross-modal task where the responses do not depend upon attending to the spoken sentence, aphasic patients may need the extra stimulus of an instruction to attend more than controls do in order to focus on the sentence while doing the task. Syntactic processing may still be automatic in these patients, once it is engaged. The problem may be in the attention control system, not the on-line sentence comprehension process.

Overall, these studies do not provide strong evidence that off-line syntactic comprehension disturbances is related to failures of on-line parsing. The view that off-line syntactic comprehension disturbances are always due to failures of on-line parsing is also contradicted by other studies that document abnormal off-line comprehension but appear to show good on-line syntactic processing in aphasic patients.

Tyler (1985) studied one such patient, DE, an agrammatic aphasic, whose off-line, end-of sentence, anomaly judgments were much less accurate than those of normals. DE's on-line performances gave evidence of syntactic processing. In a word monitoring task, he had longer latencies for word targets in syntactically correct, semantically anomalous prose than in normal prose, indicating a sensitivity to sentential meaning, and even longer latencies in word salad, indicating a sensitivity to syntactic structure. His word monitoring latencies also showed a normal tendency to have faster reaction times in the second and last thirds of normal prose but not in word salad, another indication of sensitivity to the accruing syntactic and sentential semantic representation. He showed normal effects of semantic and syntactic anomalies on monitoring times for the words following an anomalous word. These performances show on-line sensitivity to syntactic structure, though the specific aspects of syntactic structure to which DE was sensitive are not characterized.

A second study that reported a similar dissociation between at least partially preserved on-line syntactic processing and poor off-line syntactic comprehension is that of Shankweiler, Crain, Gorrell, and Tuller (1989). These authors had six agrammatic aphasics make on-line wellformedness judgments to auditorily presented sentences. Several features of the data suggested that the patients were sensitive to syntactic structure. As in Tyler's (1985) case, the patients' judgments were faster as each sentence progressed. In addition, the distance between the anomalous and licensing segments affected reaction times, with faster RTs when this distance was smaller, and patients were better able to detect anomalies that involved between-grammatical-class substitutions than those that involved within-class substitutions. Four of the patients performed poorly on an off-line comprehension test of semantically reversible passive sentences. Like DE, this pattern suggests that the patients were sensitive to aspects of grammatical form on-line but could not use syntactic information to determine sentence meaning off-line.

A third study, by Swinney et al. (1996) and Swinney and Zurif (1995), also showed good on-line syntactic processing along with poor off-line comprehension in four fluent aphasic patients. The fluent aphasics performed poorly in a standard off-line sentence–picture matching task with reversible passive sentences. Swinney and his colleagues tested these patients for on-line processing on the CMLP task described above. They showed the normal pattern of priming for the head of the relative clause when they encountered the verb of that clause, and not before. The authors took this priming pattern as an indication that these patients re-activated the head of the relative clause normally when they encountered the verb of that clause.

These studies suggest that aphasic patients may show disorders of off-line comprehension of sentences that require syntactic analyses to understand correctly even though they retain normal on-line processing of syntactic structure. This is a counter-intuitive conclusion, that would require developing a new analysis of the locus of the deficit in these patients. However, it may be premature to draw this conclusion, because there are also problems in interpreting the results of these three studies.

In the Tyler and Shankweiler studies, the patients' on-line performances were not completely normal. DE did not show faster reaction times in the second and last thirds of anomalous prose, which suggests some failure of sensitivity to syntactic structure. Tyler argued that, in conjunction with his other performances, this result suggested that DE's on-line syntactic processing was restricted to local constituents and did not extend to global sentence-wide structures. The patients in the Shankweiler et al. (1989) study only detected an average of 75% of the anomalies that were presented and rejected an average of 20% of well-formed sentences as anomalous on-line—far worse performances than found in the control group, whose hit and correct rejection rates were both 97%. In addition, neither the Shankweiler et al. nor the Swinney et al. studies tested on-line processing of the same structures tested for off-line comprehension. Finally, in the Swinney et al. (1996) study, McKoon, Ratcliff, and Ward (1994) and McKoon and Ratcliff (1994) argued that the priming effects found with these materials may be due to a better pragmatic fit of the related than the unrelated words to the pragmatic context at the point of the verb (see Nicol, Fodor, & Swinney, 1994, for a reply). Thus, it is premature to conclude that aphasic patients can retain normal on-line syntactic processing but show off-line syntactic comprehension deficits.

Swinney and his colleagues (see also Grodzinsky, 1990, 1995, 2000) have introduced the idea that some aphasic patients' off-line syntactic comprehension disorders result from disturbances of on-line syntactic processing, while others' do not. These authors have argued that Broca's aphasics have impairments of the on-line process of co-indexation of a trace, whereas fluent aphasics do not. The caveats discussed above about the interpretation of the Swinney et al. results raise questions about this hypothesis.

To summarize this literature, a few studies have shown disordered on-line syntactic processing in some aphasic patients. Off-line syntactic comprehension disorders have been attributed to the observed on-line deficits in some patients. Particular claims have been made regarding one aspect of on-line parsing in Broca's aphasics. In addition, some authors have described what they take to be preserved on-line syntactic processing in some aphasic patients. The presence of normal on-line processing might simply indicate that the aphasic impairments in these patients did not extend to syntactic processing in comprehension, but these studies also documented abnormal off-line syntactic comprehension in these patients. However, there are questions about the interpretation of all these studies. The relationship between on-line processing of syntactic structure and off-line performance on tests of syntactic comprehension in aphasia remains to be clarified. In the present study, we re-examined the question of the relationship between on-line and off-line processing of syntactic structure in aphasic patients.

2. Methods

2.1. Subjects

Two groups of subjects were tested: 28 aphasic patients with left hemisphere strokes, and 28 control subjects matched for age and education to the aphasic patients.

The aphasic patients were recruited from rehabilitation hospitals in Boston (21 patients) and Montreal (seven patients). All were right-handed native English speakers who had suffered from a single left hemisphere stroke. There were 13 female and 15 male patients. The patients ranged in age from 30 to 80 years (mean 57.6 years) and had from 7 to 22 years of education (mean 14.3 years). All patients were tested in a stable state, at least six months since their stroke. Demographic and clinical information about these patients is given in Table 1.

The non-brain damaged control subjects were recruited by advertisements in Boston. All were native English speakers. Twenty-eight control subjects were matched on an individual basis for age and education to the aphasic patients. They ranged in age from 28 to 80 years (mean 67.5 years) and had from 9 to 20 years of education (mean 14.3 years). There were 17 females and 11 males in the control group for the aphasic patients.

2.2. Procedures and materials

2.2.1. Screening tests

Patients were screened for auditory word recognition, auditory single word comprehension, and auditory sentence comprehension using materials from the Psycholinguistic Assessment of Language (PAL) (Caplan, 1992).

Table 1
Characteristics of the aphasic patients

Subject	Age	Gender	Education	Lesion area	Aphasia type
1	46	F	16	L posterior-frontal/parietal CVA	
2	62	M	18	L temporal-parietal hemorrhage	Fluent
3	72	F	12	L MCA and ACA frontoparietal CVA	Broca's
4	47	M	19	L basal ganglia CVA	
5	79	F	12	L basal ganglia and L parietal CVAs	
6	35	F	12	L medial temporal and basal ganglia CVA	
7	62	M	10	L CVA, left carotid artery total occlusion	Broca's
8	57	F	16	L CVA in MCA distribution	Broca's
9	68	M	18	L CVA in MCA distribution	Fluent
10	57	F	12	L frontal CVA	Broca's
11	69	F	14	L subcortical CVA	
12	30	M	12	L CVA	
13	54	M	21	L frontotemporal CVA	
14	50	M	18	L CVA, L ICA occlusion	
15	59	M	18	L temporoparietal hemorrhage	Fluent
16	54	F	22	L frontoparietal CVA	
17	38	M	11	L CVA	
18	70	F	12	L MCA aneurysm	Broca's
19	64	M	13	L temporoparietal CVA	
20	71	M	12	L CVA in MCA distribution	
21	58	M	16	L CVA	
22	68	M	7	L CVA	Fluent
23	42	F	15	L frontoparietal CVA	Broca's
24	80	F	11	L frontoparietal CVA	Broca's
25	62	M	16	L subcortical CVA	
26	45	F	14	L frontoparietal CVA	
27	65	F	9	L Fronto-temporo-parietal CVA	Broca's
28	48	M	14	L parietal CVA	Fluent

Phoneme discrimination. The ability to discriminate phonemes was tested by a same–different task with 40 pairs of mono-syllabic non-words (20 different and 20 identical trials). Different stimuli differed in a single consonantal phoneme with respect to place of articulation, manner of articulation, or voicing. The changed phoneme occurred in stimulus-initial or stimulus-final position, either as a single consonant or as a member of a cluster. Stimuli were presented auditorily using Psyscope (Cohen, MacWhinney, Flatt, & Provost, 1993). Stimulus items within pairs were separated by 500 ms and pairs of stimuli followed responses by 1 s. Subjects indicated orally whether the two items were identical or different.

Auditory lexical decision. Auditory lexical access was assessed using a lexical decision task for words and specifically constructed non-words. The words consisted of 40 concrete nouns. They varied in frequency (>40/million or <5/million) in Francis and Kucera (1982) and length (one syllable vs three or more syllables). Half the foils were constructed by changing a single distinctive feature in a single phoneme in different syllabic positions in comparable words. The other 20 foils were created by changing the form of words matched to the positive targets so as to resemble possible words (e.g., *harpisform* from *harpsichord*). Stimuli were presented auditorily via Psyscope and a yes/no (word/non-word) decision was made orally to each item.

Word–picture matching. Thirty-two concrete nouns were presented auditorily via Psyscope and the subject's task was to select one of two pictures as the match to the word. Targets were of either high or low frequency, and were either short (monosyllabic) or long (tri- or quadrisyllabic). They included examples from the categories of animals, fruits and vegetables, and tools. Foils were both semantically and visually similar to the targets (e.g., *deer* as target and *moose* as foil). Stimuli were presented auditorily via Psyscope to the subject, who selected the appropriate picture by pushing one of the two response keys.

Sentence comprehension. Sentence comprehension was tested with the sentence–picture matching test from the PAL. Two aspects of sentence comprehension were evaluated. The ability to understand sentences based upon lexical–pragmatic inferences was tested with twenty semantically irreversible sentences. Pictures for these stimuli consisted of the correct interpretation and a foil that varied with respect to one of the categories (e.g., Target: *The car was waxed by the man*; Foil: *The car was washed by the man*). Twenty sentences varying as to voice (active and passive) and nature of foil (verb, preposition, particle) were presented auditorily via Psychlab to the subject, who selected the appropriate picture by pushing one of the two response keys.

