Test-retest reliability of the auditory Computerized Revised Token Test (CRTT) and three experimental reading CRTT-R versions in normal elderly individuals and persons with aphasia

Introduction

The formal definition of aphasia proposed by McNeil and Pratt (2001) specifies that the language deficits are expected to cross language processing modalities so that a person with aphasia (PWA) will show deficits both in primary input (reading and listening) and output (talking and writing) modalities: albeit to potentially different levels of severity and perhaps with differing underlying psycholinguistic mechanisms. Although listening and reading comparisons are essential for diagnosis, classification and treatment purposes, no psychometrically sound test has been published that allows for a direct comparison using the same test stimuli, performance tasks and scoring procedures; making direct comparisons between the modalities virtually impossible on a patient-by-patient basis.

While the *Revised Token Test* (*RTT*) (McNeil & Prescott, 1978) was originally developed as a test of auditory comprehension only, the theoretical basis for it is equally applicable to reading comprehension. Reading comprehension is a cognitive-linguistic task entailing similar task demands as auditory comprehension such as perceptual analysis and interpretation, lexical, semantic, and phonologic activation and mapping, among many other cognitive and linguistic processes. Most of the psycholinguistic variables that affect auditory comprehension also affect reading comprehension such as stimulus length, word frequency, semantic, and syntactic complexity. Recently, the *RTT* has been computerized (*CRTT*) which allows for increased control over test administration and scoring. With this increased control, there is a high potential for development of an equivalent reading version of the *CRTT*. With this long term goal in mind, three experimental reading versions were constructed which consisted of: 1) a full-sentence presentation (*CRTT-R-FS*); 2) a participant-paced word-by-word moving window, with all words remaining on the screen (word constant; *CRTT-R-WC*); and 3) a participant-paced word-by-word moving window presentation version, with each previous word disappearing with the onset of the following word (word fade; *CRTT-R-WF*). The *CRTT-R-WC* condition was based on the self-paced reading method with each word presented immediately after a touch in the textbox and with accumulation of the words across successive touches. The *CRTT-R-WF* condition was same as the *CRTT-R-WC* condition except for the word accumulation, and was designed to control for participants who selected all words before reading the sentence, thus diminishing important chronometric information available with the *CRTT-R-WF* condition, such as time spent on each word in the sentence. This version also served to increase the working memory load of the task, and was perhaps more equivalent to that encountered in the auditory presentation in the *CRTT*.

The purpose of the current study was to assess the test-retest reliability of all four versions in NEI and PWA.

Method

Thirty NEI and 30 PWA participated in the study. The NEI group ranged in age from 38 to 83 years (*mean=65, SD=12*), and passed hearing, vision, memory, and language screens, and reported no history of communication, neurological, or psychiatric disorder. The PWA ranged in age from 40 to 90 years (*mean*=63, *SD*=13), also passed the same

sensory, memory and psychiatric history criteria as the NEI and were defined by their performance on the *Porch Index of Communicative Ability* (*PICA*) (Porch, 1981), the C*RTT* (based on McNeil & Prescott, 1978) and on the immediate and delayed language recall task from the *Assessment Battery of Communication in Dementia* (Bayles & Tomoeda, 1993). All participants also completed the *Reading Comprehension Battery for Aphasia* (*RCBA*) (LaPointe & Horner, 1998) and the *PICA*. Biographical and selection data are summarized for PWA in **Table 1** and for NEI in **Table 2.**

All participants completed four conditions that consisted of the *CRTT* with the commands presented acoustically, and three versions of the *CRTT-R* with the commands presented visually through text. In the auditory condition, all of the commands were pre-recorded and presented acoustically via loudspeakers at 75 dB SPL as measured at the level of each participant's ear. In the reading conditions, the printed commands were presented in a textbox at the bottom of a touch-screen with the three different stimulus presentation methods describe above: 1) The *CRTT-R-FS*, 2) *CRTT-R-WC*, and 3) *CRTT-R-WF*. All of the participants were retested between one and four weeks after the initial test to assess reliability across the conditions for each group. All test and retest conditions were presented randomly across subjects following the collection of all criteria and descriptive information on the initial assessment. Data collection ranged from three to nine sessions across individual participants.

