Published in final edited form as:
Brain Lang. 2013 October ; 127(1): 21–33. doi:10.1016/j.bandl.2013.07.007.

Dissociations and Associations of Performance in Syntactic Comprehension in Aphasia and their Implications for the Nature of Aphasic Deficits

David Caplan, **Jennifer Michaud**, and **Rebecca Hufford**
Neuropsychology Laboratory, Department of Neurology, Massachusetts General Hospital

Abstract

Sixty one pwa were tested on syntactic comprehension in three tasks: sentence-picture matching, sentence-picture matching with auditory moving window presentation, and object manipulation. There were significant correlations of performances on sentences across tasks. First factors in unrotated factor analyses accounted for most of the variance on which all sentence types loaded in each task. Dissociations in performance between sentence types that differed minimally in their syntactic structures were not consistent across tasks. These results replicate previous results with smaller samples and provide important validation of basic aspects of aphasic performance in this area of language processing. They point to the role of a reduction in processing resources and of the interaction of task demands and parsing and interpretive abilities in the genesis of patient performance.

Introduction

Studies of performance of people with aphasia (pwa) on tasks requiring syntactically based comprehension have reported both dissociations and associations of performance on different sentence types, and both these types of data have been used to construct models of aphasic deficits in the ability to construct syntactic structure and use it to determine aspects of sentence meaning. This study provides new data regarding aphasic dissociations and associations of performance on different sentence types on tasks requiring syntactically based comprehension. We introduce the present study with a brief review of previous literature reporting dissociations and associations and the role of these patterns of performance in using aphasic data to develop models of sentence structure and processing.

The first study to document abnormalities in comprehension of semantically reversible sentences was Caramazza and Zurif (1976). These authors built upon dissociations of performances to develop a model of sentence comprehension and of aphasic disorders. Caramazza and Zurif (1976) argued that the combination of three findings -- a) chance performance on semantically reversible syntactically complex sentences, b) good performance on semantically irreversible sentences with the same structures, and c) good performance on semantically reversible sentences with simple syntactic structures -- in some

Address Correspondence to: David Caplan, M.D., Ph.D., Neuropsychology Laboratory, 175 Cambridge Street, Suite 340, Fruit Street, Boston, MA, 02114, Tel: 617-726-3274, FAX: 617-724-7836, dcaplan@partners.org.

pwa pointed to a deficit in the ability to apply “algorithmic” analyses of syntactic structure and a preserved ability to apply “heuristic” analyses to comprehend sentences. By “algothmic processes,” Caramazza and Zurif (1976) meant parsing and interpretation – the assignment of syntactic structure by a set of rules that map input onto structures specified by a grammar, and the use of the structure that had been assigned to determine aspects of sentence meaning such as thematic roles. Subsequent work further specified details of “algorithmic” processing.

The best known work in this area is that of Grodzinky (1986; 1995a, b; 2000), who argued that dissociation of performances of patients with agrammatic Broca’s aphasia on sentence-picture matching tasks indicated that these patients had lost the ability to represent or co-index abstract representations of noun phrases (NPs) that remained in the original positions from which they had been “moved” in a syntactic “derivation” according to Chomsky’s (1975, 1981, 1995) model of syntax (the “trace deletion hypothesis”). Grodzinsky (1986; 1995a, b; 2000) argued that this pattern provided empirical support for Chomsky’s model of syntactic structure over models that did not postulate “traces.”

Several researchers have challenged aspects of Grodzinky’s analysis of the deficits in agrammatism from a within-theory perspective, arguing that “traces” are not lost but subject to other disruptions (e.g., the “double dependency hypothesis” (Mauner, 1995; Mauner et al, 1993)). These “theory-internal” critiques deal with the details of the disruption of processing of “traces” in these patients. Other researchers have argued that Grodzinsky was wrong to attribute a deficit affecting traces to all agrammatic Broca’s aphasics or to only agrammatic Broca’s aphasics (Caramazza et al, 2000). Like the “theory-internal” critiques, these “neurological” or “clinical” critiques do not question the assumption that the dissociation of performance that is sometimes found in agrammatic Broca’s aphasics in sentence-picture matching tasks may provide evidence for a model of syntax that includes “traces” or equivalent structures.

Other researchers have reported that performance on different sentence types tends to be highly correlated within individual patients. Evidence for this was summarized by Caplan (2012) and includes factor analyses of performance, which have routinely yielded first factors in unrotated analyses on which all sentence types in a study load significantly, which account for the great majority of variance in accuracy (up to 85%) (Caplan et al, 1996, 2007); Rasch models of aphasic performance (Gutman et al, 2010, 2011); and simulations of aphasic performance (Miyake et al, 1994, 1995). Coupled with the finding that more complex sentences are more often misunderstood than simpler sentences (Caplan et al, 1985, 1996, 2007), associations of performance have led to various “processing resource reduction” models of aphasic syntactic comprehension (Caplan, 2012: Miyake et al, 1994, 1995; Haarman and Kolk, 1991a, b; McNeil et al, 1991; see below for discussion).

One issue regarding both these types of models is the quality of the data upon which they are based. Many studies have relied on average data from small groups of pwa with a particular clinical syndrome (e.g., Broca’s aphasia or agrammatism) as the basis for developing a model. This practice has been criticized for obscuring potentially disconfirming cases (see Caramazza et al (2000), Drai and Grodzinksy (1999) for discussion). This criticism is somewhat muted by the presence of reports that include single case data (Caplan and Hildebrandt, 1988). However, other issues are less clearly resolved. One is that very few studies have reported the reliability of performance within individual pwa. The second is that a widespread practice is to test pwa on one task, assuming that the result provides evidence regarding the state of parsing/interpretive functioning. Dissociations of performance that are the basis for specific deficit analyses have most often been reported in sentence-picture matching, but have at times been based on performance in object manipulation (Caplan and Hildebrandt, 1988). There is, however, evidence that pwa perform differently on the same sentence type in different tasks.

Cupples and Inglis (1993) reported poor performance on passive sentences in a sentence-picture matching (SPM) task but good performance in an enactment (or "object manipulation (OM)") task in one patient. Caplan et al (1997) found that performance across SPM and OM tasks correlated at about the $r = .6$ level for ten different sentence types in 17 aphasic patients – a level of correlation that indicates that many patients differed in their accuracy relative to one another on a given sentence type in the two tasks. Caplan et al (2006, 2007) studied 42 patients using SPM and OM. Participants were presented ten examples of each sentence type in each of six baseline/experimental pairs. Caplan et al (2006) found that instances of good (i.e., above chance or normal) performance on a baseline sentence and poor performance on the corresponding experimental sentence in individual patients almost always occurred in only one task. The fact that patients demonstrated good performance on an experimental sentence type on one task indicates that their deficit could not have been a failure of the comprehension system to assign the structure of that type of sentence or to use that structure to determine sentence meaning.

There is also a question about what the existence of both dissociations and associations of performance on different sentence types implies about the nature of aphasic deficits. Although dissociations have been interpreted as evidence for deficits affecting specific structures or operations, most dissociations are "one-way" in the sense that they always affect a sentence type that has a particular element and not a sentence type that does not. Such one-way dissociations could be due to resource reductions that affect the ability to process the more demanding of two sentence types. Caplan and Hildebrandt (1988) addressed this question by documenting a set of different dissociations in different pwa, such that different syntactic elements (trace, PRO, reflexives, pronouns) were affected and spared in different cases. Employing the logic of double dissociations, they argued that, when one element but not a second was affected in one pwa and the opposite pattern was found in a second pwa, at least one of these performances must be due to a specific deficit. However, to our knowledge, this is the only study that documented such patterns, and none of the "double dissociations" were completely "clean;" in all, performance was mildly abnormal in at least some sentence types that contained the second ("unaffected") structure.

Conversely, associations of performance and complexity effects could be due to multiple specific deficits (Martin, 1995). Caplan and Hildebrandt (1988) also addressed this question, by documenting performances of individual pwa that consisted of poor performance on sentences that contained combinations of elements each of which the pwa could comprehend in isolation. Again, however, this is the only study that documents these patterns to our knowledge, and very few cases showed this pattern.

Based on the evidence summarized above, Caplan and Hildebrandt (1988) reached the conclusion that both a general reduction in the ability to construct and interpret syntactic structures and deficits affecting particular psycholinguistic operations or syntactic elements contribute to the impairment in an individual pwa. They argued that the visibility of particular specific deficits depended on the degree of resource reduction, because greater degrees of resource reduction necessarily affected elements that required more processing, such as traces and PRO, so that specific deficits in these types of elements could only been seen in pwa with lower degrees of resource reduction.

Caplan et al (2006) questioned the need for postulating specific deficits. They pointed out that, in a small number of instances (~5%), a pwa showed significantly better performance on a baseline than on an experimental sentence in a task. Assuming that the assignment of structure and meaning in the experimental sentence requires all the operations needed in the baseline sentence of such pairs, this pattern could only be attributed to a factor that affects comprehension randomly (“noise”), which would on some occasions affect performance on more instances of the baseline than the experimental sentence type. Caplan et al (2006) raised the possibility that what appear to be specific deficits are, in fact, the result of resource reduction and this noise factor.