The ability to use syntactic structure to determine sentence meaning was evaluated using 20 semantically reversible sentences with four syntactic structures—active, passive, dative-passive, subject–object relative. These were presented in a sentence–picture matching test with correct pictures and syntactically incorrect foils (e.g., Target: *The man was pushed by the woman*; Foil: *The man pushed the woman*). Sentences were presented auditorily via Psyscope to the subject, who selected the appropriate picture by pushing one of two response keys.

Sentence–picture matching. Sentence comprehension was also tested with a larger sentence–picture matching test (Caplan et al., 1997), consisting of 100 semantically reversible sentences with 10 syntactic structures: active, active with conjoined theme, dative, passive, truncated passive, dative passive, cleft–object, conjoined, object–subject, and subject–object. Procedures were as in Waters, Caplan, and Rochon (1995) and Caplan et al. (1997).

2.2.2. On-line syntactic processing

On-line syntactic processing was assessed using the Auditory Moving Windows paradigm (Ferreira, Henderson, Anes, Weeks, & McFarlane, 1996). The methods and materials were taken from previous studies with college students and elderly individuals (Waters & Caplan, 2001). In this task, on each trial subjects heard a sentence that had been digitized and segmented into a series of phrases. On half the trials the sentence was plausible and on the other half it was implausible. Subjects' task was to pace their way through the sentence as quickly as possible, by pressing a button on a box interfaced with the computer for the successive presentation of each phrase, and then to make a plausibility judgment about the sentence they had just heard. Reaction times for each button press, as well as response time and accuracy on the plausibility judgment, were recorded in Psyscope. The assumption underlying the self-paced listening task is that, when lexical factors are eliminated, listening times to words or phrases presented one at a time reflect the time it takes to integrate lexical items into an accruing syntactic and semantic structure, and are therefore longer when this integration is more difficult.

The target stimuli consisted of 104 semantically plausible and 104 semantically implausible sentences divided equally among the four sentence types cleft–subject (CS), cleft–object (CO), object–subject (OS), and subject–object (SO), shown in Table 2. The memory storage and computational requirements of processing this

Table 2
Sample stimuli used in the auditory moving windows paradigm

Cleft–subject sentences (CS)					
Phrase	Intro	NP1	V	NP2	
Acceptable	It was/	the food/	that nourished/	the child	
Unacceptable	It was/	the car/	that drove/	the woman	
Cleft–object sentences (CO)					
Phrase	Intro	NP1	NP2	V	
Acceptable	It was/	the woman/	that the toy/	amazed	
Unacceptable	It was/	the coffee/	that the secretary/	disappointed	
Object–subject sentences (OS)					
Phrase	NP1	V1	NP2	V2	NP3
Acceptable	The father/	read/	the book/	that terrified/	the child
Unacceptable	The girl/	drank/	the boy/	that entered/	the hospital
Subject–object sentences (SO)					
Phrase	NP1	NP2	V1	V2	NP3
Acceptable	The man/	that the fire/	injured/	called/	the doctor
Unacceptable	The secretary/	that the camera/	met/	drove/	the car

structure make the processing load greater in CO and SO sentences than in CS and OS sentences, respectively (Gibson, 1998). Both theoretical considerations and empirical results point to loci where increased processing load is expected in CO and SO sentences.

In the comparison of CO and CS sentences, the point of increased syntactic processing load is expected to be the verb (V) of CO sentences (Caplan, Hildebrandt, & Waters, 1994). Listening times are therefore predicted to be longer for V of CO sentences compared to V of CS sentences. However, this comparison might also reflect the fact that V is the sentence final word of CO sentences, which is associated with "wrap-up" effects (Balogh, Zurif, Prather, Swinney, & Finkel, 1998). Such effects would also be expected to make listening times for the second noun phrase (NP2) longer in CS sentences than in CO sentences. The effects of syntactic load (and possibly discourse factors—see Section 4) would, however, be expected to be seen as longer listening times for V in CO than for NP2 in CS sentences, both of which are sentence-final words.

Similar considerations apply to SO and OS sentences. Listening times are expected to be longer at the first verb (V1) in SO than in OS sentences (Waters & Caplan, 2001). This may be either a syntactic or a clause-final effect. Clause-final effects are expected to also lead to longer listening times at the second noun (NP2) in OS than in SO sentences. Syntactic factors are expected to lead to longer listening times at V1 in SO sentences than at NP2 in OS sentences (both clause-final words) and at the second verb (V2) in SO sentences compared to V2 in OS sentences (King & Just, 1991).

The sentences were closely matched for length, with the mean number of words being 8.2 in both CS and CO plausible sentences and 8.1 in CS and CO implausible sentences. OS plausible and implausible sentences had a mean of 9.2 words and SO sentences had a mean of 9.3 words.

The stimuli were designed so that plausibility judgments did not require judgments based upon detailed semantic knowledge. Sentences had verbs that require either animate objects or animate subjects (e.g., *It was the girl that the food nourished*). Implausible sentences violated these constrains. Detection of implausibility could thus be based on fairly accessible, general semantic features and did not require extensive searches through semantic memory for item-specific information.

Sentences were constructed to prevent subjects responding correctly on the basis of the order of animate and inanimate nouns in a sentence. The animacy of the noun not affected by animacy constraints of the verb was always opposite to that of the noun affected by these constraints. In the plausible sentences, the subject noun was inanimate, because previous research has found that this makes for the most difficulty in object relativized sentences (Caplan et al., 1994). In half of the implausible sentences, the subject noun was animate and the object noun inanimate, and in half the pairing was reversed. The point at which the sentences became implausible occurred in the first clause in half of the sentences, and in the second clause for the other half, to ensure that subjects had to pace their way through the entire sentence before making a decision regarding plausibility.

A male speaker read the stimulus sentences to be recorded in a semianechoic chamber. The stimuli were recorded and digitized by SoundEdit (Dunn, 1994) at a sampling rate of 44.1 kHz and a 16 bits sample size, stored as a waveform file, and edited using SoundEdit. A marker (referred to as a "tag") was placed in the waveform at the locations defining the boundaries of presentation of segments. Table 2 shows the location of tags (indicated by /) for the four sentence types used in the experiment. In order to make segment-to-segment transitions smooth, tags were placed in the waveform at areas of low signal amplitude, as indicated by auditory and visual inspection, whenever possible. When word boundaries did not coincide with areas of low signal amplitude, the tags were placed so as to maximize the intelligibility of the words. A tone was appended to the waveform of every sentence immediately following the offset of visible and auditory activity associated with the sentence-final word, since work by Ferreira and her colleagues had shown that with this technique, subjects often cannot tell when a spoken sentence has ended.

The resulting waveform files were then entered into Psyscope (Cohen et al., 1993). The experiment was controlled by a Macintosh PowerPC computer equipped with an additional digital I/O board and button box for gathering response times. When the subject pushed a button, the time of the button-press was recorded, and the waveform up to the first tag was converted from digital to analogue format and played on headphones. Subjects pushed the button again when they were ready to hear the next segment. The button-press time was recorded again and the material between the next tags was played. In order to discourage subjects from pressing the button before they had heard and processed each segment, if a subject pressed the button before the end of a segment, the segment was truncated at the point of the button press. Inter-response times were computed. These times included the time required to play out and process each segment of the sentence. The computer also recorded plausibility judgment times (time from the end of the sentence to the subject's button-press) and plausibility judgment accuracy for each sentence.

3. Results

3.1. Results of screening tests

Non-brain-damaged subjects have been tested on all the screening tests and their performances have been at or very close to ceiling. The performances of the aphasic patients on the screening tests are listed in Table 3.

Table 3
Proportion of items correct on the screening tests for the aphasic patients

Subject	Phoneme discrimination	Auditory lexical decision	Auditory wordpicture matching	Constrained sentence comprehension	Reversible sentence comprehension	Sentence picture matching
1	NA	0.95	1.00	1.00	0.85	0.91
2	0.90	0.96	0.97	1.00	0.95	0.87
3	0.70	0.94	1.00	0.85	0.65	0.66
4	0.90	0.99	1.00	0.95	0.95	0.96
5	0.60	0.49	1.00	1.00	0.85	0.80
6	0.95	0.925	1.00	0.95	0.85	0.94
7	NA	0.83	0.97	0.89	0.50	0.79
8	0.82	0.91	0.90	1.00	0.35	0.60
9	0.75	NA	0.97	0.90	0.75	0.80
10	0.95	NA	1.00	0.89	0.70	0.81
11	0.98	0.95	1.00	0.95	0.84	0.93
12	0.78	0.85	1.00	0.95	0.75	0.73
13	0.43	0.83	0.97	0.90	0.55	0.62
14	1.00	0.98	0.97	1.00	0.85	0.80
15	0.98	0.91	1.00	N/A	N/A	0.78
16	0.98	0.98	0.91	0.70	0.65	0.51
17	0.75	0.73	0.88	0.85	0.65	0.81
18	0.75	0.95	0.91	0.65	0.45	0.62
19	0.50	0.89	0.97	0.85	0.55	0.68
20	0.88	0.90	1.00	0.95	0.80	0.83
21	0.90	0.96	1.00	0.80	0.65	0.54
22	0.83	0.89	1.00	1.00	0.95	0.83
23	0.83	0.96	1.00	1.00	0.95	0.97
24	0.65	0.94	1.00	0.9	0.7	0.56
25	0.98	0.83	0.97	1.00	0.95	0.92
26	0.93	0.99	1.00	0.95	0.95	0.89
27	0.78	0.94	0.97	0.85	0.5	0.5
28	0.85	0.84	NA	0.95	65	0.73
Patients mean	0.82	0.9	0.98	0.91	0.92	0.76
Controls mean	92.7[a]	93.2[a]	96.8[a]	95[a]	90[a]	91.3[b]

[a] Based on performance of 100 normal subjects.
[b] Based on performance of 14 normal subjects.