Results

Pearson correlation coefficients were computed between the first and second test administrations for the overall and subtest scores for each condition. As summarized in **Table 3**, correlations for the overall score for each *CRTT* and *CRTT-R* experimental condition was significant and high, ranging from .89 to .97 for the PWA. The overall score for each condition also correlated highly and significantly for the NEI (ranging from .74 to .78) with the exception of a significant but low-to-moderate correlation of .43 for the *CRTT* condition. Inspection of the data for this test-retest computation revealed a poor distribution of scores for this group compared to the PWA, especially on the *CRTT*, accounting for the lower correlations. All test-retest subtests scores for each of the four conditions were significantly and moderately to highly correlated for the PWA. Test-retest correlations for each subtest for the NEI were low and nonsignificantly different from zero for all subtests for the *CRTT* condition except for subtests 1 and 8, which were significantly different from zero but relatively moderate (.58 and .65 respectively). However, most of the subtests for the three reading conditions correlated significantly and moderately highly (exceptions were subtests 2 and 6 for the *CRTT-R-FS* condition and subtests 9 and 10 for the *CRTT-R-WF* condition).

Tables 4 and 5 summarize the test and retest performance scores, the differences between them and the standard error of measurement for each test and for the overall for the PWA and NEI respectively. Paired t-tests with Bonferroni adjusted alpha levels were calculated to examine the magnitude of the differences in performance between the test and retests for each condition for each group. Differences were nonsignificant for the overall (except for the *CRTT*) and each subtest (except for that of subtests 3 and 5 for *CRTT-R-WF*) for the PWA. Likewise, differences were also nonsignificant for the overall (except the CRTT-R-WF condition) and for each subtest except for that of subtest 1 for *CRTT-R-WF* for the NEI.

Discussion

The purpose of the study is to obtain test and retest reliability among the different types of the *CRTT*. Correlation coefficients between the test and retest for both the overall and subtest scores were high for each of the conditions for the PWA. Change scores between the administrations also were small and nonsignificant, suggesting high test-retest reliability for the *CRTT* and each of the *CRTT-R* versions for PWA. Change scores also were small and nonsignificant for the NEI participants; however, the correlations were low and nonsignificant for the *CRTT* condition due to a limited distribution of scores for this group. Overall, these data are interpreted as evidence for a highly reliable Computerized Revised Token Test and for all three experimental reading versions of the test. Additional data will also be discussed relative to the reliability of chronometric data for the self-paced reading versions of the *CRTT-R*.

References

Bayles, K.A. & Tomoeda, C.K. (1993). *Arizona Battery for Communication Disorders of Dementia*. Tucson, AZ: Canyonlands Publishing, Inc.

LaPointe, L. & Horner, J. (1978). *Reading Comprehension Battery for Aphasia*-2 (revised edition), Austin, TX: Pro-Ed. 1998.

McNeil, M.R. & Pratt, S.R. (2001). Defining aphasia: Some theoretical and clinical implications of operating from a formal definition, *Aphasiology, 15*, 901-911.

McNeil, M.R. & Prescott, T.E. (1978). *Revised Token Test*. Austin, Pro-Ed.

Porch, B.E. (1967). *The Porch Index of Communicative Ability*. Palo Alto, CA: Consulting Psychologists Press.

Table 1. Descriptive and selection scores for the Persons With Aphasia

PWA	PICA (%ile)	RCBA (OA)	Age (Yrs)	Education (Yrs)	MPO	Gender
1	73	178	63	14	456	F
2	76	172	66	12	192	M
3	49	158	70	12	96	F
4	66	181	72	14	444	F
5	72	178	60	16	24	M
6	86	185	66	13	25	F
7	52	179	45	16	13	F
8	84	184	49	16	71	F
9	66	186	61	16	15	F
10	76	179	65	12	201	M
11	53	160	38	14	25	M
12	57	166	76	12	564	F
13	88	187	43	14	91	M
14	69	174	62	16	60	M
15	89	190	53	18	88	F
16	71	176	69	10	453	F
17	71	162	59	12	24	M
18	88	182	56	18	31	M
19	29	86	83	12	30	F
20	69	166	40	18	12	M
21	89	190	51	18	139	F
22	83	189	90	12	58	M
23	59	184	63	18	46	M
24	70	157	70	12	29	F
25	76	179	82	16	106	M
26	41	102	77	12	MD	M
27	66	185	64	18	68	M
28	48	175	75	12	180	M
29	26	124	69	12	59	M
30	66	175	49	14	6	M
Mean	66.93	169.63	62.87	14.56	124.34	(F; 13/ M; 17)
SD	16.80	24.63	12.95	2.49	155.08	

