Acquisition, Maintenance and Generalization of Script Training: A Comparison of Errorful and Error-Reducing Conditions

Relatively few studies have investigated errorless versus errorful learning in aphasia, and most have focused on the single word level. In a review of a series of anomia studies, Fillingham et al. (2003) found that rate of success was equivalent for errorless and errorful conditions; the number of therapies using errorful techniques outweighed those based on errorless learning; errorless approaches were likely to achieve positive immediate effects, but many of the errorless studies did not report long term effects and generalization; and there was a lack of studies reporting a direct comparison of errorful and errorless learning.

More recent work has continued to support the findings that there is no essential difference between error-free and errorless learning approaches (Middleton & Schwartz, 2012). However, none have investigated and contrasted the errorless/errorful learning paradigm in the production of phrases and sentences.

This pilot study directly measures and compare the acquisition, maintenance and generalization of script-training under both error-reducing and errorful conditions. We hypothesized that error-reducing training would improve acquisition of a trained script while errorful learning would improve maintenance of the trained script and generalization to untrained scripts

<u>Methods</u>

Design: A cross-over study with participants randomized to one of two script training conditions: error-reducing and errorful.

Participants: Four right-handed individuals with chronic aphasia due to a left-hemisphere stroke participated. Table 1 shows their demographic data and aphasia severity based on results of the Western Aphasia Battery-Revised.

Treatment was provided using a computer program, *AphasiaScripts*™ . Script training was delivered by an anthropomorphically accurate "digital" therapist capable of visually modeling speech and interactively guiding treatment, thereby ensuring treatment fidelity by removing clinician-related variables (e.g., clinician expertise, personality factors) that potentially influence treatment outcomes. The treatment software has experimental support regarding its efficacy (references omitted for anonymity). Furthermore, it accommodates the manipulation of variables such as cueing, allowing comparison of different conditions as described below:

1. Error-reducing condition (more help/cues): AphasiaScripts was administered in its current form, with all cues and supports provided prior to attempts at production.	2. Errorful condition (less help/cues): AphasiaScripts was modified so that during sentence practice and conversation practice, only written sentences were provided, without any auditory or oral-motor cues from the digital therapist. The PWA attempted to produce the phrases/sentences on his/her own for three trials. Only at the end of the practice sequence did the digital therapist provide a correct model, allowing the PWA to listen to the correct productions and watch the oral motor movements.

We developed six scripts of equal length and grammatical complexity with each PWA - three for the error-reducing condition and three for the errorful condition. In each condition, one script served as the trained script, one as the untrained script, and one as a generalization script. All scripts were dialogues, with the PWA as the responder on a topic of personal interest.

Training in each condition continued for three weeks, with a three-week washout period between. Script training occurred 6 days a week with PWAs practicing for three 30-minute sessions per day at home on a loaned laptop. Treatment and probe sessions were set up automatically via a calendar function integrated into the computer software.

During probes sessions, the lines of the script appeared on the screen and the PWA was required to read them aloud without any cues from the digital therapist. Trained and untrained scripts were probed at baseline three times, twice weekly during the training phase, at the end of the three weeks of treatment, and at the three-week maintenance visit. Generalization scripts were probed only at baseline, at the end of the three weeks of treatment, and at the three week maintenance visit.

Audio-recordings of the probes were captured by the computer software. One of three clinicians transcribed and scored production of script-related words on a 6-point scale that ranged from 0 (no response) and 1 (unintelligible or unrelated response) to 5 (accurate and immediate response). Rate was defined as the number of script-related words produced per minute. Ten percent of the script production probes, including trained, untrained, and generalization probes, were randomly selected for scoring by a second rater. Inter-rater reliability for accuracy of script related words was 94%. Timing was automatically calculated by the computer software and therefore rate was deemed to be reliable.

Effect sizes were calculated by dividing the difference between the mean of three baseline probes and the post-treatment or maintenance probe by the standard deviation for the baseline scores (Beeson & Robey, 2006).

Results

Results for all subjects are shown in Tables 2 and 3. Figures 1 and 2 are an example of the plotted data from one subject. Table 3 summarizes the comparison of the effect sizes data.

