CHAPTER 4

RESULTS AND DISCUSSION

4.1 Introduction

This chapter will present and discuss the results of the experimental procedures conducted. It focuses on answering the research question, the main aim and the sub-aims of the study as presented in Chapter 3 (section 3.2). The dependant variable under study was the number of symbols recognized following training using each treatment approach. Recognition levels were measured during the different probe measures. The independent variables were: the two treatment conditions (T1- self-generation and T2- non self-generation), the two symbol sets (symbol set 1 and symbol set 2) and the different time intervals or experimental sessions (day1 - E1, day 2 - E2, day 7 - E3). The discussion will proceed to present the interactions between these three factors or independent variables with each other but mainly the resultant effect of their interactions on the dependant variable.

4.2 Treatment Integrity and Reliability

Poor implementation or a failure to adhere to planned treatment protocols can pose a serious threat to the reliability and internal validity of a study (Schlosser, 2003). Hence an evaluation of treatment integrity is important as it would ascertain how well the described procedures and protocols have been adhered to. Treatment integrity helps to define the level of treatment reliability. Reliability refers to the consistency with which a method or procedure has been applied (Macmillian & Schumacher, 2001). Specifically, inter-rater reliability was used in this study. It refers to the match between ratings performed by different individuals who are required to evaluate the integrity of the experimental procedures conducted during data collection.

Two independent raters viewed three randomly selected video-recorded sessions (see section 3.9, chapter 3 for a full description of the procedures used). They rated the

video recorded sessions according to a checklist (see appendix 9, section 3.9, chapter 3). According to Schlosser (2003), it is important that at least 20%-40% of sessions be reviewed for treatment integrity and more importantly; there should be an equal review of sessions across all study phases. To meet this requirement, the two raters were asked to review sessions E1, E2 and E3 for three randomly selected participants. Table 4.1 below shows the spread of sessions which were rated.

Table 4.1 Sessions rated during inter-rater test	
Sessions	**Participant (P)**
E1	P1
E2	P4
E3	P7

Once the raters completed the checklist, their scores were tallied. Maxwell & Satake (2003) recommend a method based on probability theory for calculating inter-rater reliability. Using their method, the agreement levels were obtained by dividing the number of agreements by the total number of ratings (agreements and disagreements) per training area evaluated. The equation used is presented below. The scores were then converted to a percentage which is presented in table 4.2.

$$\frac{Agreements}{Agreements + Disagreements} \ X \ 100\% = Inter\text{-}rater \ agreement \ value$$

The agreement levels obtained fell between 80% - 100%. This is above the 70% agreement rate recommended by Macmillian & Schumacher (2001) for good inter-rater reliability. Hence inter-rater reliability was favourable.

Table 4.2 Inter-rater agreement levels			
Training Area Evaluated	**E1**	**E2**	**E3**
A. Self-Generation Protocol	86%	86%	100%
B. Non Self-Generation Protocol	83%	83%	100%
C. Probe Measures	100%	100%	100%

4.3 Scores Obtained

Table 4.3 presents a summary of the different scores that were obtained during the three experimental sessions (day 1-E1, day 2-E2 and day 7-E3). Experimental session 2 (day 2-E2) and experimental session 3 (day 7-E3) were derived from the two withdrawal periods imposed during testing. Therefore, the different experimental sessions over the different withdrawal periods represented the time variable in the analyses that were conducted.

As previously presented in Chapter 3 (section 3.7), there were two types of probe measures conducted. These were:

(i)   recognition probes (RP) conducted after training on each day (RPE1, RPE2, RPE3) which measured recognition levels for the specific symbol set administered (see table 3.1, chapter 3 for administration of sets and treatments)

(ii)   retention probes (RTP) conducted before re-training on day 2 – E2 and day 7 – E3 (RTPE2, RTPE3) which measured retention of recognition levels for both symbol sets.

Table 4.3 Summary of the probe measures		
Experimental Sessions TIME	**Procedure**	**Description**
Day 1 – E1	Training (treatments and sets allocated as per table 3.1, chapter 3)	This was the first session where the participant was trained to recognise symbols using a specific treatment and set (see table 3.1, chapter 3).
	Recognition probe (RPE1)	This was the first recognition probe conducted directly after the above training was completed. The recognition levels for the specific treatment approach and symbol set used was measured.
	Training (treatments and sets allocated as per table 3.1, chapter 3)	Training using the next treatment and symbol set was conducted.
	Recognition Probe (RPE1)	This recognition probe measured recognition levels for the specific treatment and symbol set used during the above training procedure(see table 3.1, chapter 3).

