Introduction

The goal of this study was to address three issues regarding aphasic disturbances of syntactic comprehension:

1. Are they best explained as resulting from reductions in the processing resources available for comprehension ("resource reduction models" [1]) or as resulting from disturbances affecting particular linguistic representations or operations ("specific deficit" models [2])?
2. If resource reduction models are correct, does resource availability differ for different tasks?
3. If resource reduction models are correct, can subgroups of patients be identified who differ in their resource availability?

We analyzed previously reported data [1] using Rasch models to provide information relevant to these questions.

Methods

Forty-two aphasic patients were tested for the ability to construct and understand three syntactic structures -- passives, relative clauses, and sentences with reflexive pronouns -- in object manipulation (OM) and picture matching (SPM). Each structure was tested with two experimental/baseline contrasts, with 10 examples of each sentence type (Table 1).

TABLE 1 HERE

We explored a series of models of aphasic disturbances of syntactic comprehension. In each model, Rasch models of the probability of a specified response (e.g. right/wrong answer) were developed on the basis of assumptions about the factors that determined patient ability and sentence difficulty. The simplest models assumed that each patient's ability was determined by his or her overall processing capacity; this can be conceived of as a single resource that applied to all sentences in both tasks. More complex models assumed that patients' abilities differed in the two tasks, for different sets of sentences, or for different groups of patients. Models that grouped sentences considered a grouping into the constructions listed above (passives, relative clauses, and sentences with reflexive pronouns, each set of sentences including the relevant baseline and experimental sentences), and other groupings that separated sentences with particular syntactic features from the corresponding baseline sentences (e.g., a grouping into four sets of sentences, consisting of the baseline sentences in one group and three sets of sentences each containing the experimental sentences used to test each structure). Models that grouped patients were based on Generalized Linear Model and Generalized Linear Mixed Model analyses of patient grouping.

Assessing the adequacy of each of the models tested was done in several ways. For the Rasch models that did not include patient groups, we determined the probability that features of the observed data (test statistics) fell within the distribution of these statistics in the set of matrixes generated by Monte Carlo simulation for each Rasch model. The first test statistic evaluated was the correlation of performance on the SPM and OM tasks. This test statistic was considered important because whether the observed correlation fell within the range of correlations predicted by the model tests the claim made by "specific deficit" models that patients have deficits that affect parsing or interpretive operations independently of task. The second test statistic was the frequency and magnitude of "reversals" of the number of correct performances on experimental and baseline sentences. Reversals refer to cases in which a patient's accuracy was greater on experimental than baseline sentences. This test statistic was considered an important feature of the performance data because it requires the presence of a random factor (reversals cannot be explained by either type of model without this additional assumption). We also examined several other test statistics that we do not report here, including the correlation of the difference between each experimental sentence and its baseline across tasks, the correlation of the number of responses that fell within the normal range of responses for each sentence type across tasks, this statistic considered differently for experimental and baseline sentences, and a statistic that counts the number of patients who correctly answered more questions on an experimental sentence than the corresponding baseline sentence without consideration of the magnitude of the difference. The overall pattern of results did not change when these statistics were considered. For models that included clusters of patients, properties of the clustering procedures prevented use of *p* values of the likelihood of occurrence of the test statistic in the simulation. Instead, we used a Bayesian approach (posterior predictive checking, which uses prior assumptions about the distribution of the probability of a test statistic) to assess whether these test statistics fell within the range generated by posterior distributions of each Rasch model.

The second set of estimates of the models' fit were based on the deviance of the model's performance from the observed data. Deviance is most easily assessed using the χ^2 test, but this test is less conservative than other tests. We therefore also used two other measures – the AIC (Akaike's information criterion [3]) and the BIC (Bayesian information criterion [4]). The AIC and BIC criterion address problems associated with the use of the χ^2 test by placing a penalty on each additional term added to the equation.

Implications of Outcomes

If specific deficit models are correct, Rasch models that postulate that patients' abilities differ for different sets of sentences will provide better fits to the observed data than models that postulate that patients' abilities are determined the same way for all sentences. If models that postulate different resource reductions for different tasks are correct, Rasch models that postulate that patients' abilities differ in different tasks will provide better fits to the observed data than models that postulate that patients' abilities are determined the same way in both tasks. If models that postulate different resource reductions for different patients are correct, Rasch models that include patient groupings will provide better fits to the observed data than models that do not.

Results

Space limitations preclude us from showing all the results here. Representative results regarding the effect of grouping sentences into sets are presented in Table 2 and 3.