With the trained script, there were improvements in both conditions for both accuracy and rate of production of script-related words for all participants. At three weeks post treatment, accuracy and rate scores continued to be greater than at baseline, although there was typically a slight reduction for percent accuracy, and some variability in both directions for rate when compared to the end of treatment.

When effect sizes were compared across conditions for the trained script, three of four subjects showed greater change in rate in the error-reducing (more help) condition while three of four subjects showed greater change in accuracy in the errorful (less help) condition.

For the untrained and generalization scripts, comparison of effect sizes across conditions showed different patterns across subjects, outcomes, and time. For example, both ABEJO and AMBDE favored the error-reducing condition for measures of both accuracy and rate from baseline to the end of treatment. However, when performance at the three-week follow-up was considered, ABEJO tended to favor the errorful condition (less help) for rate but not accuracy, while AMBDE favored the error-reducing condition for both accuracy and rate.

Discussion

This study extends previous work in aphasia by directly comparing error-reducing and errorful conditions in a script-training task and including assessment of maintenance and generalization as well as acquisition.

Consistent with previous work, no clear differences between errorful and error-reducing learning approaches emerged. Several factors unrelated to the error-reducing and errorful conditions may have affected results, including severity and type of aphasia. For example, participants presented with aphasia on the milder end of the continuum of severity. Different results may occur with participants with more severe aphasia. Protocol factors such as outcome measures (i.e. accuracy versus rate) and the timing of when these measures were taken (at the end of treatment versus 3-week maintenance) may also have impacted the results.

Interestingly, feedback has been identified as an important treatment variable in aphasia. McKissock & Ward (2007) found that both errorless and errorful conditions with feedback were equivocal, and both were significantly better than errorful learning without feedback. They concluded that whether or not the PWA makes errors during training is not important; of importance is that a correct response is given as feedback. Further discussion of the importance of feedback will be addressed as well as the implications for clinical practice.

Acknowledgement

This study was supported by Grant 1R01 DC011754 from the National Institute on Deafness and Communication Disorders.

References

Beeson, P.M., Robey, R.R. (2006). Evaluating single-subject treatment research: lessons learned from the aphasia literature. Neuropsychology Review, 16(4), 161-9.

Fillingham, J. K., Hodgson, C., Sage, K., & Lambon Ralph, M. A. (2003). The application of errorless learning to aphasic disorders: A review of theory and practice. Neuropsych Rehabil, 13, 337-363.

McKissock, S. & Ward, J. (2007). Do errors matter? Errorless and errorful learning in anomic picture naming. Neuropsychological Rehabilitation, 17, 355-373.

Middleton, E. L. & Schwartz, M. F. (2012). Errorless learning in cognitive rehabilitation: A critical review. Neuropsychological Rehabilitation: An international Journal, 22, 138-168.

Table 1. Participant demographics

Subject	Age (years)	TPO (months)	Education	Aphasia Quotient	Language Quotient	Cortical Quotient	Aphasia Type	Treatment Sequence
ABEJO	51.8	48	16	68.8	73.6	76.70	NonFluent	2-1
AMBDE	61.6	21	11	74.5	72.9	77.85	Fluent	1-2
PIESH	66.4	59	14	67.6	63.9	68.83	NonFluent	1-2
WELED	64.5	37	14	80.1	85.6	86.37	NonFluent	2-1

1=error-reducing condition (more help)
2=errorful condition (less help)

Table 2

Descriptive Data for Error-Reducing (More Help) and Errorful (Less Help) Conditions for Words per Minute (wmp)