Experimental Sessions TIME	Procedure	Description
Day 2 – E2	Retention probe (RTPE2),	This retention probe was conducted to determine the level of recognition for the symbols retained following the withdrawal period of one day. This probe measured recognition outcomes for both symbol sets and treatments which were retained from day 1-E1.
	Training	Training began after retention probe (RTPE2) was completed. One specific treatment and symbol set was used.
	Recognition probe (RPE2)	This probe (RPE2) took place directly after the above training procedure was completed. It measured the recognition levels for the specific treatment and symbol set used during training.
	Training	Training on the next symbol set using the next treatment was conducted.
	Recognition probe (RPE2)	Recognition levels were measured for the specific treatment and symbol set used directly after training.
Day 7 – E3	Retention probe (RTPE3)	This retention probe was conducted to determine the level of recognition for the symbols retained following the withdrawal period of seven days. This probe measured recognition outcomes which were retained for both treatments from day 2-E2.
	Training	Training was conducted after the probe (RTPE3) using a specific treatment and symbol set.
	Recognition probe (RPE3)	This probe measured recognition levels for the treatment and symbol set used during the training procedure above.
	Training	Training using the next treatment and symbol set was conducted.
	Recognition probe (RPE3)	This probe measured the recognition levels for the symbol set and treatment used during the training procedure above.

4.4. Statistical Analysis Procedure

The 2X2X3 factorial design used for this study allowed for the analysis of the three independent variables and their interactional effects on the dependant variable using ANOVA. Hence the raw data was analysed in order to describe the effects of the two symbols sets, the two treatment approaches and the three time lines on the recognition levels for the Blissymbols taught (for details on data analysis procedures see section 3.8, chapter 3). The recognition probe scores (RPE1, RPE2 and RPE3) were analysed in the initial ANOVA analysis and the retention probe scores(RTPE2 and RTPE3) were analysed in the repeated ANOVA measure.

The initial ANOVA analysis considered the interactional effects of the three independent variables in the following manner:

(i)   Symbol sets: the two symbol sets interactions with the treatments and their resultant effect on the recognition of symbols were analysed. The symbol sets and their interactions with the time variable and its resultant effect on symbol recognition were also analysed. Finally, the symbol sets interactions with both the treatment and the time variable were analysed in order to isolate its influence on the recognition of symbols.

(ii)   Treatments: the two treatments interactions with the symbols sets were analysed. The two treatments interaction with the time variable was considered. Then the two treatments interactions with time and symbol sets were analysed.

(iii)   Time: the overall influence of the three time lines (day 1-E1, day 2-E2 and day 7-E3) on the recognition of symbols was analysed. Thereafter the interactions of the time variable in combination with the two treatments were analysed for its effect on the recognition of the symbols. The time variable's interactions with the combination of symbol sets and treatments were also looked at.

In order to isolate the significance of the retention probe scores, a repeated ANOVA measure was performed (rANOVA). The following retention probe scores were

compared for their interactions with the two treatments and their resultant effect on the recognition of the symbols: (i) comparison of scores for RPE1 and RTPE2 and (ii) comparison of scores for RPE2 and RTP3.

4.5 Presentation of Results

Table 4.4 presents the results of the initial ANOVA analysis which analysed the significance of the recognition probe scores (RPE1, RPE2 and RPE3). The variables included were: the *time* variable which represented the different experimental sessions or day1-E1, day 2-E2 and day 7-E3, the *treatment* variable which included treatment 1 (T1 Self-Generation SGE) and treatment 2 (T2 – Non Self-Generation NSG), and the *set* variable which was symbol set 1(S1) and symbol set 2 (S2). The results presented in table 4.4 and its significance will be discussed in the sections to follow.