Performance on phoneme discrimination ranged from 50 to 90% correct. All but one patient scored above 90% on word–picture matching (the remaining patient scored 88%). Twenty-three patients scored 85% correct or higher on auditory lexical decision. One patient scored 50% correct on lexical decision, but his word–picture matching performance was perfect. These scores indicate that the patients could recognize and understand simple spoken words. Scores on the sentence comprehension tests were variable. On the PAL, all patients scored 80% or better on semantically irreversible sentences. Scores on reversible sentences were more variable, with a number of performances at chance. Scores on the more extensive sentence–picture matching task were also variable, with performances ranging from chance to 94% correct. Performances were worse on the more complex sentences in this task. Performances on the semantically reversible sentences of the PAL and on the longer sentence–picture matching test were significantly correlated ($r = .61$). Performances on these two tasks indicate that these patients had variable degrees of difficulty with syntactically based comprehension as judged by performance on an off-line task.

3.2. Results of auditory moving windows task

Accuracy in making the end-of-sentence plausibility judgment (expressed as A′) were analyzed in 2 (Group) × 4 (Sentence Type) ANOVAs by subjects (item analyses cannot be performed on A′s). RTs for plausible sentences that were responded to correctly were analyzed in 2 (Group) × 4 (Sentence Type) ANOVAs by subjects and items, after RTs greater and less than 3 SD from the mean for each condition for each subject had been removed as outliers.

The dependent measure for the auditory moving windows task consisted of the response time for each phrase in plausible sentences that the subject correctly judged to be plausible. Since, as noted above, the response times for each of the segments in this task included the duration of the segment as well as the time required by the subject to process it, corrected response times ("listening times") were calculated by subtracting the segment's tag-to-tag duration from the response time. Listening times were eliminated if they were beyond a predetermined cut-off point (2000 ms for normal subjects; 5000 ms for patients). Listening times may have been affected by the frequency of the lexical items in the sentences (Ferreira et al., 1996), which differed for the same segments in different sentences. To correct for the effects of frequency, we used a technique originally proposed by Ferreira and Clifton (1986) to correct for effects of letter length and word frequency in self-paced reading experiments effects. In this approach, a regression analysis is performed in which listening times are regressed against log frequency for each word in each subject, and the difference between the actual and predicted listening time for each word (the "residuals" of this analysis) are utilized in analyses of the AMW data. Finally, residual listening times greater and less than 3 SDs from the mean for each condition for each subject were removed as outliers. Because there were four segments in the CS and CO sentences and five in the OS and SO sentences, the CS and CO sentences and the OS and SO sentences were analyzed separately. Residual listening times for cleft–subject and cleft–object sentences were analyzed in 2 (Group) × 2 (SentenceType) × 4 (Phrase) ANOVAs by subjects and items. Residual listening times for object–subject and subject–object sentences were analyzed in 2 (Group) × 2 (SentenceType) × 5 (Position) ANOVA by subjects and by items.

We also analyzed listening times for plausible cleft–object and subject–object sentences that the aphasics responded to incorrectly. The control subjects made too few errors to analyze their listening times in sentences that provoked incorrect judgments, and there were also too few errors made by the aphasics on cleft–subject and object–subject sentences to analyze listening times in these sentences in the aphasics. In addition, since errors were scattered over sentences of a given type, item statistics of the aphasics' listening times to CO and SO sentences would have been based on responses made by a very small number of subjects in many sentences. We therefore report only the F1 values for these data. The data for CO sentences were analyzed in a 2 (Accuracy) × 4 (Phrase) ANOVA by subjects and the data for SO sentences were analyzed in a 2 (Accuracy) × 5 (Phrase) ANOVA by subjects. So that these analyses would be based on reliable listening time measurements in individual subjects, we undertook them for subjects who made 5 or more errors on a sentence type.

Reaction times on the plausibility judgments and listening times for implausible sentences were not analyzed because of the differences in the points at which the sentences became implausible.

A large number of main effects were significant in the analyses of listening times, but were qualified in most analyses by interactions among variables. We report the highest order interactions that achieved a significance at least a level of $p < .05$ in one of the analyses and a level of $p < .1$ in the other (i.e., that were significant in at least one analysis and at a level of a trend in the other). In all analyses, Tukey's test was used to examine for significant differences in terms of significant interactions. We report the results of Tukey's tests by subjects and by items.

3.3. Aphasic patients and control subjects

3.3.1. Plausibility judgment results

Results of the plausibility judgments are shown in Fig. 1.

In the analysis of A′ scores, there were main effects of Group ($F1(1,54) = 53.6$, $p < .001$) and Sentence Type ($F1(3,162) = 18.8$, $p < .001$). A′s were higher for controls than for the aphasics, and higher for CS sentences than for CO sentences and for OS than for SO sentences. There was a significant interaction of Group and Sentence Type ($F1(3,162) = 7.9$, $p < .001$). A′s were significantly higher for CS compared to CO sentences and for OS compared to SO sentences in the aphasics but not in the controls. A′s were lower for aphasics than controls on CS, CO, and SO sentences.

In the analysis of reaction times, there were main effects of Group ($F1(1,54) = 4.2$, $p < .05$; $F2(1,200) = 40.3$, $p < .001$) and Sentence Type ($F1(3,162) = 9.5$, $p < .001$; $F2(3,200) = 11.0$, $p < .001$). RTs were longer for the aphasics than for the control subjects. RTs were longer for CO sentences than for CS sentences and for SO sentences than for OS sentences.The interaction of Group and Sentence Type was not significant.

3.4. Self-paced listening results

Fig. 2 shows the mean residual listening time for CS, CO, OS, and SO sentences at each phrase for the aphasic patients and control subjects.

For CS and CO sentences, there was a three-way interaction between group, sentence type, and phrase ($F_1(3,162) = 8.1$, $p < .001$; $F_2(3,200) = 46.1$, $p < .001$). Post-hoc tests showed that both controls and aphasics had longer residual listening times to the second noun phrase (NP2) in CS sentences than to NP2 in CO sentences, to the verb (V) in CO sentences than to V in CS sentences, and to V in CO sentences than to NP2 in CS sentences. The three-way interaction arose because residual listening times were longer at V of CO and at NP2 of CS sentences in the aphasics than in the controls. The difference between V in CO and NP2 in CS sentences did not differ in the two groups (281 ms for controls; 289 ms for aphasics).

For the OS and SO sentences, there was a significant interaction between sentence type and phrase ($F_1(4,216) = 24.9$, $p < .001$; $F_2(4,250) = 46.3$, $p < .001$). Post-hoc tests by subjects showed that residual listening times were longer at V1, V2, and NP3 in SO than in OS sentences, at V1 in SO sentences than at NP2 in OS sentences, and at NP2 in OS than in SO sentences. Post-hoc tests by items showed that residual listening times were longer at V2 in SO than in OS sentences. There was also an interaction between group and phrase ($F_1(4,216) = 14.7$, $p < .001$; $F_2(4,250) = 10.4$, $p < .001$). This interaction was due to longer residual listening times at NP3 in aphasics than in controls and not elsewhere.

3.5. Listening times in incorrect responses

Fig. 3 compares residual listening times for the aphasic patients for the plausible CO and SO sentences that they correctly judged to be plausible with residual listening times for plausible CO and SO sentences on which they made judgment errors. There was a main effect of Phrase in the analysis of both sentence types (for CO sentences, $F1(3,81) = 52.6$, $p < .001$; for SO sentences, $F1(4,84) = 31.1$, $p < .001$).

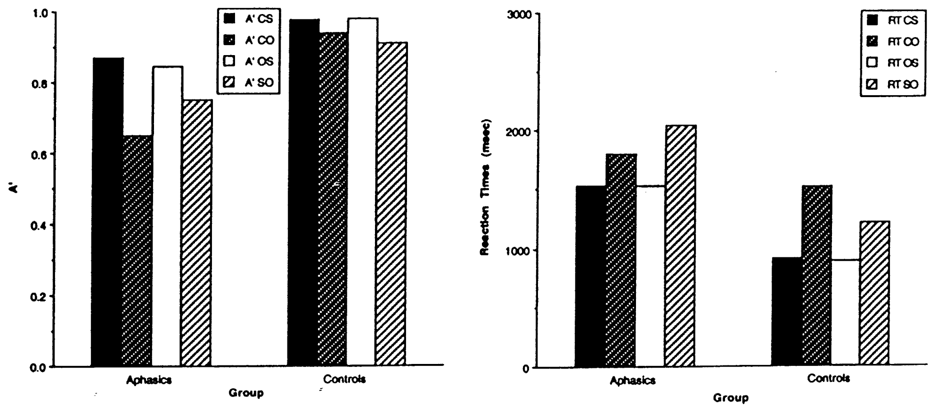

Fig. 1. A′ scores and reaction times (in milliseconds) for plausibility judgments made by aphasic and control subjects.

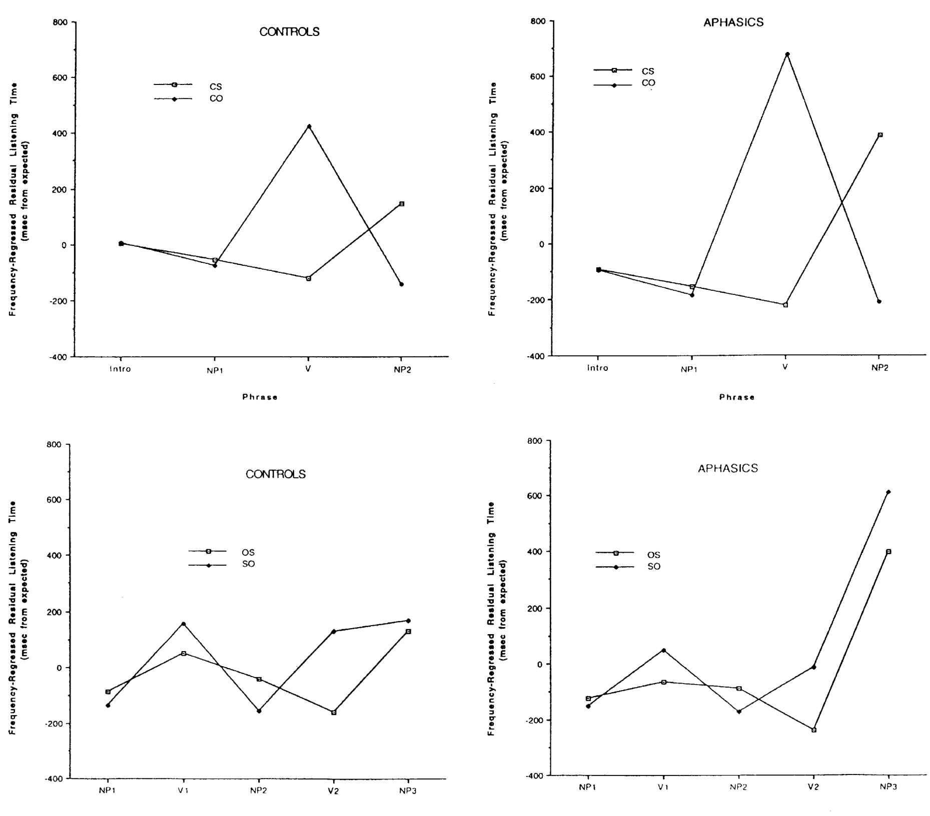

Fig. 2. Residual listening times for aphasic patients and control subjects for plausible sentences to which correct judgments were made. CS: cleft–subject sentences; CO: cleft–object sentences; OS: object–subject sentences; SO: subject–object sentences. Intro: "It was" in CS and CO sentences; NP1: first noun phrase; NP2: second noun phrase; NP3: third noun phrase; V: verb of CS and CO sentences; V1: first verb of OS and SO sentences; V2: second verb of OS and SO sentences.