This possibility was tested by Gutman et al (2010) in Rasch models of the data in Caplan et al (2006). The estimate of the ability of a pwa (i.e., the extent of resource reduction) was his/her accuracy on all sentences in one or both tasks used (SPM and OM) and the estimate of sentence demand was the total correct on each sentence in all pwa in one or both tasks. Models with different groupings of sentence type, with and without groupings of pwa, and with and without the factor of task were compared. The best fitting models required the factors of task and patient group but not sentence type. The analyses in Gutman et al (2010) suggest that, when considered within the broader context of performance of many pwa on many sentence types, what appear to be specific deficits in individual pwa are in fact are the result of differential demand made by different sentence types in different tasks and different levels of ability in different pwa, and that end-of-sentence accuracy data do not require postulating specific deficits in syntactically based comprehension in individual pwa. The Gutman et al. analyses may not have had sufficient power to detect an effect of sentence type as well as those of task and group, however.

Gutman et al’s (2010) conclusion that end-of-sentence accuracy data do not require postulating specific deficits in syntactically based comprehension in individual pwa differs from much current thinking about these deficits. Given the importance of this conclusion for the nature of aphasic deficits, it seemed important to determine if the results in Caplan et al (2006) were replicable. This seems even more important given that the results in Caplan et al (2006) were based on the use of only 10 sentences of each type, and repeating the study with a larger number of examples of each sentence type would provide evidence regarding reliability of the data. The present study addresses these issues by presenting data from 61 new pwa, using 20 sentences of each type. It also extends the observations to a third task – a second version of SPM.

Methods

Participants

Sixty-one aphasic patients (mean age 64.2 years (sd: 12.9; range: 28 – 96); mean education, 15.3 years (sd: 3.2; range: 7 – 22); M:F = 44:17) who met entry criteria were tested. Patient were required to be aphasic as determined by a physician or speech-language pathologist, right handed, have a single left hemisphere stroke, be able to perform the tests, and be judged to have adequate single word comprehension to not fail because of lexical semantic disturbances (patients were screened for disturbances of phoneme discrimination, auditory lexical decision and spoken word-picture matching using the Psycholinguistic Assessment of Aphasia, a test developed in our lab (see Caplan, 1992, for description). Patients were tested for their ability to match the nouns in the sentences to pictures, and trained as necessary until they achieved 100% correct matches. Forty-six age and education matched controls (mean age 64.2 years sd: 11.3; (range: 38 – 87); mean education, 16.3 years (sd: 2.3; range: 8 – 28); M:F = 20:26) were also tested.

Participants were tested in sessions lasting 1 – 2 hours, depending on tolerance. Testing on each sentence comprehension task was performed in a separate block from testing on other sentence comprehension tasks. Sentence types were presented a different pseudorandomized order in each task, such that three or more examples of the same sentence type never occurred in succession, and half the participants of each type were given the sentences in the forward order and half given the sentences in the reversed order. Testing on sentence comprehension tasks was interspersed with testing on tasks of short term memory (not reported here). Sentence comprehension and memory tasks were presented in different orders and groups to different participants to make maximally efficient use of the time each participant could tolerate testing.

Materials

The ability to parse and interpret passives, object extracted clefts, object extracted relative clauses, and sentences with reflexives and pronouns was tested by having patients respond to pairs of sentences in which the baseline sentence did not contain the construction/element in question or could be interpreted on the basis of a heuristic and the experimental sentence contained the structure/element and required the assignment of a complex syntactic structure to be understood (Table 1). In addition, the ability to combine operations was tested by presenting sentences with relative clauses and either a reflexive or a pronoun. Subjects were tested in object manipulation (OM) and picture matching (SPM) tasks, the latter with both whole sentence (Full SPM) and self-paced listening (auditory moving windows – AMW-SPM) presentation conditions, using digitized computer-delivered auditory stimuli. Accuracy and end-of-sentence reaction time in Full SPM and AMW-SPM were measured. Twenty examples of each sentence type were presented.

Procedures

Procedures were as in previous studies (Caplan et al, 1985, 1996, 2006, 2007; Caplan and Hildrebrant, 1988; Caplan and Waters, 2003) and will be summarized briefly. All sentence stimuli were recorded by one of the authors (DC) at a normal, but slow, speaking rate, and digitized using SoundEdit software. The files were edited for the auditory moving window task as follows. A marker (referred to as a “tag”) was placed in the waveform at each word boundary. In order to make word-to-word 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. Segments (tag to tag portions of the waveform) were saved as individual audio files. Pictures were created as JPEG files. Auditory and visual stimuli were used to create experiments in Superlab, which were presented on MacIntosh iBook computers, with auditory files presented over headphones.

Enactment (Object Manipulation (OM))—In the OM task, participants indicated thematic roles and co-indexation by manipulating paper dolls. Patients were told that the purpose of the experiment was to test their abilities to understand "who did what to whom" in the sentences. They were instructed to indicate "who did what to whom" by acting out the sentence using the items provided. The experimenter emphasized that the patient did not need to show details of the action of the verb, but had to clearly demonstrate which item was accomplishing the action and which item was receiving it. A practice session was given, during which some easy and some difficult sentence types were presented. During the practice session, the experimenter did not correct errors that a patient made, but did ask for repetitions and revisions of responses in which it was not clear which item initiated and which item received an action. Practice continued until the participant's actions could be clearly interpreted. Previous work has shown that inter-observer reliability is high for scoring these responses (Caplan et al., 1985; Caplan & Hildebrandt, 1988). The experimenter then proceeded with the digitized sentences and videotaped the participant's responses. See Caplan et al. (1985) and Caplan and Hildebrandt (1988) for a fuller description of the task.

Sentence-Picture Matching with uninterrupted auditory presentation (Full SPM)—In this test, each sentence was played auditorily with the two black and white line drawings in full view of the participant, and the participant was required to choose the drawing that matched the sentence by pressing one of two keys on the computer keyboard using fingers on the non-paretic hand. Drawings depicted the thematic roles in the sentence and a reversed set of thematic roles. Correct and incorrect pictures were displayed equally frequently on the right and left side of the computer screen; order of presentation on each side was randomized within sentence type. A practice session consisting of ten sentences of different sentence types preceded the task. Incorrect responses were not corrected by the experimenter. Responses were scored for accuracy and reaction time (RT).

Sentence-Picture Matching with auditory self paced (auditory moving window) presentation (AMW SPM)—The method was based on Ferreira et al. (1996a, b) and identical to that used in Caplan and Waters (2003) and Caplan et al (2007). On each trial, a participant heard a sentence that had been digitized and segmented as described above, and saw a pair of pictures as in Full SPM. The participant's task was to pace his/her way through the sentence as quickly as possible, by pressing a computer key for the successive presentation of each segment (the tag-to-tag portion of the waveform corresponding to each word), and to select the picture corresponding to the meaning of the sentence at the end of the sentence. If a participant pressed the button before the end of a segment, the segment was truncated at the point of the button press in order to discourage participants from pressing the button before they had heard and processed each segment.[1] Inter-response times between successive button presses, as well as response time and accuracy on the SPM task, were recorded.

Results[2]

Accuracy and RT

Table 2 shows the mean accuracy and RT, and standard deviations, for pwa and controls, as well as significant differences between experimental-baseline sentence pairs in each task.

Split half reliability

Pearsons' r and the Spearman Brown reliability co-efficients for each sentence type in each task for the pwa were all significant (all ps < .001; see Supplementary Materials

Between task correlations

Performance on each sentence type was significantly correlated across each pair of tasks for the pwa (Table 3). The mean correlation of same sentences was higher than of different sentences across all pairs of tasks (for SPM-AMW:SPM, $r_{SAME} = .78$, $r_{DIFF} = .65$; for OM:SPM, $r_{SAME} = .71$, $r_{DIFF} = .63$; for OM:SPM-AMW, $r_{SAME} = .76$, $r_{DIFF} = .66$).

Factor Analyses

Exploratory factor analysis of the accuracy results for each of the tasks showed unrotated first factors that accounted for about 90% of the variance, and Varimax rotation resulted in

[1]Both pwa and controls pushed the button for the next word before the offset of the current about 50% of the time. In about 90% of cases, the request for the next word was made after the end of the first syllable of the present word. Given the presence of the pictures, this would be late enough to identify the lexical item in the task context. There was no difference in the frequency of "early" button presses in pwa and controls or in the frequency of "early" button presses in correct and incorrect responses in with pwa or controls.

[2]Supplementary material set 1 shows the accuracy data, indicating performances in and below the normal range. Supplementary material set 2 shows the accuracy data, indicating performances above and at or below the chance range. Supplementary material set 3 shows the RT data, indicating performances in and above the normal range.

solutions in which two factors accounted for about 75% of the variance in each task (See Supplementary Materials). All sentence types loaded on the first factors in the unrotated analysis in each task, and the loadings of sentences on the retained factors in the rotated analyses did not correspond to linguistic or psycholinguistic groupings of the sentence types groupings are shown in Table 1.

Dissociations and deficits

Dissociations in performance that have been considered critical to drawing inferences about pwas' deficits consist of a performance that falls below some critical level (which we shall call "poor" performance) on one or more "experimental" sentences that require processing a particular syntactic element or structure (e.g., a "trace") and a performance that falls above that level (which we shall call "good" performance) on one or more baseline sentences that do not contain the element in question and that are otherwise matched with the experimental sentences. An issue that needs to be considered is what constitutes evidence that a person with aphasia has "poor" performance on an experiment sentence and "good" performance on a baseline sentence.