		Error-Reducing			Errorful		
Subject		T	U	G	T	U	G
ABEJO							
	BL *M*	17.2	18.0	13.2	11.8	8.6	7.5
	BL *SD*	1.7	1.3	0.5	0.2	1.9	0.4
	Tx *M*	59.1	27.2	--	43.1	11.0	--
	Final Tx day probe	79.3	25.0	18.6	68.4	12.2	10.9
	3-wk Maint. probe	66.9	18.9	14.8	65.7	20.6	11.3
AMBDE							
	BL *M*	31.5	28.1	32.4	42.2	43.8	44.9
	BL*SD*	2.5	0.9	5.3	5.2	3.9	2.1
	Tx *M*	46.7	29.2	--	49.6	43.9	--
	Final Tx day probe	53.1	40.5	32.5	50.7	45.0	44.0
	3-wk Maint. probe	53.6	37.9	43.1	52.0	40.2	31.1
PIESH							
	BL *M*	31.3	32.7	40.9	41.5	47.6	39.9
	BL *SD*	5.0	6.8	7.9	10.1	3.6	1.2
	Tx *M*	56.8	41.9	--	62.9	49.7	--
	Final Tx day probe	73.7	35.9	45.5	71.9	46.8	43.3
	3-wk Maint. probe	76.7	54.9	46.8	NA	NA	NA
WELED							
	BL *M*	57.0	56.1	56.1	49.1	46.2	47.0
	BL *SD*	0.9	5.3	5.6	3.4	7.0	3.0
	Tx *M*	111.9	68.1	--	101.5	62.3	--
	Final Tx day probe	121.9	54.9	55.2	111.3	60.3	54.0
	3-wk Maint. probe	110.6	75.9	67.1	118.1	61.1	56.1

Note. *M* = mean, *SD* = standard deviation, T = trained script, U = untrained script, G = generalization script, BL= baseline phase, Tx = treatment phase, NA = not available (probes scheduled for 1/29/12).

Table 3

Descriptive Data for Error-Reducing (More Help) and Errorful (Less Help) Conditions for Percent Accuracy

		Error-Reducing			Errorful		
Subject		T	U	G	T	U	G
ABEJO							
	BL *M*	61.7	60.0	63.5	60.3	56.0	57.7
	BL *SD*	8.6	1.7	2.1	5.7	8.7	6.5
	Tx *M*	87.4	70.9	--	79.9	63.4	--
	Final Tx day probe	97.0	69.0	71.0	91.0	63.0	70.0
	3-wk Maint. probe	91.0	75.0	68.0	88.0	66.0	64.0
AMBDE							
	BL *M*	62.0	59.7	66.3	71.0	76.3	74.3
	BL*SD*	6.9	7.6	4.0	1.7	6.0	2.5
	Tx *M*	81.9	71.1	--	80.1	74.1	--
	Final Tx day probe	79.0	75.0	68.0	84.0	68.0	70.0
	3-wk Maint. probe	72.0	61.0	68.0	75.0	64.0	53.0
PIESH							
	BL *M*	69.7	73.7	75.3	84.3	80.6	75.0
	BL *SD*	7.5	3.5	4.7	1.5	3.5	3.6
	Tx *M*	86.1	77.6	--	88.1	79.6	--
	Final Tx day probe	96.0	68.0	84.0	94.0	82.0	81.0
	3-wk Maint. probe	93.0	84.0	81.0	NA	NA	NA
WELED							
	BL *M*	82.7	85.0	87.7	84.7	83.0	85.7
	BL *SD*	2.5	0.1	3.1	3.8	5.2	3.1
	Tx *M*	89.6	83.4	--	91.9	83.9	--
	Final Tx day probe	94.0	76.0	89.0	96.0	85.0	83.0
	3-wk Maint. probe	91.0	89.0	87.0	92.0	85.0	83.0
	ES BL-3-wk Maint.	3.3	40	-0.2	1.9	0.4	-0.9

Note. *M* = mean, *SD* = standard deviation, T = trained script, U = untrained script, G = generalization script, NA = not available (probes scheduled for 1/29/12).