MPO=Months Post Onset

MD = Unrecorded data with the average based on 29 participants

Table 2. Descriptive and selection scores for the NEI

ID	PICA (%ile)	RCBA (OA)	Age (Yrs.)	Education (Yrs.)	Gender
1	10	179	55	11	F
2	55	190	74	18	M
3	27	186	71	14	M
4	30	189	59	13	M
5	4	186	50	12	M
6	10	188	77	16	M
7	20	186	66	14	M
8	25	187	70	12	M
9	12	187	64	18	M
10	95	188	69	12	F
11	5	189	68	18	F
12	7	189	56	12	F
13	2	183	77	12	M
14	25	190	64	14	F
15	7	183	77	12	F
16	4	188	70	12	M
17	3	187	76	14	F
18	25	188	71	16	M
19	25	188	83	18	F
20	10	187	81	14	M
21	4	189	78	12	M
22	22	172	71	12	M
23	22	189	52	18	M
24	20	190	76	12	F
25	3	181	42	12	M
26	35	188	54	18	F
27	35	189	38	13	F
28	58	190	56	18	F
29	2	190	50	18	M
30	7	189	48	13	M
Mean	20.30	186.83	64.77	14.27	(F;12/ M; 18)
SD	20.42	3.89	12.14	2.56	

Table 3. Correlation coefficients of the subtest scores between the test and the retest per each CRTT condition

	NEI				PWA			
	CRTT	CRTT-R-FS	CRTT-R-WC	CRTT-R-WF	CRTT	CRTT-R-FS	CRTT-R-WC	CRTT-R-WF
Subtest_1	0.58**	0.37*	0.54**	0.50*	0.90**	0.82**	0.67**	0.58**
Subtest_2	0.23	0.21	0.39*	0.47*	0.87**	0.77**	0.72**	0.81**
Subtest_3	0.25	0.63**	0.59**	0.77**	0.88**	0.90**	0.82**	0.92**
Subtest_4	0.12	0.59**	0.57**	0.56**	0.79**	0.89**	0.80**	0.83**
Subtest_5	0.30	0.53**	0.77**	0.43*	0.83**	0.68**	0.89**	0.87**
Subtest_6	0.29	0.32	0.62**	0.54**	0.88**	0.82**	0.84**	0.98**
Subtest_7	0.06	0.62**	0.73**	0.66**	0.80**	0.73**	0.87**	0.93**
Subtest_8	0.65**	0.66**	0.57**	0.54**	0.81**	0.80**	0.89**	0.92**
Subtest_9	0.31	0.60**	0.34**	0.14	0.85**	0.70**	0.88**	0.78**
Subtest_10	0.30	0.57**	0.68**	-0.01	0.91**	0.69**	0.82**	0.88**
Mean	0.31	0.51	0.58	0.46	0.85	0.78	0.82	0.85
OA	0.43*	0.78**	0.74**	0.78**	0.96**	0.89**	0.94**	0.97**

*: significant at p=0.05
**: significant at p=0.01
OA=overall scores

Table 4. Test and Retest performance for the *CRTT* and *CRTT-R* conditions for the PWA