Caramazza and Zurif (1976) defined "poor" performance as an accuracy level that was at chance, and this criterion has been widely utilized (Drai and Grodzinsky, 1999; Grodzinsky et al, 1999; see Grodzinsky, 2000, for examples). Caramazza and Zurif (1976) did not define "good" performance, simply noting that accuracy was high (90%) on semantically reversible syntactically simple sentences. Other researchers have considered performance above chance to be "good" (Drai and Grodzinsky, 1999; Grodzinsky et al, 1999; Grodzinsky, 2000). Caplan et al (2006, 2007) argued that "good" and "poor" performance should be defined relative to a control group for conceptual reasons; only performance that is within the normal range can provide evidence for the integrity of the processes that underlie a response. Caplan et al (2006, 2007) also argued for the use of the normal range as the basis for determining "good" and "poor" performance because of the occasional overlap of chance and normal performance (Caplan and Hildebrandt, 1988; Caplan et al, 2006, 2007). Performance in the chance range can be normal when the mean level of normal performance on a sentence type is low and variability of normal performance is high, and scores above chance can be abnormal if normal performance is high and variability low. If the chance criterion is used, the first possibility can result in calling performances that are normal "poor," and the second in missing mild deficits. Finally, Caplan et al (2006, 2007) argued that RT data need to be considered in determining "poor" and "good" performance because of the possibility of speed-accuracy trade-offs. However, Caplan et al (2006, 2007) found no such trade-offs in their cases, and we did not find any in the present data set either, so RT data will not be considered further.

The use of a normal or chance range to define "good" and "poor" performance captures two legitimate, but very different, intuitions about deficits. The criterion based on chance performance captures the intuition that <u>consistent</u> performance is "good" performance and that inconsistent performance and consistently <u>incorrect</u> performance is "poor" performance. The criterion based on normal performance captures the intuition that <u>normal</u> performance is "good" performance and <u>below normal</u> performance is "poor" performance. These notions of what constitutes "good" and "poor" performance are not logically incompatible. One might imagine that researchers who define "good" performance as performance within the <u>normal</u> range would take performance below the normal range as "poor" performance, but it is not logically inconsistent to consider "good" performance to be normal performance and "poor" performance to be <u>inconsistent</u> performance (i.e., performance at or below chance).

An important consideration, raised by Caplan and Hildebrant (1988), is that the use of any of these criteria could lead to a situation in which a very small, and insignificant, difference in the number of correct responses on two sentence types led to the identification of a dissociation. This would arise if performance on the baseline sentence fell just above and performance on the experimental sentence fell just below the criterion level. To guard against this, Caplan and Hildebrant (1988) required that performance on the baseline and experimental sentences differ significantly to be considered a dissociation. Operationally, they used a Chi-square test to establish the significance of a difference in performance on two sentence types in a pwa.

Other definitions of "good" and "bad" performance are possible, though none have been used by researchers to our knowledge (see Supplementary Materials 4 for discussion). One that merits discussion in the text is to consider a difference score that is outside the normal range of difference scores for correct performance on two sentence types to be abnormal. The sense of a difference in performance on two sentence types being significant differs in this approach and in the standard approach, discussed above, that examines performance on an experimental and a baseline sentence type separately. The use of a difference score that is outside the range of difference scores indicates that the difference in an individual pwa is greater than that expected in the control population. A significant Chi-square value in the comparison of performance on two sentence types indicates that the difference in performance on the two sentence types in a pwa is greater than what would be expected by chance. As with the use of the normal or chance range of performance as the basis for saying that a performance is "good" or "poor," these notions of a dissocation can yield different results. It is possible for a significant difference by Chi-square to be within the normal range of differences, and for a difference that is outside the normal range of differences not to be significant by Chi-square.

It is possible that empirical criteria could favor one of these intuitions about "good and "poor" performance. For instance, it is possible that there is an interesting relation between neural lesions and deficits defined in terms of one type of dissociation and not another. However, no such empirical basis for selecting an approach to defining deficits exists to our knowledge. We therefore believe that, at this point in the study of aphasic deficits in syntactically based comprehension, the choice of a criterion for what constitutes a deficit should reflect the purpose to which deficit analysis is to be put. In this study, we are interested in the implications of deficits for models of normal language structure and processing. If pwa have patterns of behavior that are correctly analyzed as deficits affecting specific syntactic elements or particular psycholinguistic operations, those patterns of behavior provide support for the existence of those syntactic elements or psycholinguistic operations. If this is the purpose of identifying deficits, it is desirable to provide the strongest possible evidence of a structure-specific deficit, because strong dissociations permit stronger inferences about the existence of syntactic elements or psycholinguistic operations (Shallice, 1988). These considerations led us to adopt the most conservative criteria we can think of for dissociations. In addition, to relate the results to earlier work, we used a criterion based on performance in both an experimental and baseline sentence type, rather than a difference score (but see below).

We therefore recognized the presence of a dissociation when three criteria were met: 1) normal performance on a baseline sentence; 2) performance on the experimental sentence that is both at or below chance and below the normal range; 3) a statistically significant difference in performance levels for the baseline and experimental sentences.

Operationally, the criteria were ascertained in the following ways.

1. "Normal accuracy criterion." The normal accuracy range was determined on the basis of the mean percent correct and standard deviation of the control performances corrected for sample size (Crawford and Howell, 1998). By these

criteria, patients whose percent correct for any sentence type was within 1.74 sds of the control mean percent correct for that sentence type on a given task were considered to have “good” performance on that sentence type in that task, and performance below that level was considered to be “poor” performance.[3]

2. “Chance criterion.” The calculation of chance depends upon the number of possible responses. In our version of SPM, the number of choices is always 2. In OM, this is not always the case. For instance, if any noun phrase (NP) mentioned in a sentence could play any thematic role around any verb, there are 36 possible responses (including reflexive actions) in sentences with three noun phrases and two verbs. However, Caplan and Hildebrandt (1988) showed that pwa never make most of these responses in OM, but restrict their responses to "linear" responses in which the first post-verbal NP is the theme of that verb and either the nearest preverbal NP or the first NP in the sentence is the agent. We have also found that pwa do not interpret reflexives as pronouns or *vice versa* (Caplan and Hildebrandt, 1988). Finally, we considered that, in active and CS sentences, a possible error was the inversion of NPs as agent and theme. Based on these considerations, we set the number of alternatives over which chance should be computed as 2 for all sentences except SO sentences (including SO sentences with reflexives or pronouns), where that number was 3. Expansion of the binomial then yields the result that, in OM, 14 or more correct responses in any sentence type except SO, SOREF and SOPRO, and 11 or more correct responses in SO, SOPRO and SOREF sentences, constitutes above chance performance.

3. Statistical significance. To determine whether a difference established by criterion (1) or (2) was significant, a 2 X 2 Chi-square test that compared the number of correct and incorrect responses on the experimental and baseline sentence types was performed.

Dissociations were examined for what they showed about deficits. Following the terminology in Caplan et al (2006), a *task-independent <u>construction</u>-specific deficit* was identified when a participant performed poorly on <u>one</u> experimental sentence type and well on the corresponding baseline in all tasks, and a *task-independent <u>structure</u>-specific deficit* was identified when a participant performed poorly on <u>all versions</u> of an experimental sentence type and well on the corresponding baselines in all tasks. Task-independent deficits are the strongest available evidence for a deficit affecting a linguistic representation or psycholinguistic operation. Caplan et al (2006) considered poor performance on an experimental sentence type and good performance on a baseline sentence type on only one of the two tasks they used (SPM and OM) as a *task-dependent construction- (or structure-) specific deficit*. An anonymous reviewer pointed out that this criterion is too weak; it is also necessary to show that the number of correct responses on the experimental and baseline sentence types differs in the tasks that are used. Accordingly, we analyzed patterns of poor performance on an experimental sentence type and good performance on a baseline sentence type on only one or two of the three tasks used here in 2 X 3 Chi-square analyses to determine if the number of correct responses on the experimental and baseline sentence types differed in the three tasks. If the Chi-square value was significant, we labeled the pattern a *task-dependent construction-specific deficit*. If this pattern occurred in all versions of an experimental-baseline sentence type contrast, it was considered a *task-dependent structure-specific deficit*.

[3]Non-normal distributions of accuracy scores make the use of standard deviations to establish the normal range a conservative measure in sentences where accuracy is close to ceiling.

All instances in which a pwa had a dissociation in performance on an experimental and baseline sentence type in a task by the criteria listed above are shown in Table 4. There were 62 dissociations in a task *by* the three-part criteria above.

Task-independent deficits

No patient had the same dissociation in all tasks; i.e, by criteria (1) – (3) above, there were no task-independent deficits.

Task-dependent deficits

Of the 62 dissociations, 36 occurred in one task only. By the criteria in Caplan et al (2006), these would be considered task-specific structure- or construction-specific deficits. However, of these 36, the 2 X 3 Chi-square (number correct in the experimental and baseline sentence X task) was not significant in 31 instances. By the criteria above, a task-specific structure- or construction-specific deficit cannot be diagnosed. In the remaining 5 instances, the 2 X 3 Chi-square was significant; by the criteria adopted here, these pwa have task-specific structure- or construction-specific deficits. In 18 instances, the same dissociation was found in 2 tasks in the same pwa (36 dissociations). Of these, the 2 X 3 Chi-square was not significant in 9 pwa (18 dissociations), and was significant in 4 (8 dissociations). By the criteria adopted here, the latter 4 pwa have task-specific structure- or construction-specific deficits in two tasks. Table 5 shows the 13 pwa whose performances qualify as task-specific structure- or construction-specific deficits.