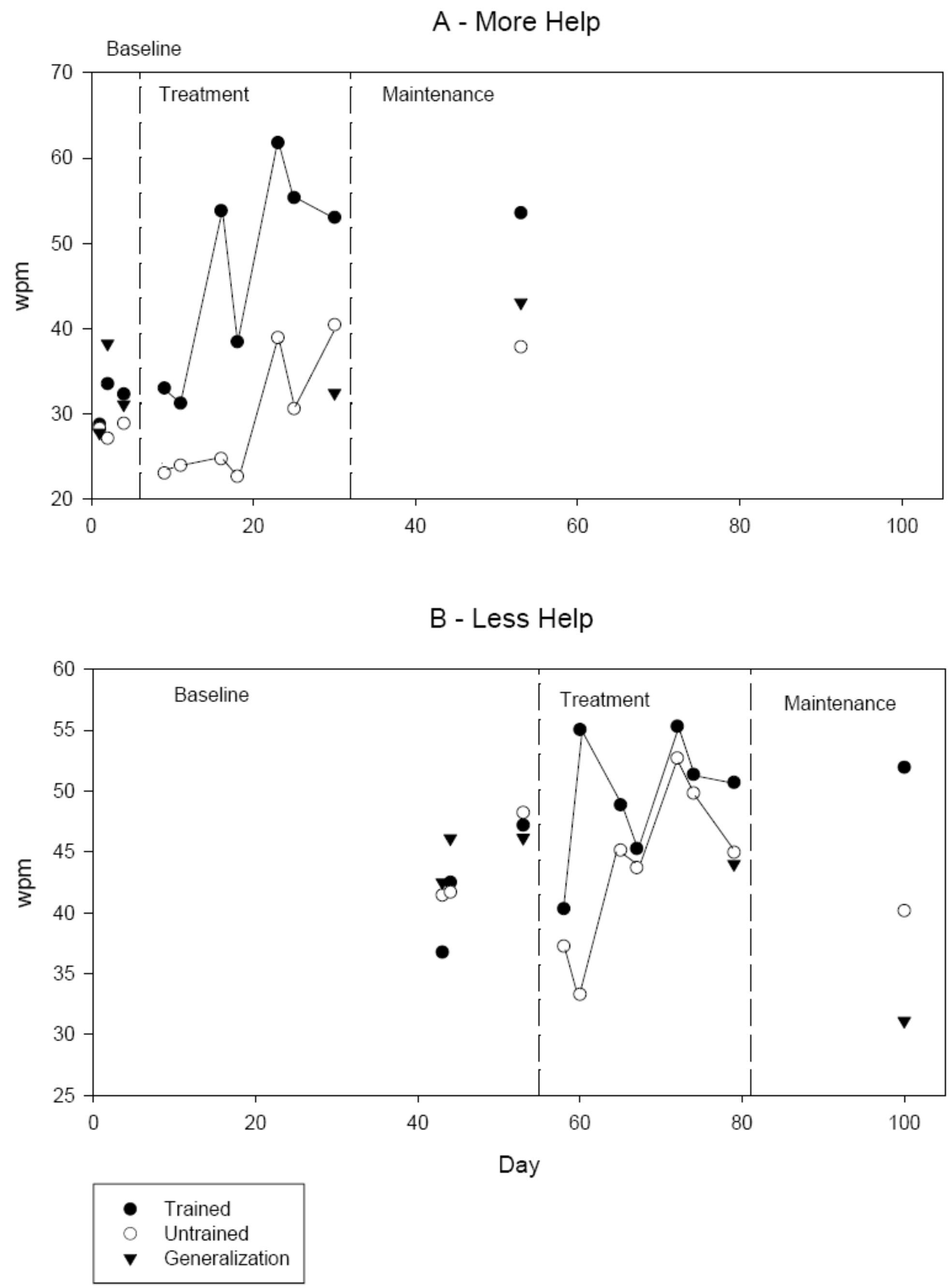
Table 4
Comparison of Effect Size Data for Error-Reducing (More Help) and Errorful (Less Help) Conditions

