Introduction

People with aphasia (PWA) often have impairments in sentence comprehension, but information about stability of performance is limited because very few patients have been tested on the same sentence types on multiple tasks. This is problematic both for theoretical explanations of such impairments, which are sometimes based on deficit analyses in groups of patients (e.g., the Trace Deletion Hypothesis; Grodzinsky, 2000) and for clinicians, who frequently base initial evaluations and discharge testing on a single session.

In the present study, consistency of performance was assessed in a group of PWA who were tested on the same sentence types over 5 tasks. Split-half reliability was calculated to determine reliability on each task and on sentence types within tasks. Correlational analyses of accuracy and response time data for sentences within and across tasks examined stability of performance on sentence types across tasks. Finally, individual patient's performances were analyzed to identify patients with stable deficits in particular syntactic operations.

Methods

Participants

42 PWA with single left hemisphere strokes and 25 non-brain damaged controls participated in the study. All were native English speakers. Participant characteristics are given in table 1. PWA completed background testing to ensure adequate single word comprehension to complete the tasks and to characterize their aphasic symptoms.

Procedure

Each participant completed five tasks. Object manipulation (OM) and sentence picture matching (SPM-Whole) with whole sentence presentation assessed end-of-sentence comprehension. Grammaticality judgment with whole sentence presentation (GJ-Whole) assessed appreciation of grammatical well-formedness. On-line syntactic processing was assessed with the Auditory Moving Windows (AMW) technique (Ferreira et al., 1996). AMW presentation was used with two tasks -- sentence-picture matching and grammaticality judgment -- in separate experiments.

In OM, participants listened to sentences and indicated thematic roles and co-indexation by manipulating paper dolls. Responses were scored for accuracy.

In SPM-Whole, participants listened to sentences and chose the drawing that matched the sentence by pressing a button on a timer interfaced with the computer. Responses were scored for accuracy and reaction time (RT).

In the GJ-Whole, participants listened to sentences and indicated whether they were grammatical by pressing a button on a timer interfaced with the computer. Responses were scored for accuracy and RT.

In the on-line tasks (SPM-AMW & GJ-AMW), the participants paced their way through the sentences by pressing a button on the response box interfaced with the computer, and, depending on the task, did either SPM or GJ. RT's for each button press and accuracy on the associated task were recorded.

Here we report on end-of-sentence accuracy data from the 5 tasks.

Stimuli

Participants were tested on three types of constructions -- active/passive; subject/object relative; baseline sentences/sentences with reflexives -- using two pairs of baseline/experimental constructions for each contrast. Sentences were generated in pairs to control for effects of lexical

frequency and semantic meaning. There were 10 exemplars of each sentence type (examples are given in Table 2).

Pictures in the SPM tasks were line drawings depicting the actors and actions in the correct and reversed thematic roles (foils). In the GJ tasks, additional sentence types that violated syntactic rules of well-formedness were included in the experiment (see Table2).

Sentences were recorded and digitized using SoundEdit (Dunn, 1994). Stimuli for the AMW tasks were broken into words, also using SoundEdit. The waveforms were then entered into Psyscope (Cohen, MacWhinney, Flatt, & Provost, 1993) to create the experiment, which was run with a Macintosh PowerPC laptop.

Results

Within-task analyses

Split-half reliability was measured for each task as a whole (half the sentences of each type being assigned to the two halves) and for each sentence type within each task. For the tasks as a whole, Pearson's r and Spearman-Brown reliability coefficients were high ($r > .8$; see Table 3). For sentence types within each task, these measures were also generally significant. The few that were non-significant occurred in sentences with high overall accuracy, and probably reflected limited variance. The within-sentence type reliability measures were similar to the average correlations of accuracy on different sentence types within each task (see Table 4).

Between-task analyses

Correlation coefficients were calculated for sentences between three pairs of related tasks: the whole sentence version of SPM and OM, and the AMW and whole sentence versions of GJ and SPM. The mean r-values for the same sentences across these three tasks were significant, but they were only slightly higher than the averages of the r-values for different sentences across these tasks (for GJ AMW/full, $r_{same} = .61$, $r_{different} = .51$; for SPM AMW/full, $r_{same} = .55$, $r_{different} = .50$; for SPM full/OM, $r_{same} = .51$, $r_{different} = .46$).

Individual Patient Analyses

Since individual patients' accuracy on OM and SPM tasks has been the basis for most deficit analysis in this area of aphasiology, we looked for specific deficits in those tasks. A deficit was judged to be present if a patient scored within 2SD of the normal mean on a baseline sentence and below normal on the matched sentence with the construction for both tasks.