Task-independent deficits by other criteria

Because these criteria are so conservative and restrictive, and the existence of task-independent deficits so important for linguistic and psycholinguistic theory, we also used three other less restrictive approaches to defining dissociations: 1) as performance on a baseline sentence that is within the normal range and performance on an experimental sentence that is below the normal range; 2) as performance on a baseline sentence that is above chance and performance on an experimental sentence that is at or below chance; 3) as a difference between the number of correct responses on an experimental and a corresponding baseline sentence that was outside the normal range of differences based on Crawford and Howell (1988), and above chance performance on the baseline sentence. The latter requirement is necessary to avoid identifying differences between two scores, both of which are in the chance range, as a dissociation that allows the inference that operations required only for the experimental sentence and not the corresponding baseline sentence are affected. In keeping with our view that performance in the normal range is the only basis for inferring the (possible) integrity of the processes that underlie a response, we also set the criterion that differences outside the normal range would be considered as dissociations if performance on the baseline sentence was within the normal range.

The results of the first two analyses are presented in Supplementary Materials 5. Although these less conservative criteria led to more dissociations, the overall picture of the data did not change regardless of the criterion used. Using the "normal" performance criterion, one individual had a task-independent structure-specific deficit and one had a construction-specific, task-independent deficit. By the "chance" criterion, one individual had a task-independent structure-specific deficit and three had a construction-specific, task-independent deficits.

The "difference" criterion with above-chance requirement for performance on the baseline sentence resulted in the largest number of task-independent deficits (15), which were seen in 10 pwa (Table 6). When performance on the baseline sentence was required to be within the normal range, the number of task-independent deficits fell to 4 (indicated in Table 6). We will discuss the results using the less restrictive criterion (baseline sentence performance above chance). In no case, did a pwa have a task-independent structure-specific deficit; i.e., a deficit affecting all sentences containing a particular structure in all tasks. Case 54003 had task-independent deficits affecting CO and SO sentences, which would be considered a task-independent structure-specific deficit were it not for the fact that s/he did not have a significant difference between SOPRO and SOREF sentences and their SS baselines (see discussion of this pwa below). One pwa, Case 54001, had three task-independent construction-specific deficits, affecting passives, cleft object sentences, and sentences with pronouns. Two pwa had two task-independent construction-specific deficits. Case 54042 had task-independent deficits affecting cleft object sentences and sentences with reflexives. Case 54054 had task-independent deficits affecting cleft object sentences and sentences with object relative clauses and a pronoun compared to sentences with object relative clauses and a sentence final third NP (i.e., a deficit affecting the ability to co-index a pronoun only in this syntactic context). The remaining 6 pwa had task-independent deficits affecting one construction.

Discussion

Our focus in this section is on the implications of the patterns of performance seen in pwa. We will first discuss associations in performance and their implications, then turn to dissociations.

The between-task correlations and factor analyses replicate existing data (Caplan et al, 1996, 2006, 2007; Gutman et al, 2010, 2011). These results show that performance on all sentence types is significantly correlated in aphasia. The factor analyses show that pwa who perform poorly (or well) on one sentence type in one task tend to perform poorly (or well) on other sentence types in that task. The unrotated factor analyses are very similar for the three tasks, indicating that this is true in all tasks. The across-task correlations show that patients who perform poorly (or well) on one sentence type in one task tend to perform poorly (or well) on that sentence type in other tasks.

These findings have suggested that a large part of performance is determined by a single factor that captures the overall ability of a person with aphasia to understand sentences that require syntactic analysis to be understood and to demonstrate that understanding in picture matching and enactment tasks. A second finding in the literature (e.g., Caplan et al, 1985, 2007) is that sentences that are more psycholinguistically complex are less well comprehended. This was true in the present study as well. Because this single factor is sensitive to the overall processing load imposed by a sentence type, it has been termed "resource reduction."

We have commented extensively on what is meant by "resource reduction" elsewhere (Caplan, 2012), and will only briefly review our views on this concept here. Parsing and interpretation have increased computational demands at certain points in sentences. For instance, in surprisal models (Hale, 2001; Levy, 2008), the parser projects structure, and when the input does not match the anticipated structure, new structure(s) must be created, increasing parsing load. Surprisal is measured as the negative log probability of a structure or word w_i in context $\{w_1 \ldots w_{i-1}\}$, which Levy (2008) showed is equivalent to the difference between two successive probability distributions of structures (i.e., their "relative entropy"). In retrieval models, retrieval of items in memory is subject to variable amounts of interference (Lewis and Vasishth, 2005; Lewis et al, 2006; Van Dyke & Lewis, 2003) and involves more complex searches at particular points in sentences (Alcocer and Phillips, submitted). The extra processing associated with higher surprisal values, more interference at retrieval, more complex searches, and other aspects of parsing and interpretation, is reflected in behaviors such as eye fixation durations, self paced reading times, cross-modal lexical decision times, and other measures. Individual differences in the effects of these loads have been modeled as results of differences in the efficiency with which these operations are carried out (Just and Carpenter, 1992; Miyake et al, 1994). We have suggested that reductions in processing efficiency (in our terms, in processing resource capacity) are partially responsible for aphasic syntactic comprehension. A more specific version of this model is the claim that parsing and interpretation is slowed in pwa (Burkhardt et al, 2008; Haarman and Kolk, 1991a, b).

Reductions in the efficiency of operations other than those of the parser/interpreter would be expected to have similar effects. In particular, slowed or otherwise disrupted activation of lexical representations has been suggested as the source of these types of effects (Love et al, 2008; but see Dickey et al (2007) for an alternate view). It is also possible that a more general disturbance affecting skilled performances, such as an across-the-board slowing of processing speed, could produce these patterns. Finally, resource reduction could affect operations that are needed to perform a task to which comprehension is directed, such as planning actions, inspecting pictures, etc. Performing tasks also makes demands on executive functions in the form of deployment of attention, maintenance of task goals, uploading mechanisms that support task performance (scene analysis, action planning, and other task demands), executing those mechanisms, response selection, assessment of success on a trial, and other processes.

Limitations in any of the functions listed above would give rise to some degree of correlation in how performance was affected. They would also give rise to effects of complexity, since more complex operations require more of these processes. The results in this study are consistent with any of these types of operations being affected. As noted, the literature contains more specific suggestions regarding specific mechanisms (such as slowed lexical access) that could lead to these aspects of performance, but more work is required to know if any of these suggestions is correct in any pwa.

The data presented here also replicate previous reports regarding dissociations of performance (Caplan et al., 2006, 2007). The critical finding is that, *although performance on individual sentence types is highly correlated across tasks, critically large differences in performance on pairs of sentence types that are selected to reflect linguistic or psycholinguistic operations are not*. As in previous studies, dissociations of performance did not occur in the same constructions, let alone the same structures, on different tasks. The major concern regarding Caplan et al (2006, 2007) was that, with only 10 examples of each sentence type in that study, it was unclear that each pwa had been examined carefully enough to be sure that his/her performance was reliably captured. The present study addresses this concern by using 20 examples of each sentence. The high split half reliability indicates that performance on each sentence type is stable within each task. We have also approached the identification of deficits in a novel way, by looking for differences between experimental and baseline sentences that are outside the range of differences expected in normal individuals. This approach also failed to identify any pwa in whom a task-independent structure-specific deficit was found. The present study demonstrates quite convincingly that pwa rarely, if ever, have deficits affecting one structure on different tasks. [4]

[4]Newhart et al (2012) reported that 6/53 pwa in the acute post-stroke period had task-independent structure specific deficits affecting either passive (compared to acive) or cleft object (cmpared to cleft subject) sentences, when tested on SPM and OM. However, these pwa were tested ononly 10 examples of reversible A, P, CS and CO sentences; a dissociation was identified if a pwa had at or below chance performance on P or CO sentences and performance on P or CO sentences that was at least 10% (i.e., one sentence) different than A or CS sentences; and the only construction that tested object extraction was CO. These criteria are too weak to establish a dissociation (see text) and it is unclear from the data available that any of the pwa met reasonable criteria for deficits in even one task.

A novel feature of this study is the use of self-paced word-by-word presentation (auditory moving window – “AMW-SPM”). This task is identical to “Full” SPM except for allowing participants to do whatever processing they wish before having to encounter the next word and requiring them to request the next word. The fact that there were differences in performance on the same sentence type in Full SPM and AMW SPM despite these seemingly small differences between the two tasks points to the extent to which details of task conditions affect comprehension.

The fact that task-independent structure-independent deficits are exceptionally rare in aphasia, and may not exist at all, affects the implications of data from syntactic comprehension tasks in pwa for theories of syntactic structure and processing, for the relation of other cognitive processes to parsing and interpretation, and for the neural basis of the parser/interpreter. Specifically, conclusions regarding the nature of syntactic structures (e.g., that “traces” exist) cannot be based on performance in one task; correlations between performance on a memory task and the level of performance of pwa on a type of sentence cannot be taken as evidence that the memory functions tested are involved in the operation of the parser/interpreter; and finding that lesions in a brain area lead to poor performance on sentences containing a particular syntactic structure or element in pwa in one task does not imply that that brain area is necessary for the parsing and interpretive operations that apply in that sentence.

An important practical question therefore is: how many tasks must show a dissociation before it can be interpreted as evidence for a task-independent structure- or construction independent deficit? There is no clear answer to this question, other than to say “the more, the better.” The differences seen between AMW-SPM and Full-SPM suggest that caution is called for in such interpretations.

This study documents two other types of deficits: task-independent construction-specific deficits, and task-dependent construction-specific deficits.

