

Supplemental Table S10. Summary of test-retest results for the Cinderella fictional narrative.

Given that sample length (in both subject groups) and aphasia severity (Aphasia group only) may affect discourse measure reliability, here we stratify ICC, Spearman's rho, and Wilcoxon's signed rank test *p*-value by each factor. Sample length is the average sample length across all discourse tasks (*n* = 12 in each sample length group for NBD group (Med = 511); *n* = 11 in long sample and *n* = 12 in short sample length group for Aphasia group (Med = 264.5)).

Proxy	Group	Measure	Long sample length	Short sample length	Mild or Latent Severity (<i>n</i> = 14)	Moderate or Severe Severity (<i>n</i> = 9)	Latent Aphasia (<i>n</i> = 6)	Clinical Aphasia (<i>n</i> = 17)
			ICC (95% CI) [Spearman's rho] Wilcoxon <i>p</i> -value	ICC (95% CI) [Spearman's rho] Wilcoxon <i>p</i> -value	ICC (95% CI) [Spearman's rho] Wilcoxon <i>p</i> -value	ICC (95% CI) [Spearman's rho] Wilcoxon <i>p</i> -value	ICC (95% CI) [Spearman's rho] Wilcoxon <i>p</i> -value	ICC (95% CI) [Spearman's rho] Wilcoxon <i>p</i> -value
Lexical and informativeness	NBD	%CIU	0.68 (0.22, 0.90) [0.63], <i>p</i> = .23	0.59 (0.09, 0.86) [0.50], <i>p</i> = .13	N/A	N/A	N/A	N/A
	Aphasia		0.04 (-0.42, 0.56) [0.19], <i>p</i> = .08	0.93 (0.63, 0.98) [0.89], <i>p</i> = .009	0.45 (-0.03, 0.77) [0.59], <i>p</i> = .04	0.83 (0.31, 0.96) [0.88], <i>p</i> = .02	0.28 (-0.50, 0.85) [0.20], <i>p</i> = .31	0.87 (0.58, 0.96) [0.76], <i>p</i> = .007
	NBD	PI Density	-0.09 (-0.70, 0.51) [-0.04], <i>p</i> = .81	0.37 (-0.27, 0.77) [0.35], <i>p</i> = .68	N/A	N/A	N/A	N/A
	Aphasia		0.71 (0.24, 0.91) [0.83], <i>p</i> = .11	0.76 (0.38, 0.92) [0.89], <i>p</i> = .52	0.53 (-0.002, 0.82) [0.58], <i>p</i> = .80	0.77 (0.27, 0.94) [0.89], <i>p</i> = .13	0.22 (-0.91, 0.85) [0.37], <i>p</i> = .69	0.79 (0.52, 0.92) [0.89], <i>p</i> = .46
	NBD	TTR	0.60 (0.12, 0.86) [0.47], <i>p</i> = .15	0.64 (0.16, 0.88) [0.62], <i>p</i> = .06	N/A	N/A	N/A	N/A
	Aphasia		0.62 (0.02, 0.88) [0.75], <i>p</i> = .02	0.81 (0.47, 0.94) [0.66], <i>p</i> = .12	0.68 (0.20, 0.89) [0.70], <i>p</i> = .04	0.94 (0.76, 0.99) [0.93], <i>p</i> = .18	0.64 (-0.10, 0.94) [0.83], <i>p</i> = .06	0.91 (0.76, 0.97) [0.92], <i>p</i> = .07
	NBD	Tokens	0.65 (0.18, 0.88) [0.45], <i>p</i> = .27	0.58 (0.002, 0.86) [0.81], <i>p</i> = .01	N/A	N/A	N/A	N/A
	Aphasia		0.67 (-0.05, 0.91) [0.59], <i>p</i> = .004	0.78 (0.42, 0.93) [0.86], <i>p</i> = .21	0.77 (0.27, 0.93) [0.71], <i>p</i> = .01	0.96 (0.77, 0.99) [0.996], <i>p</i> = .04	0.63 (-0.11, 0.94) [0.82], <i>p</i> = .04	0.92 (0.75, 0.97) [0.97], <i>p</i> = .02