For CO sentences, residual listening times were longer on the verb than on other segments. For SO sentences, residual listening times were longer at NP3 than at any other segment. The main effect of Accuracy and the interaction of Accuracy × Phrase were not significant in either analysis.

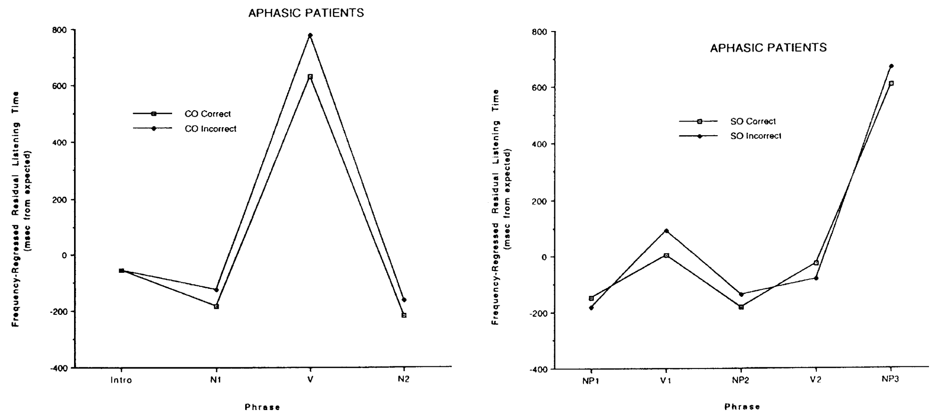

Fig. 3. Residual listening times for aphasic patients and control subjects for plausible cleft–object (CO) and subject–object (SO) sentences to which correct and incorrect judgments were made.

3.6. Discussion of results for aphasics and controls

The results of this study showed effects of syntactic complexity on plausibility judgments. Both aphasics and controls took longer to make these judgments about syntactically complex sentences. Aphasics also showed effects of syntactic complexity on the accuracy of their judgments while the controls did not. These results are consistent with other reports that aphasic patients show effects of syntactic structure on accuracy of performance on tasks such as sentence–picture matching and enactment, while normal subjects often only show effects of syntactic structure on RTs in comprehension tasks. Aphasics' RTs were also slower than controls', a result that is also consistent with reports in the literature.

In the self-paced listening data for controls, residual listening times were longer at points where processing load is expected to be high. In the comparison of CO and CS sentences, these points were V in CO compared to V and NP2 in CS sentences and NP2 in CS compared to NP2 in CO sentences. As noted above, longer residual listening times at V in CO and NP2 in CS sentences are likely to reflect sentence-final processing effects. The longer residual listening time for V in CO than for NP2 in CS sentences is attributable to syntactic and discourse factors, which we will consider in more detail in Section 4. For OS and SO sentences, residual listening times were longer at V1, V2, and NP3 in SO than in OS sentences, and at NP2 in OS than in SO sentences. The longer residual listening times at V1 in SO than in OS sentences and at NP2 in OS than in SO sentences are likely to reflect end-of clause effects. The longer residual listening times for V1 in SO than for NP2 in OS sentences and for V2 in SO than for V2 in OS sentences are attributable to the syntactic processing load at that point. All these effects are expected on the basis of theoretical considerations and are consistent with results of previous studies. The longer residual listening times for NP3 in SO than in OS sentences may reflect review processes that take place, which are more demanding in more complex sentences (see Section 4).

The listening time data for the aphasics are surprising. Despite their poorer off-line performance and their difficulties understanding the syntactically more complex object-relativized CO and SO sentences, the aphasic patients' residual listening times for individual segments in the plausible sentences that they correctly judged to be plausible were similar to those of normal subjects. The only differences between the aphasics and the controls were those found in the end-of-sentence effects—longer residual listening times at V in CO sentences, NP2 in CS sentences, and NP3 in SO and OS sentences in aphasics than in controls. In addition, patients showed the same pattern of residual listening times for segments in the more difficult plausible sentences (types CO and SO) on which they made errors as on sentences of these types they correctly judged to be plausible. This suggests that they were assigning syntactic structure and propositional meaning on-line in the same fashion in sentences they responded to correctly and those they responded to incorrectly.

The results provide evidence for the integrity of on-line syntactic processing in patients who manifest off-line disturbances in syntactic comprehension. As noted above, this raises questions about the locus of the deficit in these patients. However, it is possible that these results are strongly influenced by the performance of aphasic patients with mild impairments, whose on-line performances might be expected not to differ from controls. Therefore, we examined the performance of good- and poor-performing patients.

3.7. Good and poor comprehender aphasic patients

Patients were divided into good and poor comprehenders based on their performance on the reversible sentences of the PAL. Patients with scores higher than 75% correct on the semantically reversible sentences of the PAL were grouped as good comprehenders; those with scores below 75% were considered poor comprehenders. Twelve subjects were classified as good comprehenders and 16 as poor comprehenders by these criteria. The good comprehenders achieved a mean of 90% correct responses, and the poor comprehenders a mean of 63% correct responses on this test.

3.8. Plausibility judgment results

Reaction times and A′ scores on the plausibility judgment task are shown in Fig. 4.

Analysis of A′ scores showed main effects of Group ($F1(3,76) = 16.7$, $p < .001$) and Sentence Type ($F1(3,76) = 19.4$, $p < .001$). Good comprehenders had higher A′s than poor comprehenders. Performance was more accurate on CS than on CO sentences and on OS than on SO sentences. The interaction of Group and Sentence Type was significant ($F1(3,76) = 7.3$, $p < .001$). A′s were higher for OS than for SO sentences and for CS than for CO sentences in the poor comprehenders, but not in

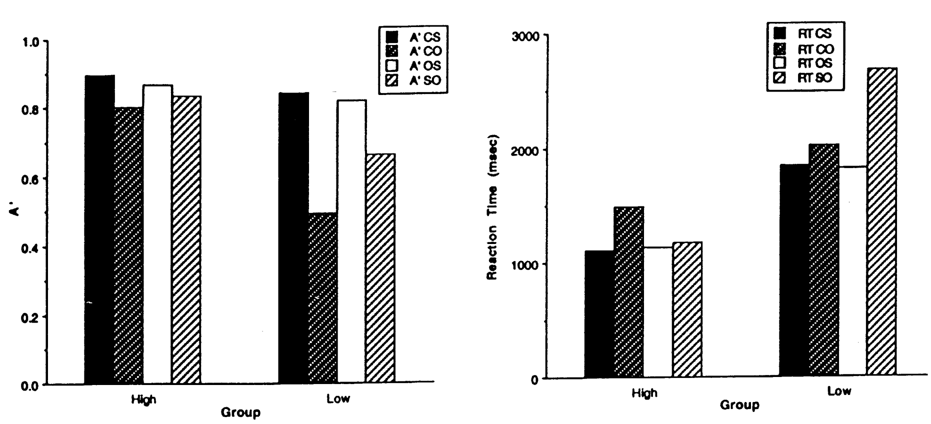

Fig. 4. A′ scores and reaction times (in milliseconds) for plausibility judgments made by good and poor comprehender aphasic patients.

the good comprehenders. Good comprehenders had higher A′s than poor comprehenders on CO and SO sentences.

The reaction time data from the plausibility judgments showed a main effect of Sentence Type ($F1(3,78) = 3.3$, $p < .05$; $F2(3,200) = 4.1$, $p < .001$). Post-hoc analyses showed that reaction times were longer for SO sentences than for OS sentences, but not for CO compared to CS sentences. There was an interaction of Group and Sentence Type ($F1(3,78) = 3.2$, $p < .05$; $F2(3,200) = 2.8$, $p < .05$). RTs were longer for SO than for the OS sentences in the poor comprehenders but not in the good comprehenders; the difference between CO and CS sentences was not significant in either group. RTs were longer for the poor comprehenders than for the good comprehenders on CS and SO sentences.

3.9. Self-paced listening results

Residual listening times for the good and poor comprehenders on plausible sentences judged as plausible are presented in Fig. 5.