In this study, when dissociations were established by “difference” criteria (with what we consider a weak criterion for considering performance on the baseline sentence to be “good”), 10 patients had construction-specific, task independent deficits. What might produce this pattern? One possibility is that these performances result from the demands made by the need to process two structures or elements in one sentence. This could be the case in Case 54054, whose differences between SOREF and SO sentences suggests that the combination of the demands of SO sentences (object extraction) and co-indexing a pronoun exceeded his/her processing capacity, whereas the demands of reflexives or SO sentences in isolation did not. This is an example of the type of performance that Caplan and Hildebrandt (1988) took as evidence for resource reduction in an individual pwa. This pattern is not the rule, however. The remaining construction-specific, task independent deficits occurred in the simpler syntactic contexts, despite above-floor performance on the sentence types that required the two operations. For instance, Case 54001 had a task independent, construction-specific deficit seen in the difference score for CO-CS sentences but in no other contrast involving object *vs* subject extraction, although s/he had performances in SO, SS, SOPRO, SSPRO, SOREF and SSREF sentences in the same range as CO and CS sentences (Table 7). The same is true for Case 54003, mentioned above, whose difference scores showed task-independent deficits for CO-CS and SO-SS sentences, but not SOREF-SSREF or SOPRO-SSPRO sentences (Table 7). Supplementary Materials 1 and 2 show that other pwa with construction-specific, task independent deficits did not have ceiling or floor effects that obscured effects for other, related, sentence types. A second argument that not all construction-specific, task independent deficits are due to resource reduction is that there are double dissociations affecting these deficits. For instance, Case 54017 had a task independent, construction-specific deficit affecting passives but not CO, and Case 54003 had the opposite pattern.

We note that we have used the term "construction" in an informal way, to refer to syntactic structures that include a particular linguistic element (e.g., trace (in Chomsky's model), a reflexive or a pronoun) together with their sentential context. A question arises as to whether the task independent, construction-specific deficits documented here provide evidence for the types of syntactic units that have been called "constructions" in the linguistic literature (e.g., Goldberg, 1995, 2006). It is unclear whether they do. In the sense used the literature, "constructions" are syntactic forms that have particular meanings (e.g., the sense of intentional receipt of an item by the post-verbal NP in an English inner dative). The aspects of meaning that are tested in the protocols used here are not construction-specific contributions to meaning. For instance, the construction-specific contribution of a cleft structure to meaning is to place an item in linguistic focus, but the measure of performance here was not whether a pwa understood that the clefted item is in focus but whether s/he could assign thematic roles in clefts, an aspect of meaning is common to part of the structure of clefts, relative clauses, and *wh* questions. Clear evidence from aphasia for the existence of constructions as discussed in the literature would be selective deficits affecting the unique aspect of the form-meaning mapping in a construction. This has not been studied, to our knowledge. On the other hand, the existence of task-independent deficits that affect the aspect of the form-meaning mapping common to several constructions in only one construction does raise concerns about a view of linguistic structure that Chomsky has articulated, according to which surface forms such as passives or clefts are not real syntactic entities but are rather the product of combinations of more basic units and operations, such as those that are responsible for representing certain types of moved constituents (traces). The present results provide no evidence for these more abstract, "atomic" linguistic units and operations, and do suggest that what Chomsky considers surface manifestations of such units and operations contain features that are subject to processing disruptions in aphasia. More work into the nature of task independent, construction-specific deficits is needed to understand the reasons for these performances and their implications for linguistic and psycholinguistic theory.

The last, and most common, pattern of dissociation seen in this study and in Caplan et al (2006) is poor performance on an experimental sentence and good performance on the corresponding baseline sentence in one task – what Caplan et al (2006) called *task-specific, construction-specific deficits*. These deficits arose in all tasks, indicating that specific combinations of construction- and task-related operations trigger comprehension failure. We will briefly comment on how such deficits might arise.

In pwa, comprehension has been shown to incrementally map onto aspects of visual arrays. For instance, Dickey and Thompson (2004, 2009) and Dickey et al (2007) documented increased eye fixations on pictures of objects at and after the verb of an object-extracted wh-question in normal individuals and in pwa when their responses to those questions were correct, but, when the pwa answered the questions incorrectly, these increased fixations were not found. The increased fixations on the item depicting the object, which reinforce the correct interpretation of the sentence, are not available in OM, and could be critical to the ability of a person with aphasia to interpret a wh-question.

There is also evidence for effects of particular constructions on these eye fixation patterns. Unlike the results with wh-questions, Dickey et al. (2007) did not find a theme preference following the verb for either normal or aphasic participants when object cleft structures were tested. The different effects of the two constructions could result from the nature of the information in a cleft as opposed to a wh-question.[5]

The consideration that task differences of this sort could interact with sentences in determining performance of pwa opens the door to many more possibilities. The studies cited above, like those using the visual world paradigm in normals (Tanenhaus et al, 1995; Altmann and Kamide, 2007; Farmer et al, 2007), use visual stimuli that consist of arrays of objects mentioned in a stimulus sentence, plus distracters. SPM tasks used with pwa standardly depict actors accomplishing thematic roles, and usually require a choice between a picture depicting the syntactically encoded thematic roles and a syntactically reversed foil. There is evidence suggesting that this difference might affect pwa. Hanne et al (2011) found that, in a SPM task with depictions of thematic roles, errors made by pwa showed early and sustained fixations on the foil, a contrast to the pattern in Dickey's studies, where errors showed late returns to an agent rather than late sustained fixation on a theme (seen in controls). The difference could be due to the depiction of thematic roles, rather than single entities, encouraging early commitment to an interpretation. Other factors that could affect comprehension are the detail and clarity with which items and actions are depicted (many studies use cartoon figures, which may require more visual processing than realistic pictures of actors), the timing of the presentation of a picture and a sentence (preview might affect the interaction, and presenting a picture only after a sentence was over would deny the comprehender the advantage of on-line inspection of potentially reinforcing items and prevent the disadvantage of on-line inspection of competitors).

All these potential accounts indicate how task-specific construction-specific deficits could arise: the additional aid to comprehension of particular sentences available in sentence-picture matching is not available in object manipulation. The picture-matching advantage could, however, be reversed. If a person with aphasia could assign and interpret a construction with an object-extracted clause, it is possible s/he could produce a motor movement that demonstrates his/her understanding but that the presence of a distracter element in a picture could interfere with his/her ability to sustain that interpretation until making a response in a sentence-picture matching task.

The discussion above illustrates possible mechanisms that could produce task-specific, construction-specific deficits. Appendices 1 and 2 show that task-specific, construction-specific deficits are extremely variable in their nature, and it remains an important, if distant, goal to model the combination of disorders of parsing and interpretation and fulfillment of task demands that lead to these deficits.

Considering the results within a broader perspective, the finding that vascular aphasic patients have task-specific deficits in the domain of syntactically based comprehension can be compared to recent observations regarding lexical impairments. Jefferies and Lambon Ralph (2006) studied the semantic deficits of 10 aphasic stroke patients and 10 patients with semantic dementia. The patient groups differed along several dimensions, one of which was the consistency of performance on similar and dissimilar tasks. While both stroke and SD cases had high consistency of performance across different versions of a task (e.g., pyramids and palm trees presented as words and as pictures), only the SD patients showed consistent performances across different types of tasks (e.g., pyramids and palm trees, naming, word-picture matching). Jefferies and Lambon Ralph commented regarding the stroke patients that "Even though the different semantic tasks contained similar or identical concepts … their executive/control requirements varied (p. 2143) … core amodal semantic representations interact with a semantic control system that shapes or regulates the activation of the

[5]The difference could also be due to the fact that the wh-question was the stimulus that required a response whereas a second question had to be posed about a cleft. Dickey and Thompson (2009) suggest that the difference may be due to the set of the sentences presented – in Dickey et al (2007), subject-extracted sentences were not part of the stimulus set whereas subject-extracted wh-questions were presented in Dickey and Thompson (2009). If this is the correct account of the difference, it indicates that task-related strategies interact with task-specific aspects of the comprehension task, further adding to the factors that might affect performance.

information associated with a concept in order to produce task-appropriate behaviour. It is this aspect of semantic cognition that we believe is compromised in our comprehension-impaired stroke patients. (p. 2145)" (see also Jefferies et al, 2007, 2008). This general model is very similar to the one we are proposing regarding syntactic deficits, and suggests that a common feature of the psychopathogenesis of vascular aphasic symptoms in many, perhaps all, language domains is a disturbance of the interaction of language-specific processes with task-related operations, including operations at the level of executive control. The idea that difficulties at the level of control systems are part of the mechanism that produces task-specific deficits affecting certain types of representations in vascular aphasia also overlaps with the idea we introduced earlier that intermittent disruption of control mechanisms could be a source of what we have called "resource reduction" in the domain of syntactically based comprehension in the same group of patients.

We conclude this discussion with two cautionary notes.

The first was raised by an anonymous reviewer, who questioned our interpretation of the results on the basis of the view that any hard-and-fast criterion to identify task-independent structure-specific deficits is misguided. S/he argued that it is more appropriate to see such deficits as falling along a continuum; that structure-specific task-independent deficits are like resource reductions in being graded, rather than all-or-none. The reviewer then argued that that structure-specific task-independent deficits would interact with task demand, such that what is in fact a true task-independent structure-specific deficit might be obscured by the fact that tasks vary in difficulty.