	CRTT (Auditory)				*CRTT-R-FS*				*CRTT-R-WC*				*CRTT-R-WF*			
	test	retest	Δ	SEM	test	retest	Δ	SEM	test	retest	Δ	SEM	test	retest	Δ	SEM
Subtest_1	13.69	13.83	0.14	0.43	13.42	13.71	0.29	0.37	13.09	13.21	0.11	0.27	12.95	13.20	0.25	0.42
Subtest_2	13.48	13.70	0.21	0.56	13.37	13.42	0.05	0.50	13.39	13.29	-0.10	0.38	13.32	13.51	0.19	0.36
Subtest_3	13.40	13.53	0.13	0.53	12.90	13.25	0.35	0.45	13.01	13.11	0.10	0.50	13.01	13.49	0.47**	0.48
Subtest_4	13.27	13.42	0.14	0.71	12.66	12.91	0.25	0.51	12.77	12.81	0.04	0.65	13.01	13.23	0.22	0.78
Subtest_5	12.42	12.74	0.32	0.69	12.40	12.58	0.18	0.63	12.31	12.70	0.39	0.41	12.52	12.99	0.46**	0.54
Subtest_6	12.41	12.64	0.23	0.65	12.14	12.37	0.23	0.46	12.24	12.41	0.17	0.50	12.45	12.79	0.34	0.22
Subtest_7	12.44	12.69	0.25	0.80	12.58	12.76	0.18	0.62	12.47	12.65	0.18	0.49	12.56	12.77	0.21	0.46
Subtest_8	12.45	12.77	0.31	0.79	12.28	12.60	0.32	0.54	12.52	12.38	-0.14	0.47	12.52	12.83	0.30	0.51
Subtest_9	13.43	13.72	0.29	0.57	13.20	13.06	-0.14	0.66	13.02	13.19	0.17	0.48	13.36	13.56	0.21	0.59
Subtest_10	13.47	13.42	-0.05	0.50	12.78	13.10	0.32	0.74	13.01	13.06	0.05	0.57	13.33	13.39	0.06	0.52
OA mean[a]	13.05	13.24	0.20	0.62	12.77	12.97	0.20	0.55	12.78	12.88	0.10	0.47	12.90	13.17	0.27*	0.49
SD[a]	0.54	0.48	0.11	0.12	0.45	0.42	0.15	0.11	0.38	0.34	0.15	0.10	0.37	0.31	0.08	0.15

OA=overall scores

SEM=Standard Error Measurement

mean[a] = weighted means

SD[a] = weighted standard deviations

Δ= difference between test and retest (scores from retest – scores from the first test)

*: significant (p<0.01)

**: significant (p<0.005)

Table 5. Test and Retest performance for the *CRTT* and *CRTT-R* conditions for the NEI

	CRTT (Auditory)				*CRTT-R-FS*				*CRTT-R-WC*				*CRTT-R-WF*			
	test	retest	Δ	SEM	test	retest	Δ	SEM	test	retest	Δ	SEM	test	retest	Δ	SEM
Subtest_1	14.74	14.89	0.15	0.18	14.43	14.68	0.25	0.15	13.50	13.69	0.19	0.24	13.23	13.64	0.42**	0.22
Subtest_2	14.84	14.94	0.10	0.22	14.70	14.70	0.00	0.13	14.01	14.21	0.20	0.30	13.91	14.10	0.19	0.28
Subtest_3	14.74	14.86	0.12	0.35	14.36	14.33	-0.03	0.18	13.95	14.14	0.20	0.29	14.23	14.49	0.26	0.24
Subtest_4	14.59	14.79	0.20	0.36	13.90	13.93	0.03	0.43	13.90	14.19	0.29	0.45	14.44	14.63	0.19	0.26
Subtest_5	14.37	14.53	0.16	0.20	13.67	13.55	-0.12	0.44	13.53	13.74	0.21	0.23	13.82	14.29	0.47	0.41
Subtest_6	14.17	14.44	0.26	0.35	13.51	13.60	0.09	0.48	13.56	13.69	0.13	0.35	14.07	14.35	0.28	0.26
Subtest_7	14.39	14.63	0.24	0.32	13.69	13.83	0.14	0.36	13.46	13.59	0.13	0.22	13.88	14.23	0.35	0.39
Subtest_8	14.40	14.73	0.33	0.21	13.59	13.59	0.00	0.31	13.52	13.64	0.13	0.45	14.18	14.27	0.09	0.32
Subtest_9	14.83	14.83	0.01	0.08	14.22	13.99	-0.23	0.40	14.54	14.43	-0.11	0.42	14.58	14.64	0.06	0.41
Subtest_10	14.82	14.75	-0.07	0.10	13.92	13.89	-0.03	0.41	14.31	14.48	0.17	0.29	14.75	14.52	-0.23	0.44
OA mean[a]	14.59	14.74	0.15	0.24	14.00	14.01	0.01	0.33	13.85	14.00	0.15	0.32	14.14	14.32	0.18*	0.32
SD[a]	0.24	0.16	0.12	0.10	0.40	0.43	0.13	0.13	0.38	0.34	0.10	0.09	0.44	0.30	0.20	0.08

OA=overall scores

SEM=Standard Error Measurement

mean[a] = weighted means

SD[a] = weighted standard deviations

Δ= difference between test and retest (scores from retest – scores from the first test)

*: significant (p<0.01)

**: significant (p<0.005)