For the CS and CO sentences, there were interactions between group and sentence type ($F1(3,78) = 4.4$, $p < .05$; $F2(3,200) = 6.3$, $p < .01$) and between phrase and sentence type ($F1(3,78) = 43.7$, $p < .001$; $F2(3,200) = 130.2$, $p < .001$). The group by sentence type interaction was due to good but not poor comprehenders showing

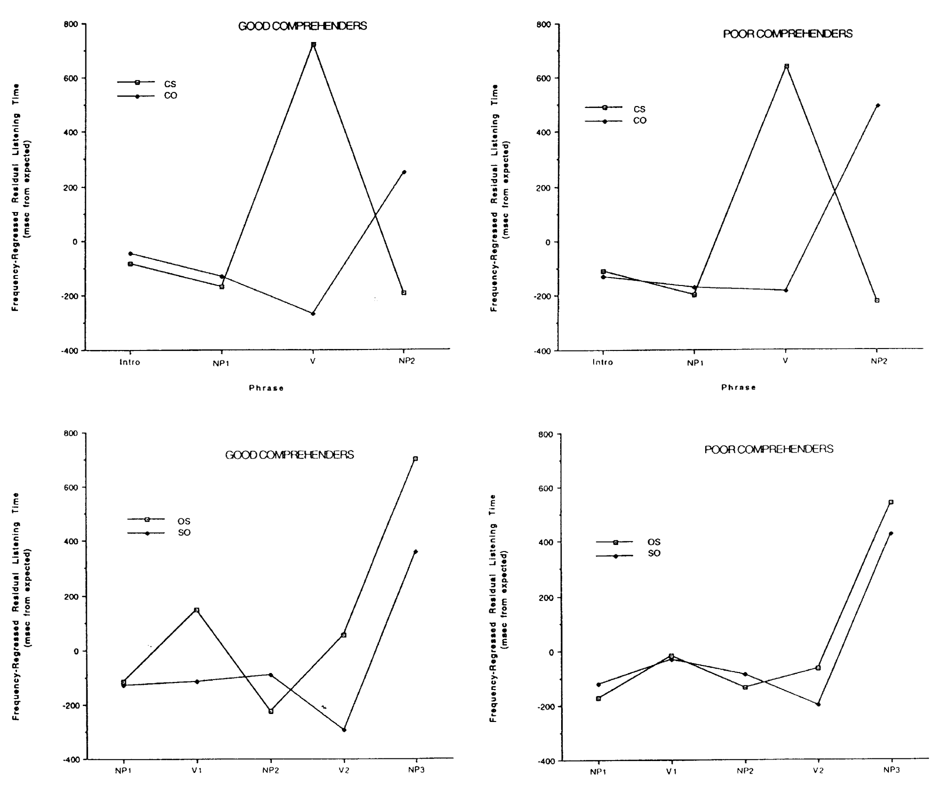

Fig. 5. Residual listening times for good and poor comprehender aphasic patients for plausible sentences to which correct judgments were made. CS: cleft–subject sentences; CO: cleft–object sentences; OS: object–subject sentences; SO: subject–object sentences. Intro: "It was" in CS and CO sentences; NP1: first noun phrase; NP2: second noun phrase; NP3: third noun phrase; V: verb of CS and CO sentences; V1: first verb of OS and SO sentences; V2: second verb of OS and SO sentences.

overall longer residual listening times for CO than for CS sentences. The phrase and sentence type interaction was due to longer residual listening times for NP2 in CS than in CO sentences, for V in CO than in CS sentences, and for V in CO than NP2 in CS sentences. The three-way interaction of group, sentence type and phrase was not significant.

For the OS and SO sentences, there was a interaction of group, sentence type and phrase ($F1(4, 104) = 2.0, p = .1; F2(4, 250) = 4.0, p < .01$). For high comprehending patients, Tukey's tests by subjects showed that residual listening times at V1 and NP3 in SO were longer than at V1 and NP3 in OS sentences. Tukey's tests by items also showed that residual listening times at V2 in SO were longer than at V2 in OS sentences and that residual listening times at V1 in SO were longer than at NP2 in OS sentences. There were no differences in residual listening times for these pairs of segments in the two sentence types in low performing subjects.

3.10. Listening times in incorrect responses

Fig. 6 shows residual listening times for the good and poor comprehension aphasic patients for the plausible CO and SO sentences on which they did and did not make errors. These data were analyzed as described above for each group separately.

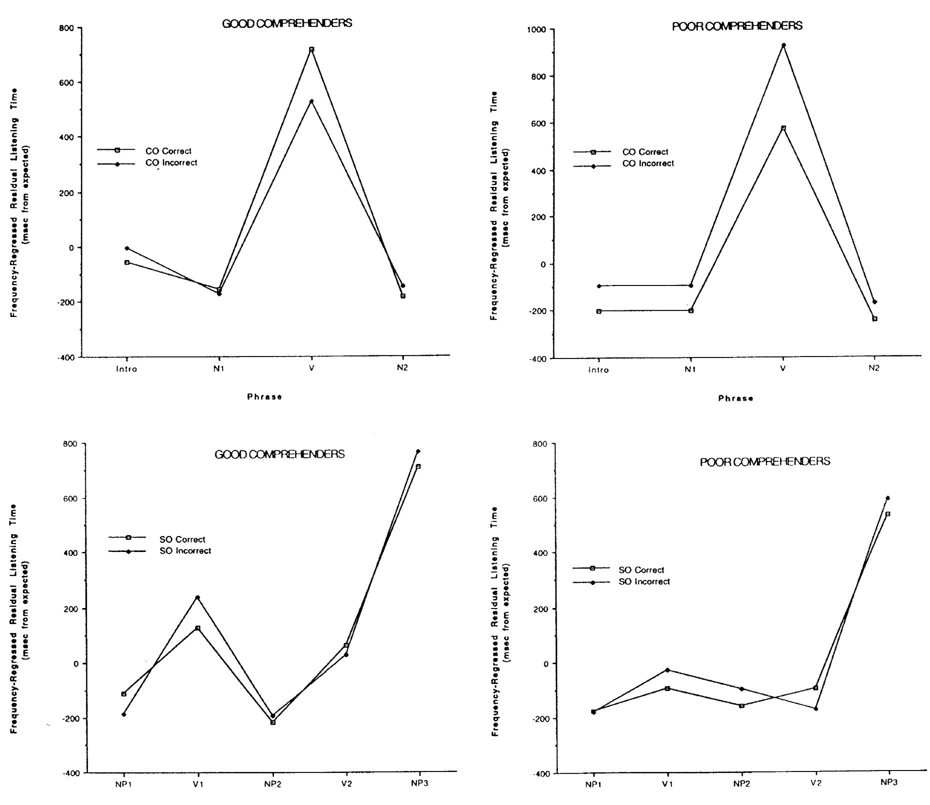

Fig. 6. Residual listening times for good and poor comprehender aphasic patients for plausible cleft–object (CO) and subject–object (SO) sentences to which correct and incorrect judgments were made.

For good performers, there were main effects of phrase in both analyses (for CO, $F1(3,33) = 19.1$, $p < .001$; for SO, $F1(4,44) = 22.1$, $p < .001$). Residual listening times did not differ for any segment as a function of accuracy of judgment.

For poor performers, there was an interaction between phrase and accuracy in CO sentences ($F1(3,45) = 3.7$, $p < .01$). Residual listening times were longer on V in sentences that were incorrectly judged to be implausible. For SO sentences, there was an effect of phrase ($F1(4,60) = 13.0$, $p < .001$). For SO sentences, residual listening times did not differ for any segment as a function of accuracy of judgment.

3.11. Correlational analyses

Correlational analyses were performed to further investigate the relationship between end-of-sentence performance and on-line processing. These analyses were performed between accuracy and reaction time differences in making plausibility judgments in the plausible more and less complex sentences and the difference in residual listening times in critical segments in the plausible more and less complex sentences to which subjects responded correctly. Results are shown in Table 4.

Two of these correlations were significant. The first was a positive correlation between the difference in plausibility judgment reaction times to cleft–object (CO) and cleft–subject (CS) sentences and the difference in residual listening times to the verb (V) of CO sentences and the second noun phrase (NP2) in CS sentences. The second was a negative correlation between the difference in plausibility judgment reaction times to subject–object (SO) and object–subject (OS) sentences and the difference in residual listening times to the first verb (V1) of SO sentences and the second noun phrase (NP2) in object–subject sentences. These correlational results are consistent with the results of the analysis of variance, as discussed below.

3.12. Discussion of results for good- and poor-performing aphasics

To summarize the results of this comparison, patients selected to be good and poor comprehenders on the basis of a pretest differed as expected in the end-of-sentence measures. A′s were lower for the poor comprehenders than for the good comprehenders on the syntactically more complex CO and SO sentences. RTs for plausibility judgments were longer for the poor comprehenders than for the good comprehenders, and RTs were longer on the SO compared to the OS sentence in the poor comprehenders but not in the good comprehenders. These results indicate that

Table 4
Correlations between end-of-sentence and on-line performance measures

On-line measure	End of sentence measure	
	A′ CS-A′ CO	RT CO-RT CS
CO V-CS V	.21	.20
CO V-CS NP2	.04	.41*
	A′ OS-A′ SO	RT SO-RT OS
SO V1-OS V1	.13	−.32
SO V1-OS NP2	.19	−.45*
SO V2-OS V2	−.31	−.19

CS: cleft–subject sentences; CO: cleft–object sentences; OS: object–subject sentences; SO: subject–object sentences. NP1: first noun phrase; NP2: second noun phrase; NP3: third noun phrase; V: verb of CS and CO sentences; V1: first verb of OS and SO sentences; V2: second verb of OS and SO sentences.

* $p < .05$.

subjects selected to be poor comprehenders on the basis of performance on a separate test were more impaired than the good comprehenders in this plausibility judgment task and that their impairment affected their performance on the more complex sentences to a greater degree than on the simpler sentences.

Residual listening times for individual segments differed in patients who performed well and those who did not. In the good comprehending patients, the pattern of residual listening times was similar to that seen in the previous analysis: residual listening times were elevated at the ends of clauses and at points of syntactic complexity. The poor comprehending patients did not show these on-line effects. The poor comprehending subjects performed differently on-line in the CO–CS sentences and in the SO–OS sentences.

For poor comprehending patients, there was no increase in residual listening times at any phrase in SO compared to OS sentences. These differences in off-line and on-line performances were reflected in correlational analyses that showed that, for the entire aphasic population, longer reaction time differences between the SO and OS sentences on the plausibility judgments were associated with smaller differences in residual listening times to V1 of plausible SO compared to NP2 of plausible OS sentences. The absence of an increase in residual listening times at any phrase in SO compared to OS sentences in the low performing patients suggests that these patients did not assign the syntactic structure of the more difficult SO sentences on-line.

The picture for the CO sentences in the low-comprehending patients is more complex. Low comprehending patients had longer residual listening times on V in CO than on V or NP2 in CS sentences to which they responded correctly. Correlational analyses showed that, for the group of aphasics, as the difference in RTs on plausibility judgments between CO and CS sentences increased, the difference in residual listening times between V of plausible CO and NP2 of plausible CS sentences also increased. In addition, the poor comprehending subjects showed longer residual listening times on V in CO sentences to which they responded erroneously than on V in CO sentences to which they responded correctly. Both these patterns contrast with their performances on SO sentences, where they showed no effects of syntactic structure or accuracy on listening times. The residual listening times in CO sentences suggests these low comprehending patients attempted to assign the structure and meaning of these sentences on-line. When they succeeded, their residual listening times showed the normal pattern (and one measure of off-line performance was correlated with one measure of on-line performance for these sentences in the aphasic group as a whole). When they made errors, they spent more time trying to structure the sentence syntactically and/or semantically and allocated additional time to processing the most demanding phrase of the sentence.