The challenge in investigating this model seems to us to be that a minor structure-specific task-independent deficit would not be easy to identify in an individual pwa because performance on the experimental sentence type would differ minimally from that on the baseline sentence type. It is, however, possible that modeling would serve to confirm or refute this type of model. One approach is Rasch modeling, which to date has not supported the view that sentence-type specific deficits are needed to account for accuracy data (see above). Rasch modeling does not examine specific mechanisms, and another approach would be to try to simulate the accuracy data by lesioning implemented models that specified both parsing/interpretive and task operations, as Patil et al (2012) have done for eye fixations in a SPM. The question that such models could address is whether it is necessary to include task-independent, structure-specific deficits in such simulations. These analyses are beyond the scope of the present paper, which examines the data using the framework currently utilized by researchers in the field (namely, to consider deficits to exist on the basis of dissociations in performance on different sentence types).

There are, however, some aspects of the data that can speak to the reviewer's suggestion, short of extensive modeling. We have considered one such aspect of the data. As we understand his/her suggestion, it implies that one reason that differences between performance on experimental and baseline sentences might be obscured is that there are ceiling effects in easy tasks and floor effects in hard or unusual tasks (such as SPM-AMW). Examination of the individual pwa data (Supplementary Materials 1) shows that, although some pwa do show ceiling and floor effects, this is not the major reason for the absence of task-independent deficits. The vast majority of responses in all tasks are in a range of accuracy that would allow a dissociation to be seen in a task, if it was present.

The second caveat is that end-of-sentence data are incomplete measures of the activity of the parser/interpreter. As on-line measures have become available, they have affected the interpretation of deficits. *For* instance, until recently, chance performance on sentences that contained a linguistic representation was interpreted as possible evidence for loss of that representation (cf, Grodzinsky's (2000) formulation of the trace deletion hypothesis). However, on-line measures of performance have been reported to be normal in some pwa in sentences that they interpret correctly and abnormal in sentences that interpret incorrectly, when overall performance is at chance (Caplan and Waters, 2003; Dickey et al, 2007s), a pattern that strongly indicates that the abnormality is an intermittent inability to process a representation, not its loss. More importantly, end-of-sentence data may be insensitive measures of the activity of the parser/interpreter. We have found that pwa whose accuracy on sentences in SPM-AMW is within the normal range at times show abnormally large effects of local processing load on self-paced listening times (Caplan et al, unpublished data). These pwa appear to have some abnormality in parsing/interpretation, or a temporally aligned operation, that is not so severe as to affect final comprehension. Associations and dissociations of different on-line performances in different sentence types and tasks are important to explore to develop better empirically based models of deficits in syntactic comprehension.

To conclude, this study replicates previous findings regarding associations and dissociations of performances in aphasia. Associations of performance replicate previous results and, along with complexity effects in the group and in individual patients, point to a single factor that affects complex sentences more than simple sentences as an important determinant of aphasic impairment, which we have called "resource reduction." The pattern of dissociations clearly establishes that task-independent deficits affecting particular structures or constructions are rare; thus far in the literature, they have not been shown to occur in isolation. This has negative and positive implications. On the negative side, it indicates that patterns of performance on sentence types in one comprehension task are not an adequate basis for attributing deficits in parsing and interpretation to pwa. On the positive side, it points to the need to develop theories of the effects of task on comprehension of specific sentence types to account for deficits in syntactically based comprehension in aphasia.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by NIDCD grant DC00942 to David Caplan.

References

Alcocer P, Phillips C. (submitted). Using relational syntactic constraints in content-addressable memory architectures for sentence processing.

Altmann GTM, Kamide Y. The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. Journal of Memory and Language. 2007; 57:502–518.

Burkhardt P, Avrutin S, Piñango MM, Ruigendijk E. Slower-Than-Normal Syntactic Processing in Agrammatic Broca's Aphasia: Evidence From Dutch. Journal of Neurolinguistics Journal of Neurolinguistics. 2008; 21(2):120–113.

Caplan, D. Language: Structure, Processing and Disorders. Cambidge, MA: MIT Press; 1992.

Caplan, D. Resource reduction accounts of syntactic comprehension disorders. In: Bastiaanse, R.; Thompson, CK., editors. Perspectives on Agrammatism. New York: Psychology Press; 2012.

Caplan, D.; Hildebrandt, N. Disorders of Syntactic Comprehension. Cambridge, MA: MIT Press (Bradford Books); 1988.

Caplan D, Waters GS. On-line syntactic processing in aphasia: studies with auditory moving windows presentation. Brain and Language. 2003; 84(2):222–249. [PubMed: 12590913]

Caplan D, DeDe G, Michaud J. Task-independent and task-specific syntactic deficits in aphasic comprehension. Aphasiology. 2006; 20:893–920.

Caplan D, Baker C, Dehaut F. Syntactic determinants of sentence comprehension in aphasia. Cognition. 1985; 21:117–175. [PubMed: 2419022]

Caplan D, Hildebrandt N, Makris N. Location of lesions in stroke patients with deficits in syntactic processing in sentence comprehension. Brain. 1996; 119:933–949. [PubMed: 8673503]

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

Caplan D, DeDe G, Michaud J. Task-independent and task-specific syntactic deficits in aphasic comprehension. Aphasiology. 2006; 20:893–920.

Caplan D, Waters GS, DeDe G, Michaud J, Reddy A. A study of syntactic processing in Aphasia I: Behavioral (psycholinguistic) aspects. Brain and Language. 2007; 101:103–150. [PubMed: 16999989]

Caramazza A, Zurif EB. Dissociation of algorithmic and heuristic processes in language comprehension: Evidence from aphasia. Brain and Language. 1976; 3:572–582. [PubMed: 974731]

Caramazza A, Capitani E, Rey A, Berndt RS. Agrammatic Broca's Aphasia is not Associated with a Single Pattern of Comprehension Performance. Brain and Language. 2000; 72:193–218. [PubMed: 10764517]

Chomsky, N. Reflections on Language. New York: Pantheon; 1975.

Chomsky, N. Lectures on Government and Binding. Dordrecht: Foris; 1981.

Chomsky, N. The Minimalist Program. Cambridge, Mass: The MIT Press; 1995.

Crawford JR, Howell DC. Comparing an individual's test score against norms derived from small samples. The Clinical Neuropsychologist. 1998; 12:482–486.

Cupples L, Inglis AL. When task demands induce "asyntactic" sentence comprehension: A study of sentence interpretation in aphasia. Cognitive Neuropsychology. 1993; 10:201–234.

Dickey MW, Thompson CK. The resolution and recovery of filler-gap dependencies in aphasia: evidence from on-line anomaly detection. Brain and Language. 2004; 88(1):108–127. [PubMed: 14698736]

Dickey MW, Thompson CK. Automatic processing of wh- and NP-movement in agrammatic aphasia: Evidence from eyetracking. Journal of Neurolinguistics. 2009; 22(6):563–583. PMC2748948. [PubMed: 20161014]

Dickey M, Choy J, Thompson CK. Real time comprehension of wh-movement in aphasia: Evidence from eyetracking, while listening. Brain and Language. 2007; 100:1–22. [PubMed: 16844211]

Drai D, Grodzinsky Y. Comprehension regularity in Broca's aphasia? There's more of it than you ever imagined. Brain and Language. 1999; 70(1):139–143. [PubMed: 10534378]

Farmer TA, Anderson SE, Spivey MJ. Gradiency and visual context in syntactic garden-paths. Journal of Memory and Language. 2007; 57:570–595. [PubMed: 18037980]

Ferreira F, Anes MD, Horine MD. Exploring the use of prosody during langauge comprehension using the auditory moving window technique. Journal of Psycholinguistic Research. 1996a; 25:273–290. [PubMed: 8667299]

Ferreira F, Henderson JM, Anes MD, Weeks PA, McFarlane DK. Effects of lexical frequency and syntactic complexity in spoken language comprehension: Evidence from the auditory moving window technique. Journal of Experimental Psychology: Learning, Memory & Cognition. 1996b; 22:324–335.

Goldberg, AE. Constructions: A Construction Grammar Approach to Argument Structure. Chicago: University of Chicago Press; 1995.

Goldberg, AE. Constructions at Work: the nature of generalization in Language. Oxford: Oxford University Press; 2006.

Grodzinsky Y. Language deficits and the theory of syntax. Brain and Language. 1986; 27:135–159. [PubMed: 3947938]

Grodzinsky Y. Trace deletion, theta-roles, and cognitive strategies. Brain and Language. 1995a; 51:467–497.

Grodzinsky Y. A restrictive theory of agrammatic comprehension. Brain and Language. 1995b; 51:26–51.

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

Grodzinsky Y, Piñango MM, Zurif E, Drai D. The critical role of group studies in neuropsychology: comprehension regularities in Broca's aphasia. Brain and Language. 1999; 67(2):134–147. [PubMed: 10092346]

Gutman R, DeDe G, Michaud J, Liu JS, Caplan D. Rasch models of aphasic performance on syntactic comprehension tests. Cognitive Neuropsychology. 2010; 27:230–244. [PubMed: 21058076]

Gutman R, DeDe G, Caplan DLiu, Jun S. Rasch model and its extensions for analysis of aphasic deficits in syntactic comprehension. Journal of the American Statistical Association. 2011; 106:1304–1316.

Hanne S, Sekerina IA, Vasishth S, Burchert F, De Bleser R. Chance in agrammatic sentence comprehension: what does it really mean? Evidence from eye movements of German agrammatic aphasic patients. Aphasiology. 2011; 25:221–244.