We looked for patients who were abnormal on 1) passives but not actives, corresponding to a deficit in the ability to construct and interpret passives (active vs. full and truncated passives), and 2) object extracted, but not subject extracted, relative clauses, corresponding to a deficit in the ability to construct and interpret object extracted relative clauses (CO vs. CS and SO vs. SS), 3) the combination of these two deficits, which would constitute a deficit in the ability to co-index traces, 4) sentences with reflexives and not the corresponding baseline sentences, reflecting a deficit in the ability to determine the reference of a pronoun. Twenty-two patients had deficits in one or more of these operations in OM, and twenty-two in SPM. One patient showed a consistent deficit in the ability to construct and interpret passives across both tasks, but examination of the RT data suggested that this may have been a speed/accuracy trade-off. No patients had consistent deficits in the ability to construct and interpret object extracted relative clauses, to co-index traces, or to determine the reference of a pronoun.

Discussion

The split-half reliability and correlation analyses suggest that performance is affected to a similar extent on most sentence types in most patients, because r-values were similar for the

same and different sentences across tasks, and because split-half reliability was similar to the mean r-value for different sentences within a task. These data suggest that factors other than the ability to assign structure and meaning in specific sentence types are important determinants of patients' performance.

These conclusions are supported by the analysis of individual patient's performances, where the pattern of accuracy differed both for sentence types that test the same operation (e.g., active vs. full and truncated passives), and for similar tasks (SPM and OM). This result is not what one might expect from some of the literature, which holds that many patients have specific deficits in one, or perhaps more than one, syntactic operation. The discrepancy between the results here and the conclusions in the literature is likely due to the larger database and the resultantly more stringent criteria that were applied to ascribe a specific deficit to a patient in this study.

References

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

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

Ferreira, F., Anes, M.D., Horine, M.D. (1996) Exploring the use of prosody during language comprehension using the auditory moving window technique. *Journal of Psycholinguistic Research*, *25*, 273-290.

Grodzinsky, Y. (2000). The neurology of syntax: Language use without Broca's Area. *Behavioral-and-Brain-Sciences*, *23(1)*, 1-71.

Table 1: Participant Information

	N	# Female	Age	Education
People with aphasia	42	16	Mean: 60.3 Range: 25-85	Mean: 14.7 Range: 9-22
Control	25	17	Mean: 68.9 Range: 53-90	Mean: 14 Range: 9-21

Table 2: Sentence Types

Abbreviation	Sentence Type	Example
A	Active	The father hit the man.
CO	Cleft Object	It was the mother who the girl kissed.
CS	Cleft Subject	It was the boy who tickled the aunt.
PF	Passive Full	The boy was kissed by the girl.
PT	Passive Truncated	The uncle was bitten.
RG	Reflexive Genitive	The wife of the man squeezed herself.
RGB	Reflexive Genitive Baseline	The brother of the woman tickled the wife.
RP	Reflexive Possessive	The girl's father hugged himself.
RPB	Reflexive Possessive Baseline	The woman's brother tickled the wife.
SO	Subject Object	The father who the girl hugged kicked the man.
SS	Subject Subject	The woman who squeezed the man followed the girl.
	Foils for Grammaticality Judgment Task	
Active Passive Unacceptable		The mother was kicked the boy.
Cleft Subject/ Cleft Object		It was the girl who the man hugged the father.
Reflexive Genitive		The sister of the man kissed himself.
Reflexive Possessive		The woman's brother tickled herself.
Subject Subject/ Subject Object		The girl who the man hugged the father kicked the man.

Table 3:Split Half Reliability

Task	Pearson r	Probability	Spearman-Brown	Range of Pearson r for Sentences in Task	Range of Spearman-Brown for Sentences in Task
GJ AMW	0.94	<.0001	0.971	.61-.85	.76-.92
GJ –Whole Sent	0.95	<.0001	0.976	.23-.78	.38-.88
SPM AMW	0.92	<.0001	0.958	.22-.76	.36-.87
SPM – Whole Sent	0.88	<.0001	0.938	.16-.73	.28-.84
OM	0.95	<.0001	0.976	.19-.92	.32-.96

Table 4: Mean split half reliability within sentence types vs. mean correlations between sentence types

Measure	GJ-AMW	GJ-Whole	SPM-AMW	SPM-Whole	OM
Within sentence type split half r (mean)	.72	.58	.55	.49	.55
Between sentence type r (mean)	.62	.65	.55	.65	.69