Table 4.4 ANOVA analysis results Interactions of time, treatments and sets on recognition levels			
Independent Variable Interactions	**p-value***	**Effect Size***	**Effect Size Rating**
Time	0.387	0.051	small
Treatment	0.322	0.027	small
Set	0.709	0.004	small
Time-Treatment	0.556	0.032	small
Time-Set	0.228	0.079	small
Treatment-Set	0.851	0.001	small
Time-Treatment-Set	0.264	0.071	small

* p-value: if $p < 0.05$, then significant at a 5% level
** Effect size: 0.01 – 0.05 small, 0.06 – 0.14 medium, >0.15 large.

4.5.1 Interactions of the symbol sets with recognition levels

An intensive development procedure was followed (as described in section 3.5.1, chapter 3) to provide an argument for the equivalency of the two symbol sets. The present ANOVA analysis ascertained whether the symbol sets produced any bias in how the participants learnt to recognise the symbols taught. The initial ANOVA analysis helped to isolate any significant interaction between the sets, treatments and time on the dependant variable (recognition levels for the symbols as determined by

RPE1, RPE2 and RPE3). Table 4.4 presents the ANOVA results together with the effect sizes. As is evident from table 4.4, a non-significant p-value together with a small effect size, was obtained for the interaction of the symbol sets with treatments, time, and the combination of treatment-time. This indicated that there was no significant interaction between the symbol sets used during training and the level of symbol recognition. Hence the symbol recognition levels were not influenced by any ease of recognisability of any one symbol set over the other.

Table 4.5 shows the mean (M) symbol recognition levels and their standard deviations for symbol set 1 and symbol set 2 across experimental sessions. These similar mean recognition levels for symbol set 1 and symbol set 2 further validated the equivalency procedure conducted during the material development phase of this study.

Table 4.5 Mean recognition levels and standard deviations across symbol sets (n 14)		
	Mean (M)	**Standard Deviation**
Set 1	9.04	2.49
Set 2	9.12	2.83

4.5.2 Interaction of treatments with symbol recognition

4.5.2.1 Treatment 1 (self-generation) and treatment 2 (non self-generation)

Table 4.4 shows the non-significant interactions between treatments and the level of symbol recognition determined during the recognition probes. Figure 4.1 shows the mean recognition levels obtained during RPE1, RPE2 and RPE3 when using treatment 1 (self-generation - SG) and treatment 2 (non self-generation – NSG). As is evident from the similar mean recognition levels across recognition probe measures, the self-generation approach was showing no recognition advantage when the recognition probe scores were analysed.
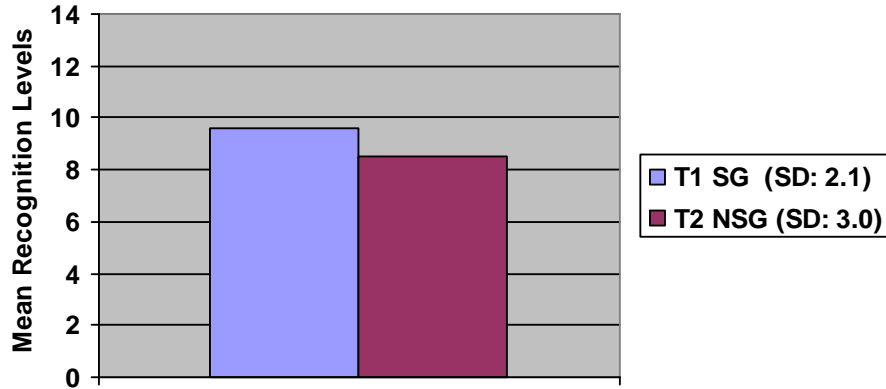
Figure 4.1 Mean recognition levels and standard deviations (SD) for treatment 1 (Self-Generation SG) and treatment 2 (Non Self-Generation NSG)

Figure 4.1 presents the mean recognition level and standard deviations (SD) of both treatment conditions. It provides as overview of the performance across all experimental sessions.

The results do indicate that the participants with severe Global aphasia and severe Broca's aphasia do have the ability to recognise Blissymbols. This is consistent with previous studies which have shown that AAC may be a viable method for establishing communication in these patients (Koul, Corwin & Hayes, 2004; Koul & Harding, 1998; Koul & Lloyd, 1998; McCall, Shelton, Weinrich & Cox, 2000).