Haarmann HJ, Kolk HHJ. A computer model of the temporal course of agrammatic sentence understanding: the effects of variation in severity and sentence complexity. Cognitive Science. 1991a; 15:49–87.

Haarmann HJ, Kolk HHJ. Syntactic priming in Broca's aphasia: evidence for a slow activation. Aphasiology. 1991b; 5:1–36.

Hale J. A probabilistic Earley parser as a psycholinguistic model. In Proceedings of NAACL. 2001; Vol.2:159–166.

Jefferies E, Lambon Ralph MA. Semantic impairment in stroke aphasia versus semantic dementia: a case-series comparison. Brain. 2006; 129:2132–2147. [PubMed: 16815878]

Jefferies E, Baker SS, Doran M, Lambon Ralph MA. Refractory effects in stroke aphasia: a consequence of poor semantic control. Neuropsychologia. 2007; 45:1065–1079. [PubMed: 17074373]

Jefferies E, Patterson K, Lambon Ralph MA. Deficits of knowledge versus executive control in semantic cognition: insights from cued naming. Neuropsychologia. 2008; 46:649–658. [PubMed: 17961610]

Just MA, Carpenter PA. A capacity theory of comprehension: Individual differences in working memory. Psychological Review. 1992; 99:122–149. 1. [PubMed: 1546114]

Levy R. Expectation-based syntactic comprehension. Cognition. 2008; 106(3):1126–1117. [PubMed: 17662975]

Lewis RL, Vasishth S. An activation-based model of sentence processing as skilled memory retrieval. Cognitive Science. 2005; 29:375–419. [PubMed: 21702779]

Lewis RL, Vasishth S, Van Dyke JA. Computational principles of working memory in sentence comprehension. Trends in Cognitive Sciences. 2006; 10:44–54.

Love T, Swinney D, Walenski M, Zurif E. How left inferior frontal cortex participates in syntactic processing: evidence from aphasia. Brain and Language. 2008; 107:203–219. [PubMed: 18158179]

Mauner G. Examining the empirical and linguistic bases of current theories of agrammatism. Brain and Language. 1995; 50:339–368. [PubMed: 7583194]

Mauner G, Fromkin V, Cornell T. Comprehension and acceptability judgments in agrammatism: Disruptions in the syntax of referential dependency. Brain and Language. 1993; 45:340–370. [PubMed: 8269330]

Martin R. Working memory doesn't work: a critique of Miyake et al.'s capacity theory of aphasic comprehension deficits. Cognitive Neuropsychology. 1995; 12:623–636.

McNeil MR, Odell K, Tseng C-H. Toward the integration of resource allocation into a general theory of aphasia. Clinical Aphasiology. 1991; 20:21–39.

Miyake A, Carpenter PA, Just MA. A capacity approach to syntactic comprehension disorders: Making normal adults perform like aphasic patients. Cognitive Neuropsychology. 1994; 11(6): 671–717.

Miyake A, Carpenter PA, Just MA. Reduced resources and specific impairments in normal and aphasic sentence comprehension. Cognitive Neuropsychology. 1995; 12(6):651–679.

Newhart M, Trupe LA, Gomez Y, Cloutman L, Molitoris JJ, Davis C, Leigh R, et al. Asyntactic Comprehension, Working Memory, and Acute Ischemia in Broca's Area versus Angular Gyrus. Cortex. 2011; 48(10):1288–1297. [PubMed: 22079684]

Patil, U.; Hanne, S.; Vasishth, S.; Burchert, F.; De Bleser, R. Modeling aphasic sentence comprehension in a cue-based retrieval architecture. CUNY Sentence Processing Conference; 2012.

Shallice, T. From Neuropsychology to Mental Structure. Cambridge, UK: Cambridge University Press; 1988.

Tanenhaus MK, Spivey-Knowlton MJ, Eberhard KM, Sedivy JE. Integration of visual and linguistic information in spoken language comprehension. Science. 1995; 268:1632–1634. [PubMed: 7777863]

Thompson CK, Dickey M. The resolution and recovery of filler-gap dependencies in aphasia: Evidence from on-line anomaly detection. Brain and Language. 2004; 88:108–127. 1. [PubMed: 14698736]

Van Dyke J, Lewis RL. Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalysed ambiguities. Journal of Memory and Language. 2003; 49:285–316.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Highlights

- syntactic comprehension deficits in 61 pwa
- no evidence for task-independent structure specific effects
- discussion of criteria for dissociations
- discussion of implications for methods and conclusions in aphasiology

Table 1

Sentence Types

Sentence Type	**Example**
Active (A)	The girl followed the boy.
Passive (P)	The girl was covered by the father.
Three Noun Phrases (3NP)	The niece said that the woman touched the girl.
Pronoun (PRO)	The daughter said that the woman held her.
Reflexive (REF)	The uncle said that the boy covered himself.
Cleft Subject (CS)	It was the boy who scratched the man.
Cleft Object (CO)	It was the man who the boy scratched.
Subject Object (SO)	The brother who the sister followed kissed the woman.
Subject Subject (SS)	The boy who hugged the girl chased the woman.
Subject Subject and Pronoun (SSPRO)	The nephew who touched the man washed him.
Subject Object and Pronoun (SOPRO)	The woman who the girl held touched her.
Subject Subject and Reflexive (SSREF)	The uncle who held the boy covered himself.
Subject Object and Reflexive (SOREF)	The nephew who the man touched scratched himself.

Experimental-Baseline contrasts used to examine effects of structures

Passive: P-A

Object Extraction: CO-CS, SO-SS, SOPRO-SSPRO, SOREF-SSREF

Pronouns: PRO-3NP, SSPRO-SS, SOPRO-SO

Reflexives: REF-3NP, SOREF-SO, SSREF-SS

Table 2

A. Performance (accuracy) of people with aphasia on sentence comprehension tests. Sentence type labels as in Table 1 * indicates a significant difference between the experimental sentence annotated and its baseline (e.g., PRO and 3NP in sentence picture matching with auditory moving window presentation. + indicates a significant difference between SOPRO and SO and between SOREF and SO sentences.

	Sentence Picture Matching with Auditory Moving Window Presentation												
	A	**P**	**3NP**	**PRO**	**REF**	**CS**	**CO**	**SS**	**SO**	**SSPRO**	**SOPRO**	**SSREF**	**SOREF**
PWA Mean	82.1	79.9	84.6	74.5*	86.3	83.4	74.3*	75.7	63.1*	75.2	63.7*	71.9	63.5*
PWA std	18.4	20.2	16.0	21.0	21.0	18.4	23.0	20.1	22.4	21.0	20.6	22.5	19.9
EC Mean	94.8	94.1	97.9	93.1	98.7	95.7	94.4	93.4	91	96.6	88.4	95.8	83.7
EC std	6.1	8.3	5.8	4.4	3.7	6.7	6.2	9.7	8.9	6.9	14.8	7.9	13.1
	Sentence Picture Matching with Full Sentence Presentation												
	A	P	3NP	PRO	REF	CS	CO	SS	SO	SSPRO	SOPRO	SSREF	SOREF
PWA Mean	83.8	77.5*	81.7	74.5*	80.4	81.2	72.6*	70.0	66.5+	71.1	60.1*,+	69.3	61.0*,+
PWA std	15.3	18.4	17.3	21.9	23.8	18.6	20.4	17.8	14.8	19.8	20.5	20.5	17.7
EC Mean	95.1	93.6	97	96.3	97.2	95	92.6	92.1	82.3	95.1	82.4	94.2	83
EC std	3.6	6.5	5.5	5.7	4.8	4.5	6.6	8	13.4	6.1	14.1	8.9	12.7
	Object Manipulation												
	A	P	3NP	PRO	REF	CS	CO	SS	SO	SSPRO	SOPRO	SSREF	SOREF
PWA Mean	88.1	77.6*	66.3	68.8	67.0	88.2	71.8*	56.2	33.5*	53.3	36.8*	60.4	37.1*
PWA std	16.9	27.0	35.6	32.9	39.9	18.3	28.8	41.5	37.8	42.6	33.3	39.5	37.7
EC Mean	100	99	97.2	97.4	95.4	100	98.8	99	84.2	91.8	83.4	96.4	87.9
EC std	0	6.7	14.9	13.4	17.8	0	6	2.7	21.2	25.6	24.2	15.1	22.9

B. Performance (RT) of people with aphasia on sentence comprehension tests. Sentence type labels as in Table 1.

	Sentence Picture Matching with Auditory Moving Window Presentation												
	A	**P**	**3NP**	**PRO**	**REF**	**CS**	**CO**	**SS**	**SO**	**SSPRO**	**SOPRO**	**SSREF**	**SOREF**
PWA Mean	1178	1242	1016	1115	940	932	1038	1180	1472	1343	1522	1145	1383
PWA std	823	885	692	822	615	652	695	848	970	922	940	766	817
EC Mean	527	599	529	559	522	501	534	584	778	650	836	600	775
EC std	236	272	270	285	277	258	278	329	393	328	402	346	414

B. Performance (RT) of people with aphasia on sentence comprehension tests. Sentence type labels as in Table 1.

	Sentence Picture Matching with Auditory Moving Window Presentation												
	A	**P**	**3NP**	**PRO**	**REF**	**CS**	**CO**	**SS**	**SO**	**SSPRO**	**SOPRO**	**SSREF**	**SOREF**
	Sentence Picture Matching with Full Sentence Presentation												
	A	P	3NP	PRO	REF	CS	CO	SS	SO	SSPRO	SOPRO	SSREF	SOREF
PWA Mean	4944	5402	6480	6212	5486*	5762	6447*	7298	7927	7543	7996	7157	7789
PWA std	1674	1874	1911	1792	1430	1722	1868	2034	2215	2156	1940	1893	2204
EC Mean	3284	3591	4502	4465	4154	3990	4482	4831	4986	4795	5243	4565	4864
EC std	801	779	834	827	733	698	753	863	737	875	720	881	781