The different patterns of on-line performance of the poor comprehending subjects on SO and CO might reflect two different types of processing deficits. Alternatively, they might reflect two different ways of coping with a single processing deficit—a failure to attempt to structure a sentence, seen in SO sentences, and an effort to structure a sentence that led to normal residual listening times when it was successful and to longer residual listening times on the more demanding portion of the sentence when unsuccessful, seen in the CO sentences.

In summary, these results indicate that off-line performance patterns are related to on-line processing. The data suggest the existence of at least two types of abnormal on-line processing, both associated with impaired performance on off-line tasks.

3.13. Broca's and fluent aphasic patients

Patients were classified as Broca's or Fluent aphasics based on a variety of factors: clinical aphasiological diagnosis made by a speech–language pathologist at the referring institution, performance on standard aphasia batteries such as the BDAE, performance on an experimental sentence completion task designed to identify agrammatic speech (Goodglass, Christiansen, & Gallagher, 1993), and lesion location. Broca's aphasics had lesions involving Broca's area; fluent aphasics had lesions involving temporal and/or parietal structures and sparing Broca's area. These factors served to identify nine patients as Broca's aphasics and five as fluent aphasics, as indicated in Table 1.

3.14. Plausibility judgment results

Reaction times and A′ scores on the plausibility judgment task in Fig. 7. These data were analyzed using the approaches utilized above.

Analysis of A′ scores showed a main effect of Sentence Type ($F1(3,36) = 17.3$, $p < .001$). A′s were lower for SO than for OS sentences and for CO than for CS sentences. There was an interaction between Group and Sentence Type ($F1(3,36) = 3.7$, $p < .05$). A′s were lower for SO than for OS sentences for both groups, and lower for CO than for CS sentences in the fluent aphasics only. A′s did not differ between groups for any sentence type.

The reaction time data from the plausibility judgments showed main effects of Group ($F1(1,12) = 5.0$, $p < .05$; $F2(3,199) = 70.8$, $p < .001$) and Sentence Type ($F1(3,36) = 3.2$, $p < .05$; $F2(3,199) = 2.9$, $p < .05$). Broca's aphasics responded more slowly than fluent aphasics to all sentence types. Reaction times were longer for SO sentences than for OS sentences.

3.15. Self-paced listening results

Residual listening times for plausible sentences judged as plausible by the Broca's and fluent aphasics are presented in Fig. 8. These data were analyzed using the approaches utilized above.

For the CS and CO sentences, there was a significant interaction between sentence type and phrase ($F1(3,36) = 20.6$, $p < .001$; $F2(3,200) = 139.5$, $p < .001$). Tukey's test by subjects showed that residual listening times were longer at V of CO than at V of CS sentences and at NP2 of CS than at NP2 of CO sentences. Tukey's test by items also showed that residual listening times were longer at V of CO than at NP2 of CS sentences. The three-way interaction of group, sentence type and phrase did not approach significance in either analysis.

For the OS and SO sentences, there were significant interactions between sentence type and group ($F1(1,12) = 4.5$, $p < .05$; $F2(4,245) = 2.7$, $p = .1$) and between sentence type and phrase ($F1(4,48) = 2.8$, $p < .05$; $F2(4,245) = 2.1$, $p = .08$). The interaction between sentence type and group was due to residual listening times being longer in fluent than in Broca's aphasics in SO but not in OS sentences. The interaction between sentence type and phrase seen in the analysis by subjects was due to residual listening times being longer at NP3 in SO than in OS sentences.

Inspection of Fig. 5 suggests that residual listening times were longer at V1 of SO sentences than at V1 or NP2 of OS sentences in the fluent but not Broca's aphasics. The three-way interaction of group, sentence type and phrase did not approach significance in the analysis by subjects ($F1(4,48) = .66$, ns)

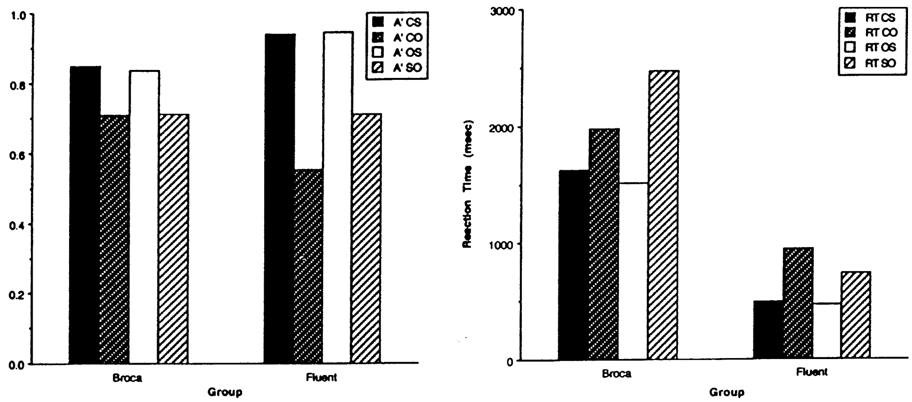

Fig. 7. A′ scores and reaction times (in milliseconds) for plausibility judgments made by for Broca's and fluent aphasic patients.

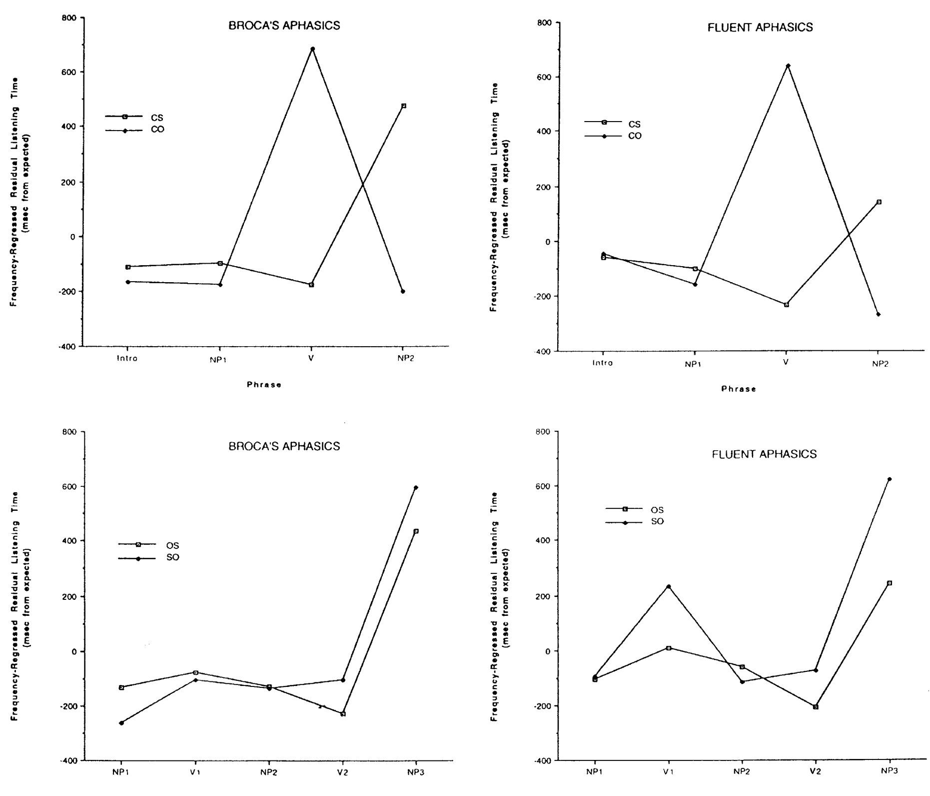

Fig. 8. Residual listening times for Broca's and fluent aphasic patients to which correct judgments were made. CS: cleft–subject sentences; CO: cleft–object sentences; OS: object–subject sentences; SO: subject–object sentences. Intro: "It was" in CS and CO sentences; NP1: first noun phrase; NP2: second noun phrase; NP3: third noun phrase; V: verb of CS and CO sentences; V1: first verb of OS and SO sentences; V2: second verb of OS and SO sentences.

but it was significant in the analysis by items ($F2(4, 245) = 3.0$, $p = .02$). Residual listening times were in fact longer at V1 of SO sentences than at V1 or NP2 of OS sentences in the fluent but not Broca's aphasics, in Tukey's test by items.

3.16. Listening times in incorrect responses

Fig. 9 shows residual listening times for Broca's and fluent patients for the plausible CO and SO sentences on which they made errors. These data were analyzed as described above. In all analyses, there were effects of phrase (for Broca's: for CO, $F1(3,24) = 14.0$, $p < .001$; for SO, $F1(4,32) = 8.4$, $p < .001$; for fluents: for CO, $F1(3,12) = 37.3$, $p < .001$; for SO, $F1(4,16) = 2.6$, $p = .07$). Residual listening times did not differ for any segment in either group in either sentence type as a function of accuracy of judgment.

3.17. Discussion of results for Broca's and fluent aphasics

Broca's aphasics took longer than fluent aphasics to make plausibility judgments, but the fluent aphasics showed a more pronounced effect of syntactic structure in their judgment accuracy, with less accurate performance on the cleft–object than on cleft–subject sentences, which was not found in Broca's aphasics. Analysis of residual listening times showed that both groups of subjects had longer residual listening times at V of CO sentences than at V or NP2 of CS sentences, and suggests that fluent, but not Broca's, aphasics had longer residual listening times at V1 of SO sentences than at V1 or NP2 of OS sentences. This difference between the groups would be consistent with the view that the fluent aphasics assigned syntactic structure in the SO sentences, while the Broca's aphasics did not.