Proxy	Group	Measure	Long sample length ICC (95% CI) [Spearman's rho] Wilcoxon <i>p</i> -value	Short sample length ICC (95% CI) [Spearman's rho] Wilcoxon <i>p</i> -value	Mild or Latent Severity (<i>n</i> = 14) ICC (95% CI) [Spearman's rho] Wilcoxon <i>p</i> -value	Moderate or Severe Severity (<i>n</i> = 9) ICC (95% CI) [Spearman's rho] Wilcoxon <i>p</i> -value	Latent Aphasia (<i>n</i> = 6) ICC (95% CI) [Spearman's rho] Wilcoxon <i>p</i> -value	Clinical Aphasia (<i>n</i> = 17) ICC (95% CI) [Spearman's rho] Wilcoxon <i>p</i> -value
Fluency / efficiency	NBD		0.78 (0.39, 0.93) [0.84], <i>p</i> = .91	0.41 (-0.17, 0.79) [0.36], <i>p</i> = .62	N/A	N/A	N/A	N/A
	Aphasia	CIUs/min	0.64 (0.07, 0.89) [0.49], <i>p</i> = .28	0.31 (-0.28, 0.74) [0.70], <i>p</i> = .38	0.92 (0.77, 0.97) [0.87], <i>p</i> = .17	0.55 (-0.11, 0.88) [0.63], <i>p</i> = .57	0.83 (0.15, 0.98) [0.94], <i>p</i> = .84	0.59 (0.17, 0.83) [0.64], <i>p</i> = .22
	NBD		0.65 (0.19, 0.89) [0.47], <i>p</i> = .23	0.49 (-0.05, 0.82) [0.76], <i>p</i> = .02	N/A	N/A	N/A	N/A
	Aphasia	SpeakingSecs	0.71 (0.19, 0.92) [0.85], <i>p</i> = .02	0.86 (0.58, 0.96) [0.74], <i>p</i> = .79	0.72 (0.30, 0.90) [0.74], <i>p</i> = .06	0.94 (0.75, 0.99) [0.87], <i>p</i> = .95	0.74 (-0.05, 0.96) [0.89], <i>p</i> = .06	0.85 (0.63, 0.94) [0.89], <i>p</i> = .54
	NBD		0.76 (0.36, 0.93) [0.65], <i>p</i> = .57	0.42 (-0.19, 0.79) [0.44], <i>p</i> = .91	N/A	N/A	N/A	N/A
	Aphasia	WPM	0.88 (0.59, 0.97) [0.73], <i>p</i> = .10	0.91 (0.66, 0.98) [0.92], <i>p</i> = .02	0.91 (0.75, 0.97) [0.90], <i>p</i> = .15	0.96 (0.68, 0.99) [0.93], <i>p</i> = .03	0.57 (-0.18, 0.92) [0.37], <i>p</i> = .56	0.96 (0.85, 0.99) [0.95], <i>p</i> = .008
Syntactic	NBD		0.64 (0.12, 0.88) [0.41], <i>p</i> = .57	0.38 (-0.24, 0.77) [0.56], <i>p</i> = .30	N/A	N/A	N/A	N/A
	Aphasia	Noun/verb	0.83 (0.43, 0.95) [0.81], <i>p</i> = .05	0.66 (0.15, 0.90) [0.65], <i>p</i> = .02	0.91 (0.74, 0.97) [0.65], <i>p</i> = .76	0.60 (-0.02, 0.90) [0.74], <i>p</i> = .31	0.38 (-0.30, 0.87) [0.54], <i>p</i> = .16	0.71 (0.35, 0.89) [0.74], <i>p</i> = .14
	NBD		-0.06 (-0.57, 0.50) [0.12], <i>p</i> = .31	0.51 (-0.10, 0.83) [0.51], <i>p</i> > .99	N/A	N/A	N/A	N/A
	Aphasia	Open/closed	0.74 (0.32, 0.92) [0.74], <i>p</i> = .28	0.57 (-0.05, 0.86) [0.47], <i>p</i> = .76	0.66 (0.23, 0.88) [0.73], <i>p</i> = .81	0.60 (-0.19, 0.91) [0.26], <i>p</i> > .99	0.87 (0.35, 0.98) [0.83], <i>p</i> = .69	0.58 (0.12, 0.83) [0.40], <i>p</i> = .90

Proxy	Group	Measure	Long sample length	Short sample length	Mild or Latent Severity ($n = 14$)	Moderate or Severe Severity ($n = 9$)	Latent Aphasia ($n = 6$)	Clinical Aphasia ($n = 17$)
			ICC (95% CI) [Spearman's rho] Wilcoxon p -value	ICC (95% CI) [Spearman's rho] Wilcoxon p -value	ICC (95% CI) [Spearman's rho] Wilcoxon p -value	ICC (95% CI) [Spearman's rho] Wilcoxon p -value	ICC (95% CI) [Spearman's rho] Wilcoxon p -value	ICC (95% CI) [Spearman's rho] Wilcoxon p -value
	NBD	MLU	0.65 (0.13, 0.89) [0.57], $p = .85$	0.50 (-0.10, 0.83) [0.55], $p = .97$	N/A	N/A	N/A	N/A
	Aphasia		0.59 (0.07, 0.87) [0.66], $p = .15$	0.94 (0.79, 0.98) [0.91], $p = .73$	0.78 (0.44, 0.92) [0.61], $p = .72$	0.94 (0.77, 0.99) [0.88], $p = .30$	0.01 (-0.52, 0.71) [-0.20], $p = .22$	0.95 (0.88, 0.98) [0.86], $p = .89$
	NBD	VerbUtt	0.59 (0.03, 0.86) [0.51], $p = .97$	0.48 (-0.10, 0.82) [0.49], $p = .23$	N/A	N/A	N/A	N/A
	Aphasia		0.32 (-0.21, 0.74) [0.33], $p = .15$	0.90 (0.70, 0.97) [0.87], $p = .38$	0.70 (0.28, 0.89) [0.45], $p = .81$	0.93 (0.72, 0.98) [0.87], $p = .82$	0.07 (-0.27, 0.68) [-0.03], $p = .09$	0.91 (0.78, 0.97) [0.82], $p = .40$

Koo and Li (2016) gives the following suggestion for interpreting intraclass correlation coefficient (ICC), including confidence intervals: below 0.50 = poor; between 0.50 and 0.75 = moderate; between 0.75 and 0.90 = good; and above 0.90 = excellent.