Table 3

Correlation of accuracy on sentences across tasks in pwa

Sent Type	OM-SPM Full	OM-SPM AMW	SPM Full - SPM AMW
A	0.80	0.76	0.83
P	0.75	0.81	0.80
CO	0.75	0.80	0.85
CS	0.72	0.76	0.73
3NP	0.83	0.83	0.78
PRO	0.79	0.79	0.79
REF	0.59	0.73	0.80
SO	0.57	0.77	0.71
SS	0.70	0.74	0.78
SOPRO	0.60	0.75	0.66
SSPRO	0.70	0.75	0.83
SOREF	0.74	0.79	0.76
SSREF	0.66	0.66	0.84

Table 4

Deficits in the experimental version of a sentence pair in a task based on three criteria:

1. accuracy within normal limits on the baseline sentence
2. accuracy BOTH below normal limits AND at/below chance on the experimental sentence
3. a 2×2 chi-square of response accuracy (#correct, #errors) and sentence type (exptl, base) was significant

Subject	Task	Sentence Pair	Number of Correct Experimental Sentences	Number of Correct Baseline Sentences	$\chi^2(2\times2)$	p
54001	OM	PRO-3NP	10	19	10.16	0.0014
54001	SPM-AMW	PRO-3NP	12	17	6.22	0.0126
54003	OM	SOPRO-SSPRO	8	16	6.67	0.0098
54003	OM	SOREF-SSREF	9	19	11.9	0.0006
54003	OM	SO-SS	4	19	23.02	0.0001
54003	SPM	SO-SS	11	17	4.29	0.0384
54003	SPM-AMW	SOREF-SSREF	11	20	10.59	0.0011
54004	OM	SOPRO-SSPRO	4	14	10.1	0.0015
54004	OM	SOREF-SSREF	6	19	18.03	0.0001
54004	SPM	SOPRO-SO	9	16	5.23	0.0222
54004	SPM-AMW	SOPRO-SSPRO	10	17	4.79	0.0286
54005	SPM	SO-SS	11	19	8.53	0.0035
54005	SPM	SSPRO-SS	13	19	5.63	0.0177
54009	OM	SOREF-SSREF	9	16	5.23	0.0222
54011	SPM	SOPRO-SO	4	13	8.29	0.004
54012	OM	SOREF-SSREF	3	20	29.57	0.0001
54015	SPM	SOREF-SSREF	12	20	10	0.0016
54016	SPM-AMW	SO-SS	10	17	5.58	0.0181
54019	OM	PRO-3NP	4	15	12.13	0.0005
54019	OM	REF-3NP	4	15	12.13	0.0005
54022	SPM	SORE-FSO	10	16	3.96	0.0467
54022	SPM	SOREF-SSREF	10	17	5.58	0.0181

Subject	Task	Sentence Pair	Number of Correct Experimental Sentences	Number of Correct Baseline Sentences	$\chi^2(2\times2)$	p
54023	SPM	SOPRO-SO	6	16	10.1	0.0015
54026	OM	SOPRO-SSPRO	8	17	8.64	0.0033
54026	OM	SOREF-SSREF	8	19	13.79	0.0002
54027	OM	CO-CS	5	20	24	0.0001
54029	OM	CO-CS	4	20	26.67	0.0001
54033	SPM	SOREF-SO	6	13	4.91	0.0267
54034	SPM	CO-CS	8	18	10.99	0.0009
54034	SPM-AMW	CO-CS	13	19	5.63	0.0177
54037	OM	SOPRO-SSPRO	2	14	15	0.0001
54037	SPM	CO-CS	7	20	19.26	0.0001
54041	OM	SOPRO-SSPRO	3	19	25.86	0.0001
54041	OM	SOREF-SSREF	3	20	29.57	0.0001
54041	OM	SO-SS	5	19	20.42	0.0001
54041	SPM	SOPRO-SSPRO	11	18	6.14	0.0132
54041	SPM-AMW	SOREF-SSREF	8	16	8.05	0.0046
54045	OM	SOPRO-SSPRO	6	15	8.12	0.0044
54045	OM	SOREF-SSREF	4	15	12.13	0.0005
54047	SPM-AMW	SOPRO-SSPRO	10	20	13.33	0.0003
54047	SPM-AMW	SO-SS	10		9.63	0.0019
54048	OM	SOPRO-SSPRO	1	16	23.02	0.0001
54048	OM	SOREF-SSREF	9	20	15.17	0.0001
54048	OM	SO-SS	8	19	13.79	0.0002
54048	SPM-AMW	CO-CS	13	20	8.48	0.0036
54048	SPM-AMW	SOPRO-SSPRO	12	18	4.8	0.0285
54048	SPM-AMW	SOREF-SSREF	9	17	7.03	0.008
54053	SPM-AMW	CO-CS	10	19	10.16	0.0014
54054	OM	SOPRO-SSPRO	7	17	10.42	0.0012
54054	OM	SOREF-SSREF	6	20	21.54	0.0001

Subject	Task	Sentence Pair	Number of Correct Experimental Sentences	Number of Correct Baseline Sentences	$\chi^2(2\times2)$	p
54054	SPM	CO-CS	10	18	7.62	0.0058
54054	SPM	SOREF-SSREF	7	16	8.29	0.004
54054	SPM-AMW	SOPRO-SSPRO	7	17	10.42	0.0012
54057	OM	SOPRO-SSPRO	3	20	29.57	0.0001
54057	OM	SOREF-SSREF	8	20	17.14	0.0001
54057	OM	SO-SS	8	20	17.14	0.0001
54057	SPM	SOPRO-SSPRO	10	17	5.58	0.0181
54059	SPM	SOPRO-SSPRO	9	18	9.23	0.0024
54065	OM	SOPRO-SSPRO	0	20	40	0.0001
54065	OM	SOREF-SSREF	2	20	32.73	0.0001
54065	OM	SO-SS	3	20	29.57	0.0001
54065	SPM-AMW	SO-SS	12	20	8.98	0.0027

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 5

Task specific deficits. A. Deficit in one task. B. Deficit in two tasks

A.

subj	**exptl**	**base**	**$\chi^2(3\times2)$**	**p**	**task deficit seen in**
54011	SOPRO	SO	9.2545	0.0098	SPM
54012	SOREF	SSREF	9.2297	0.0099	OM
54037	SOPRO	SSPRO	7.31	0.0259	OM
54065	SOPRO	SSPRO	11.5889	0.0030	OM
54065	SOREF	SSREF	8.8502	0.0120	OM

B.

subj	**exptl**	**base**	**$\chi^2(3\times2)$**	**p**	**tasks deficits seen in**
54041	SOREF	SSREF	5.8012	0.055*	OM, AMW
54048	SOPRO	SSPRO	8.8423	0.0120	OM, AMW
54057	SOPRO	SSPRO	8.0448	0.0179	OM,SPM
54065	SO	SS	8.9063	0.0116	OM, AMW

Table 6

Difference scores outside the normal range on three tasks in which performance on the baseline sentence was above chance. * indicates that performance on the baseline sentence was within the normal range

	A-P diff		
PWA	OM	SPM-Full	SPM-AMW
54001	25	50	25
54017	65	20	22.2
	CS-CO diff		
	OM	SPM-Full	SPM-AMW
54001	20	20	20
54003*	15	20	26.8
54027	75	30	48.3
54031	60	25	17.5
54034	55	50	30
54042	25	40	15
54054*	30	40	15
	SS-SO diff		
54003	75	30	25
	SSPRO-SOPRO diff		
	OM	SPM-Full	SPM-AMW
54041*	80	35	35
54054	55	45	50
	3NP-PRO diff		
	OM	SPM-Full	SPM-AMW
54001	45	20	34.4
54019*	50	30	15.8
	3NP-REF diff		
	OM	SPM-Full	SPM-AMW
54042	35	35	25.1

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 7

Performance (percent correct) of Cases 54001 and 54003 on sentences contrasting object- and subject extraction

Case 54001					
AMW-SPM		**Full-SPM**		**OM**	
CS	**CO**	**CS**	**CO**	**CS**	**CO**
70	50	75	55	95	75
SS	**SO**	**SS**	**SO**	**SS**	**SO**
75	50	65	70	25	5
SSPRO	**SOPRO**	**SSPRO**	**SOPRO**	**SSPRO**	**SOPRO**
70	55	75	75	30	25
SSREF	**SOREF**	**SSREF**	**SOREF**	**SSREF**	**SOREF**
60	50	60	50	65	35

Case 54003					
AMW-SPM		**Full-SPM**		**OM**	
CS	**CO**	**CS**	**CO**	**CS**	**CO**
70	50	75	55	95	75
SS	**SO**	**SS**	**SO**	**SS**	**SO**
75	50	65	70	25	5
SSPRO	**SOPRO**	**SSPRO**	**SOPRO**	**SSPRO**	**SOPRO**
70	55	75	75	30	25
SSREF	**SOREF**	**SSREF**	**SOREF**	**SSREF**	**SOREF**
60	50	60	50	65	35