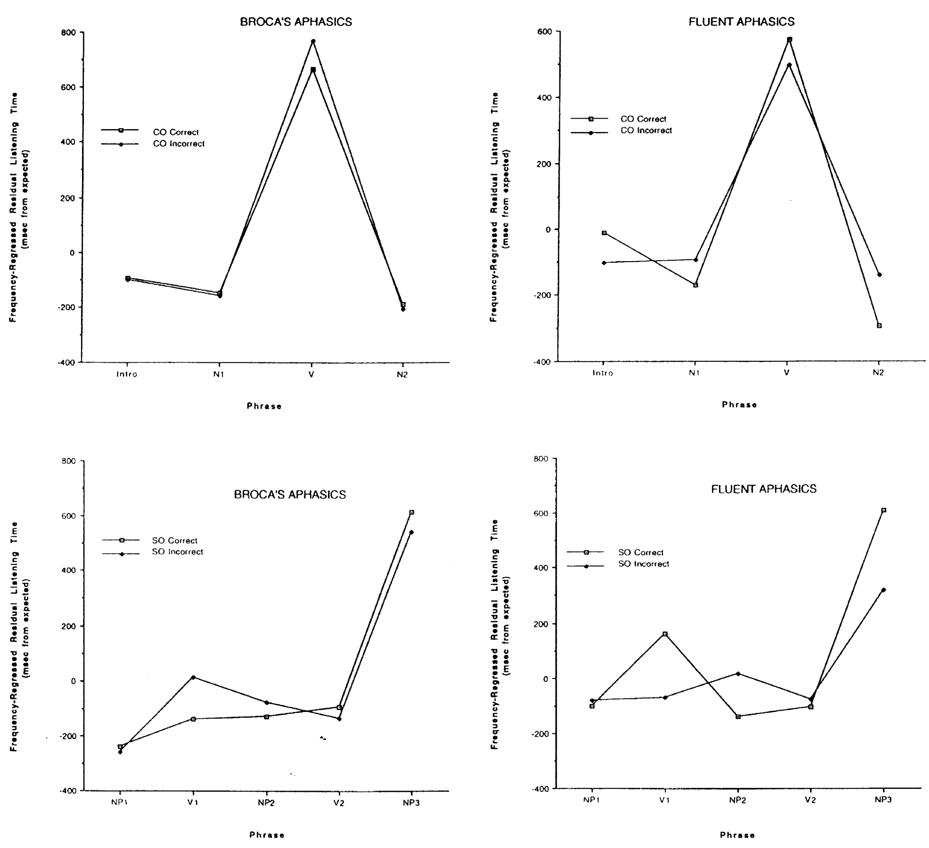

Fig. 9. Residual listening times for Broca's and fluent aphasic patients for plausible cleft–object (CO) and subject–object (SO) sentences to which correct and incorrect judgments were made.

However, these results do not provide strong support for the view that these two clinical types of aphasics differ in their on-line syntactic processing because they were not reliable by subjects.

4. General discussion

This study allows a direct comparison between an off-line measure of sentence comprehension—performance (accuracy and RT) in a plausibility judgment task—and an on-line measure of sentence processing—residual listening times in a self-paced listening task—in the same materials in a group of aphasic patients. We begin our discussion of the results by briefly reviewing the results in the normal subjects in relation to factors that determine residual listening times.

The prolongation of residual listening times at the embedded verb of object-relativized clauses compared to both the embedded verb of subject relativized clauses and to the last word in subject relativized clauses can be attributed to constructing syntactic and propositional representations. At the syntactic level, there is a "storage" cost to maintaining the prediction that there will be an embedded verb in a relative clause, which is higher in an object-relative clause because it has to be maintained over the subject of such a clause (Gibson, 1998). At the propositional semantic level, two thematic roles can be assigned at the verb of object-relativized sentences whereas only one thematic role can be assigned in subject-relativized clauses, leading to higher "integration" costs (Gibson, 1998).

Additional evidence for syntactic effects on residual listening times is the prolongation of residual listening times at the main verb (V2) in SO compared to OS sentences. This can also be related to memory costs. In SO sentences, V2 is related to the subject of the main clause (NP1) across the relative clause, whereas in OS sentences, it is related to the immediately adjacent head of the relative clause (NP2), resulting in a larger memory load at that point in SO than in OS sentences. Some of the increase in residual listening times at V2 in SO sentences may also be due to "spill-over" effects from previous words (King & Just, 1991).

Discourse factors may also play a role in determining residual listening times. The processing load at the verb of cleft–object sentences is high in part because of a discrepancy between the discourse implications of clefting and object-relativization. A cleft sentence explicitly places the clefted noun in the position of discourse focus, and the focus of a discourse usually is the subject of a sentence and the agent of a verb (Kintsch & van Dijk, 1978). The mismatch between the discourse and thematic roles played by the clefted noun in a cleft–object sentence can first be recognized at the embedded verb. The need to reconcile these discrepant properties leads to an increase in processing load, and therefore in listening times, at the verbs of CO compared to the verb or final noun of CS sentences.

Two other processing mechanisms may have influenced residual listening times. One are wrap-up effects that some investigators have suggested occur at the ends of sentences and clauses, such as reactivation of all noun phrases in a sentence (Balogh et al., 1998). This could partly underlie the longer residual listening times on sentence-final words. Longer sentence- (but not clause-) final residual listening times may also reflect subjects beginning the decision-making process regarding the plausibility of the sentence before pushing the response button for the sentence-final segment.

Finally, there were longer end-of-sentence residual listening times in CO compared to CS and in SO compared to OS sentences. The increase in residual listening times for NP3 in SO compared to OS sentences is particularly interesting because the demands on construction of syntactic, semantic, and discourse-level representations are essentially equal at that point in the two sentence types and NP3 is well past the point at which spill-over effects would be expected. Judgment times were longer in SO than in OS sentences, so the longer residual listening times on NP3 in SO sentences are not likely to have resulted from greater intrusion of decision-making into residual listening times for this segment in this sentence type. A process that may occur at the end of SO and CO sentences that would prolong residual listening times is a partial re-analysis of the presented sentence to adjudicate between different sets of thematic roles that are considered as possible meanings of the sentence. There is very strong evidence from the effects of the plausibility of the fit of noun phrases to thematic roles around verbs on self-paced reading and eye tracking times that subjects consider all possible thematic role assignments of all nouns to all verbs that occur within a clause (Frazier, 1987; MacDonald, Pearlmutter, & Siedenberg, 1994; Trueswell & Tanenhaus, 1994; for application to aphasia, see Caplan & Hildebrandt, 1988; Saffran, Schwartz, & Linebarger, 1997; Waters, Caplan, & Hildebrandt, 1991). The thematic roles assigned in this way must be rejected in favor of those licensed by the syntax. Some parsing models postulate a delayed re-analysis mechanism that is involved in this process (e.g., Frazier, 1987), and regressive eye movements provide empirical evidence for a review of parts of a sentence when processing is demanding. A mechanism that partially recomputes thematic roles on the basis of syntactic structure is likely to occur at the end of a sentence or a clause. Since this process is more resource-demanding in more complex sentences, this mechanism would account for longer residual listening times on the last word of syntactically more complex sentences.

The aphasic patients who performed well on the judgment task and who showed evidence of understanding syntactically complex sentences had the same pattern of residual listening times on the self-paced listening task as controls. This indicates that their on-line processing was affected by the same factors as normals'. These high-performing patients thus appear to have intact processing mechanisms needed to construct syntactic, propositional, and discourse representations.

In contrast, aphasic patients who performed poorly on the judgment task, and who showed effects of syntactic structure on this task, showed evidence of on-line processing impairments. Low performing patients appear to have attempted to structure CO sentences. Correct responses were associated with normal prolongations of residual listening times on the portion of the sentence with the highest processing demand—the verb. This suggests that these sentences were processed normally at times. Errors on the plausibility judgment were associated with prolonged residual listening times on the verb of CO sentences. This suggests that these subjects attempted to structure these sentences, but failed, and, when they failed, they devoted more time than normal to the attempt before abandoning the effort. In contrast to their apparent treatment of CO sentences, the low performing patients seem not to be trying to assign syntactic structure to SO sentences, as judged by their self-paced listening times.

This analysis has implications for several aspects of aphasic performance.

One is the role that guessing plays in generating correct and incorrect responses. It is widely assumed that a patient who responds correctly on a task at a rate no greater than chance is guessing among available responses. In some analyses, it is argued that patients generate two or more interpretations for a sentence and choose randomly among them to respond to task demands (Grodzinsky, 2000). In other analyses, it is argued that patients do not generate any interpretations of a sentence and choose randomly between possible responses (Caplan et al., 1985). The results of this study suggest that random performance does not always reflect guessing. As a group, the low performing patients responded at chance on CO sentences. Examination of individual cases showed that this was the case for 11 of the 16 patients in this group. The on-line performance of these subjects on these sentences provides strong evidence that their responses did not result from guesses, but rather from correct interpretations of some sentences and incorrect interpretations of others. The fact that these patients showed the same pattern of self-paced listening as normals when they responded correctly to these sentences argues that they processed these sentences in the same fashion as the normal subjects. The alternative to this view would have to be the claim that the patients failed to process these sentences normally and that they abandoned the effort to process the most demanding phrase in a sentence in exactly the same amount of time that it took the controls to correctly process this phrase. We would argue that, while possible, this is not likely.

The results also bear on the causes of patients' failure to assign syntactic structure. One suggestion that has been made is that some aphasics have a "mapping" problem that allows them to assign syntactic structures but prohibits them from using these structures to determine aspects of sentence meaning (Linebarger, 1995; Linebarger et al., 1983a, 1983b). Though this may be the correct analysis of some patients' deficits, the data do not support a mapping deficit in the patients presented here. A deficit in mapping is not compatible with the normal on-line processing seen in either the high performing subjects or in low performing patients with correctly interpreted CO sentences, because semantic interpretation occurs on-line in conjunction with building syntactic structure and contributes to performance on self-paced tasks (MacDonald et al., 1994; McClelland & St. John, 1989; Trueswell et al., 1994). Therefore, these patients' normal on-line performance presumably reflects the integrity of both structure building and mapping in these patients in these sentences. A mapping deficit would also not be compatible with the failure of low performing patients to show any effects of syntactic structure on residual self-paced listening times in SO sentences, if, as is likely, residual listening times reflect non-semantic syntactic operations to any significant degree.

A second account of a possible on-line deficit is that aphasics may have a pathologically fast decay rate of representations (Haarman & Kolk, 1994). This account predicts that patients will limit the time they allocate to syntactic and semantic processing in an effort to retain more information in memory (Just & Carpenter, 1992), resulting in shorter residual listening times, especially on demanding segments. The fact that patients (including low performing patients for correctly interpreted CO sentences) showed the same degree of on-line sensitivity to local processing load as controls suggests that rapid decay of representations is not the fundamental problem in the patients studied here. As we argued above, a normal pattern of on-line performance would be unlikely to result from aphasics failing to assign structure and meaning in the normal fashion—because of fast decay of representations or for any other reason—and terminating efforts with exactly the same time course as seen in successful assignments by normal subjects. A second problem for this account is that these patients took longer to process the demanding segment of CO sentences they failed to understand. As noted above, pathologically fast decay of representations would be expected to lead to truncation of processing time in favor of retention of representations in memory, not prolongation of processing time.