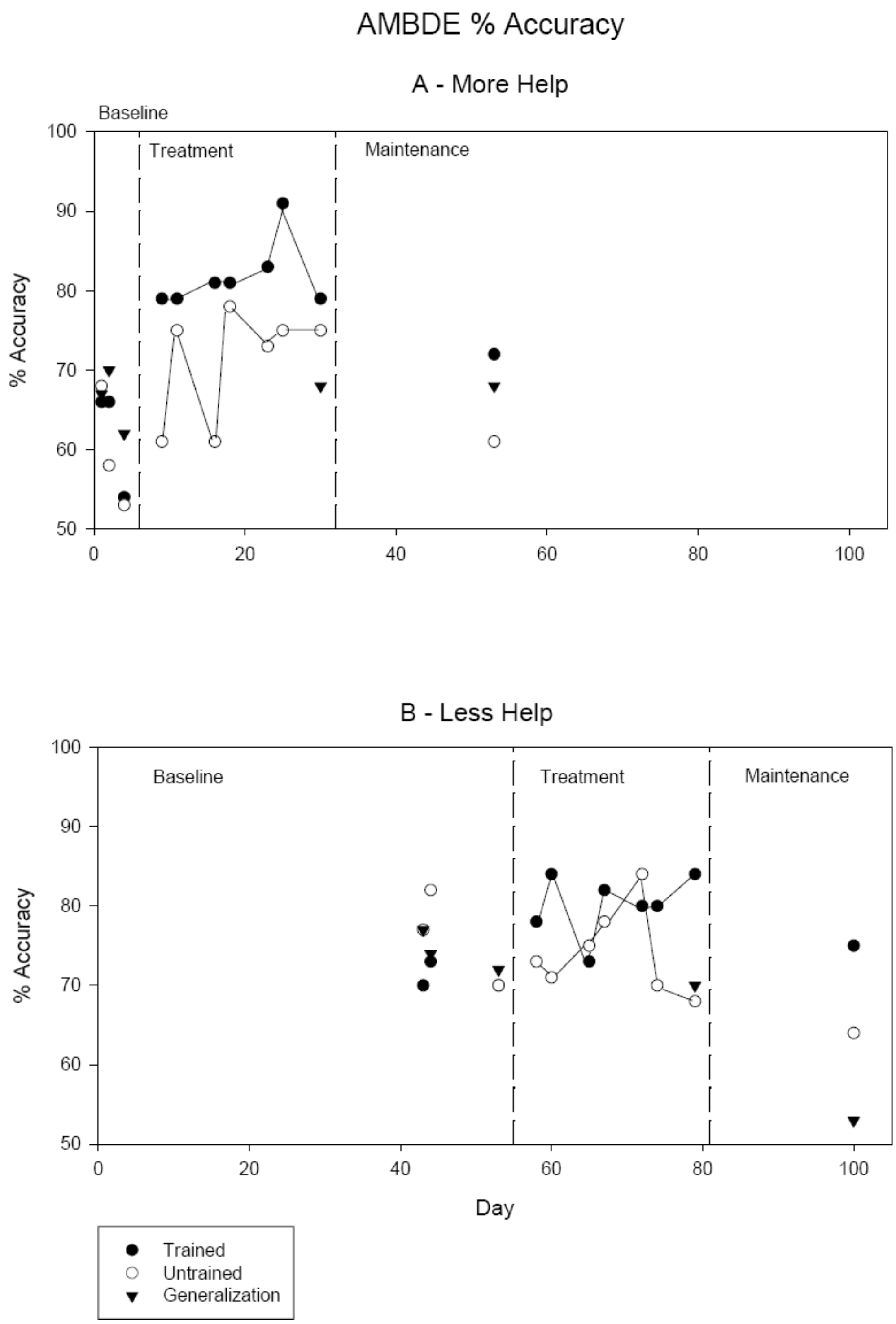
		Words per Minute			Percent Accuracy		
Subject		T	U	G	T	U	G
ABEJO							
	BL to Final Tx day probe	2	1	1	2	1	1
	BL to 3-wk Maint. probe	2	2	2	2	1	1
AMBDE							
	BL to Final Tx day probe	1	1	1	2	1	1
	BL to 3-wk Maint. probe	1	1	1	2	1	1
PIESH							
	BL to Final Tx day probe	1	1	2	2	2	1
	BL to 3-wk Maint. probe	NA	NA	NA	NA	NA	NA
WELED							
	BL to Final Tx day probe	1	2	2	1	2	1
	BL to 3-wk Maint. probe	1	1	2	1	1	1

Note. 1= Error-reducing (more help) condition is favored
2 = Errorful (less help) condition is favored
NA = not available (probes scheduled for 1/29/12).

Figures 1 and 2. Examples of Graphs Showing Probes Data for One Subject (AMBDE)

AMBDE wpm


A - More Help
Baseline
Treatment
Maintenance
wpm
70
60
50
40
30
20
0
20
40
60
80
100
B - Less Help
Baseline
Treatment
Maintenance
wpm
60
55
50
45
40
35
30
25
0
20
40
60
80
100
Day
Trained
Untrained
Generalization


AMBDE % Accuracy
A - More Help
Baseline
Treatment
Maintenance
% Accuracy
100
90
80
70
60
50
0
20
40
60
80
100
B - Less Help
Baseline
Treatment
Maintenance
% Accuracy
100
90
80
70
60
50
0
20
40
60
80
100
Day
Trained
Untrained
Generalization