TABLES 2 and 3 HERE

The critical results were as follows. Models that postulated different resource availability in different tasks were superior to those that did not. Models that included sentence type as a factor were inferior to ones that did not include a sentence type factor with respect to their ability to simulate test statistics. Models that postulated that the resource demand exerted by sentences differed for different groups of patients were superior to those that did not include patient groupings. The best models postulated both patient grouping and task effects, with three groups of patients in object manipulation and two in sentence picture matching. One quarter of the patients changed group membership in the two tasks.

Discussion

The results of these models support the view that performance of individuals with aphasia on tasks that require syntactic analysis is best understood as being determined by the ability of a patient to understand syntactic structures generally -- i.e., by reductions in the resources available for syntactic comprehension – not as deficits affecting particular syntactic structures. They further provide evidence that resource reduction differs in different tasks for some patients. Questions remain about the nature of the resource used in these tasks, in particular, its relation to other cognitive functions such as working memory, speed of processing, and inhibitory abilities.

References

1. Caplan, D., Waters, G.S., DeDe, G., Michaud, J., & Reddy, A. (2007): A study of syntactic processing in Aphasia I: Behavioral (psycholinguistic) aspects, Brain and Language, 101, 103-150.

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

3. Akaike, H (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19, 716–723

4. Schwarz, G., (1978). Estimating the dimension of a model. Annals of Statistics, 6, 461-464.

Table 1: Syntactic Structures and Sentence Types

Structure	Test Sentence	Baseline Sentence
Passive		
	The man was scratched by the boy	The man scratched the boy
	The man was scratched	(same as above)
Object Relatives		
	It was the man who the boy scratched	It was the boy who scratched the man
	The boy who the man scratched pushed the girl	The boy who scratched the man pushed the girl
Reflexives		
	The father of the boy scratched himself	The father of the boy scratched the girl
	The boy's father scratched himself	The boy's father scratched the girl

Table 2: Probability of Observed Correlation between Sentence Types in SPM and OM occurring in Rasch Models without a task factor and with no sentence type factor and two different groupings of sentence types. Sentence Grouping Type 1 grouped sentences into three experimental/baseline pairs ({A, PT, PF}, {CO, CS, SO, SS}, {RP, RPB, RG, RGB}). Sentence Grouping Type 2 grouped sentences into baseline sentences (A, CS, SS, RPB, RGB) and three groups of experimental sentences (PT, PF), (CO, SO) and (RP, RG). When the model is well fit to the date, the observed correlations fall within the range of correlations found in each model. Such correlations have non-significant *p* -values and are shown in bold italic.

		p value		
Sentence Type	Observed Correlation Across SPM and OM	No Sentence Type Grouping	Sentence Grouping Type 1	Sentence Grouping Type 2
A	0.572	0.036	0.036	0.0352
PF	0.654	0.044	0.041	0.0584
PT	0.138	***0.073***	***0.14***	0.013
RG	0.635	***0.136***	***0.09***	***0.14***
RGB	0.518	***0.82***	***0.66***	***0.84***
RP	0.434	***0.44***	***0.4***	***0.48***
RPB	0.52	***0.35***	***0.28***	***0.86***
CO	0.70	0.019	0.015	0
CS	0.518	***0.06***	***0.058***	0.056
SO	0.54	***0.94***	***0.73***	***0.25***
SS	0.45	***0.79***	***0.97***	***0.79***

Table 3: Probability of better Performance on Baseline than Experimental Sentence occurring in Rasch Models without a task factor and with no sentence type factor and two different groupings of sentence types. "Reverses" indicates average number of such occurrences; "Reverse Effect" indicates average of magnitude of these occurrences. Sentence Grouping Type 1 grouped sentences into three experimental/baseline pairs ({A, PT, PF}, {CO, CS, SO, SS}, {RP, RPB, RG, RGB}). Sentence Grouping Type 2 grouped sentences into baseline sentences (A, CS, SS, RPB, RGB) and three groups of experimental sentences (PT, PF), (CO, SO) and (RP, RG). When the model is well fit to the date, the observed correlations fall within the range of correlations found in each model. Such correlations have non-significant *p*-values and are shown in bold italic. SPM = Sentence-Picture Matching; OM = Object Manipulation (Enactment)

		p value		
	Reverses Value	No Sentence Type Factor	Sentence Type 1 Factor	Sentence Type 2 Factor
All Reverses	2.214	0.00005	0.0001	0
All Reverse Effect	4.52	0	0	0
SPM Reverses	1.21	0.0005	0.0001	0.0005
SPM Reverse Effect	2.21	0	0	0
OM Reverses	1	0.056	***0.11***	***0.24***
OM Reverse Effect	2.31	0.023	0.07	***0.059***