The deficit in these patients appears to lead to an intermittent failure to successfully structure and interpret CO sentences and a failure to attempt to structure SO sentences. The intermittent failure to structure and interpret CO sentences could result from an intermittent failure to accomplish an operation related to structuring the relative clause, to relate the head noun of the relative clause to its position around the verb of the clause, or to accomplish some other operation. The failure to attempt to structure SO sentences may be due to the fact that high processing costs, which exceed those usually available to the patient, can be predicted in these sentences before the embedded verb is reached. However, it must be said that the exact nature of the impairment in these patients and the reasons for differences in the abnormal residual listening times seen in CO and SO sentences are not fully discernable based on this one experimental task.

One final pattern of performance has implications for the possible locus of a deficit in processing. This pattern consisted of poor off-line, end-of sentence, performance on the syntactically more complex sentences and normal effects of syntactic structure on on-line measures. The clearest example of this pattern is the performance of the fluent aphasic patients in CO and CS sentences: their accuracy was very low for CO sentences (significantly lower than for CS sentences) but they showed normal prolongation of residual listening times on the verbs of CO compared to CS sentences. This pattern strongly suggests that these patients' errors arose after syntactic structure had been assigned. This could reflect a disturbance of a review process, which we suggested might also account for longer residual listening times for the final words of more complex sentences.

Finally, the results of this study are relevant to the hypothesis that there is a difference in on-line syntactic processing in Broca's and fluent aphasics. The results in the SO/OS sentences provide evidence in favor of Grodzinsky's (2000) claim that on-line processing associated with structuring and interpreting relative clauses differs in Broca's aphasics and fluent aphasics. Broca's aphasics performed similarly to other poor-comprehending patients, and showed no effects of syntactic structure on on-line measures in SO sentences. Fluent aphasics showed poor off-line but normal on-line performance. As noted above, this may reflect difficulties they have in a review stage of processing. The preservation of on-line, implicit processing of syntactic structure and disturbed performance in off-line measures of syntactic processing in fluent patients is reminiscent of their normal performance on implicit tests of lexical semantic processing (priming) and disturbances in conscious, explicit tests of this function (word-picture matching) in some studies (Blumstein et al., 1982; Milberg & Blumstein, 1981), though the mechanisms responsible for the implicit/explicit processing dissociation may differ at the lexical and syntactic levels. The data are only suggestive of differences between Broca's and fluent aphasics in on-line syntactic processing, however, because the theoretically critical interaction was only significant in the analysis by items. Studies with larger numbers of patients in these two groups are needed to further explore this question.

In summary, this study provides evidence regarding on-line syntactic processing in aphasic patients. It documents the existence of aphasic patients whose syntactic processing is unaffected, as far as can be judged from both on-line and off-line measurements. In other patients, it documents a correspondence between disorders of off-line performance in syntactic comprehension and an on-line measure of syntactic, propositional and discourse-level processing. This correspondence suggests that on-line processing deficits underlie off-line impairments in comprehension of complex syntactic structures in many aphasic patients. The deficits in on-line syntactic processing documented in this study are not likely to be due to a "mapping" impairment or to overly rapid decay of syntactic structures in aphasic patients. They may represent failures to assign certain aspects of syntactic representations on-line, and/or to be able to utilize a review process to adjudicate between pragmatically and syntactically derived sentence meaning. The results of this study suggest that there may be a difference between Broca's and fluent aphasics with respect to the nature of their deficit(s). The combined use of on-line and off-line measures with the same sentences in the same patients appears to have promise as a way to identify different sentence-level processing impairments that occur in aphasia.

References

Balogh, J., Zurif, E. B., Prather, P., Swinney, D., & Finkel, L. (1998). Gap filling and end of sentence effects in real-time language processing: Implications for modeling sentence comprehension in aphasia. *Brain and Language, 61*, 169–182.

Berndt, R., Mitchum, C., & Haendiges, A. (1996). Comprehension of reversible sentences in "agrammatism": A meta-analysis. *Cognition, 58*, 289–308.

Blumstein, S. E., & Milberg, W., et al. (1982). Semantic processing in aphasia: Evidence from an auditory lexical decision task. *Brain and Language, 17*, 301–315.

Caplan, D. (1992). *Language: Structure, processing, and disorders*. Cambridge, MA: MIT Press.

Caplan, D., & Hildebrandt, N. (1988). *Disorders of syntactic comprehension*. Cambridge, MA: MIT Press/ Bradford Books.

Caplan, D., Baker, C., & Dehaut, F. (1985). Syntactic determinants of sentence comprehension in aphasia. *Cognition, 21*, 117–175.

Caplan, D., Hildebrandt, N., & Waters, G. S. (1994). Interaction of verb selectional restrictions, noun animacy and syntactic processing. *Language and Cognitive Processes, 9*(4), 549–585.

Caplan, D., Hildebrandt, N., & Makris, N. (1996). Location of lesions in stroke patients with deficits in syntactic processing in sentence comprehension. *Brain, 119*, 933–949.

Caplan, D., Waters, G., & Hildebrandt, N. (1997). Determinants of sentence comprehension in aphasic patients in sentence–picture matching tasks. *Journal of Speech and Hearing Research, 40*, 542–555.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, N. (1986). *Knowledge of language*. New York: Praeger.

Chomsky, N. (1994). *The minimalist program*. Boston: MIT Press.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research: Methods, Instruments, and Computers, 25*, 257–271.

Dunn, J. (1994). *SoundEdit. Version 2*. Micromedia Inc./Apple Inc..

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language, 25*, 348–368.

Ferreira, F., Henderson, J. M., Anes, M. D., Weeks, P. A., Jr., & McFarlane, D. K. (1996). Effects of lexical frequency and syntactic complexity in spoken language comprehension: Evidence from the auditory moving window technique. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 324–335.

Francis, W. N., & Kucera, H. (1982). *Frequency analysis of English usage*. Boston: Houghton Mifflin.

Frazier, L. (1987). Sentence processing: A tutorial review. In *Attention and performance XII: The psychology of reading* (pp. 559–586). Coltheart: Erlbaum.

Friederici, A., & Kilborn, K. (1989). Temporal constraints on language processing: Syntactic priming in Broca's aphasia. *Journal of Cognitive Neuroscience, 1*, 262.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition, 68*, 1–76.

Goodglass, H., Christiansen, J. A., & Gallagher, R. E. (1993). Comparison of morphology and syntax in free narrative and structured tests: Fluent vs non-fluent aphasics. *Cortex, 29*, 377–407.

Grodzinsky, Y. (1990). *Theoretical perspectives on language deficits*. Cambridge, MA: MIT Press.

Grodzinsky, Y. (1995). A restrictive theory of agrammatic comprehension. *Brain and Language, 50*, 27–51.

Grodzinsky, Y. (2000). The neurology of syntax: Language use without Broca'a area. *Behavioral and Brain Sciences, 23*, 47–117.

Haarman, J., & Kolk, H. (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.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122–149.

Kilborn, K., & Friederici, A. (1994). Cognitive penetrability of syntactic priming in Broca's aphasia. *Neuropsychology, 8*, 81–90.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language, 30*, 580–602.

Kintsch, W., & van Dijk, T. A. (1978). Towards a model of text comprehension and production. *Psychological Review, 85*, 363–395.

Linebarger, M. C. (1995). Agrammatism as evidence about grammar. *Brain and Language, 50*, 52–91.

Linebarger, M. C., Schwartz, M. F., & Saffran, E. M. (1983a). Sensitivity to grammatical structure in so-called agrammatic aphasics. *Cognition, 13*, 361–392.
Linebarger, M. C., Schwartz, M. F., & Saffran, E. M. (1983b). Syntactic processing in agrammatism: A reply to Zurif and Grodzinsky. *Cognition, 15*, 207–214.
MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review, 101*, 676–703.
McClelland, J. L., & St. John, M., et al. (1989). Sentence comprehension: A parallel distributed processing approach. *Language and Cognitive Processes, 4*, 287–336.
McKoon, G., & Ratcliff, R. (1994). Sentential context and on-line lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1239–1243.
McKoon, G., Ratcliff, R., & Ward, G. (1994). Testing theories of language processing: An empirical investigation of the on-line lexical decision task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1219–1228.
Milberg, W., & Blumstein, S. E. (1981). Lexical decision and aphasia: Evidence for semantic processing. *Brain and Language, 14*, 371–385.
Nicol, J., Fodor, J., & Swinney, D. (1994). Using cross-modal lexical decision tasks to investigate sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1229–1238.
Saffran, E. M., Schwartz, M. F., & Linebarger, M. C. (1997). Semantic influences on thematic role assignment: Evidence from normals and aphasics. *Brain and Language, 62*, 255–297.
Shankweiler, D., Crain, S., Gorrell, P., & Tuller, B. (1989). Reception of language in Broca's aphasia. *Language and Cognitive Processes, 4*, 1–33.
Swinney, D., & Zurif, E. (1995). Syntactic processing in aphasia. *Brain and Language, 50*, 225–239.
Swinney, D., Zurif, E., Prather, P., & Love, T. (1996). Neurological distribution of processing resources underlying language comprehension. *Journal of Cognitive Neuroscience, 8*, 174–184.
Trueswell, J. C., & Tanenhaus, M. K., et al. (1994). Semantic influence on syntactic processing: Use of thematic information in syntactic disambiguation. *Journal of Memory and Language, 33*, 285–318.
Tyler, L. (1985). Real-time comprehension processes in agrammatism: A case study. *Brain and Language, 26*, 259–275.
Waters, G. S., & Caplan, D. (2001). Age, working memory, and on-line syntactic processing in sentence comprehension. *Psychology and Aging, 16*, 128–144.
Waters, G. S., Caplan, D., & Hildebrandt, N. (1991). On the structure of verbal short-term memory and its functional role in sentence comprehension: Evidence from neuropsychology. *Cognitive Neuropsychology, 8*(2), 81–126.
Waters, G. S., Caplan, D., & Rochon, E. (1995). Processing capacity and sentence comprehension in patients with Alzheimer's disease. *Cognitive Neuropsychology, 12*(1), 1–30.
Zurif, E., Swinney, D., Prather, P., Solomon, J., & Bushell, C. (1993). An on-line analysis of syntactic processing in Broca's and Wernicke's aphasia. *Brain and Language, 45*, 448–464.