However, the results from the recognition probes shown in figure 4.1, do not support the superiority of a self-generation type of approach when trying to teach individuals with severe aphasia to recognise symbols. There seemed to be no obvious, immediate benefit to the patient's increased involvement in the construction of the to-be-recognised symbols. The underlying premise of the SGE is the notion that the more an individual is involved in the construction of the to-be-remembered item, the greater the memorial advantage for that item can be expected. The SGE research has successfully been able to replicate the SGE phenomenon using words, numbers, sentences and pictures with non brain-damaged individuals (Graf, 1980; Gardiner & Hampton, 1985; Peynircioglu, 1989; Slamecka & Graf, 1978,). Studies conducted

101

with individuals with brain-damage have also been able to show a memory advantage for words when using the SGE (Barrett, Crucian, Schwartz, & Heilman, 2000; Chiaravalloti & DeLuca, 2002; Chiaravalloti, DeLuca, Moore & Ricker, 2005; Dick & Kean, 1989; Goverover, Chiaravalloti, Johnston & DeLuca, 2005; Lengenfelder, Chiaravalloti & DeLuca, 2003; Lipinska, Backman, Mantyla & Viitanen, 1994; Michell, Hunt & Schmitt, 1986; Multhaup & Balota 1997; O'Brien, Chiaravalloti, Arango-Lasprilla, Lengenfelder & DeLuca, 2007; Souliez, Pasquier, Lebert, Leconte & Petit, 1996). However the SGE failed to emerge during the recognition probe measures in this study. In comparison to other studies on the SGE, the present study's format differed by the specific deficits associated with the severe aphasia and the type of stimuli used to elicit the SGE. Hence, it is plausible that the reason for the SGE not emerging during the recognition probe tests was influenced by the very nature of the severe aphasia itself and the type of stimuli used to elicit the SGE.

There are many characteristics of a severe aphasia (as presented in section 2.6.1, chapter 2) that could contribute to how these individuals learn to recognise symbols. Deficits in higher level cognitive processes seem to pose the foremost threat to symbol recognition. However, O'Brien et al's (2007) study into the SGE with participants presenting with traumatic brain injury did show that even individuals with multiple deficits in most cognitive domains (i.e. working memory, episodic memory and executive functioning) were still able to benefit from the SGE. Hence ruling out deficits within the cognitive domains that are associated with a severe aphasia, some other possibilities for why the SGE did not enhance the symbol recognition levels needs to be explored.

The most relevant SGE theory points to the process of semantic activation as a contributor to the emergence of the SGE (Graf, 1980; McElroy & Slamecka, 1982; Nairne et al, 1985; Payne et al, 1986). This theory assumes that because self-generation is more effortful it activates the items location in the individual's lexical network and enhances the items retrieval from memory. However, in the instance of this present study, the opposite explanation is proposed. It is possible that the SGE failed to emerge because it was the NSG approach and not the SGE approach which allowed the participant to determine the semantic or meaning association between the symbol and its referent thus influencing its recognition. During the NSG procedure

the participant was not required to perform any actions besides studying the symbol carefully and remembering it for the test to follow. Hence, it is possible that the participant merely used the allotted time to examine the symbol and make sense of it. On the other hand, the SGE approach required the participant to draw the dot picture to make the symbol. It is possible that the participant then focused on completing the drawing successfully and failed to focus on determining the semantic link between the symbol and its referent. The drawing activity was distracting and shifted the participants attention away from making the semantic association required. Hence, lexical activation did not occur and the SGE failed to emerge on recognition testing.

Another possibility for SGE not emerging requires evaluation of the stimuli used. Some SGE studies have shown no SGE for non-words or non-meaningful stimuli (Gardiner & Hampton, 1985; Nairne, Pusen and Widner, 1985; Nairne & Widner, 1987). The semantic-association or lexical activation (McElroy & Slamecka, 1982) theories on SGE suggest that in order for the SGE to emerge, meaning must be established. The Blissymbols are in-fact new, non-meaningful picture stimuli presented to the participants for the first time. In fact one of the participant selection criteria was no previous exposure to AAC training (see section 3.3.2, chapter 3). The participants could have failed to attach meaning to the Blissymbols on a single exposure. The drawing of the dot picture could have also further distracted their attention away from forming a meaningful link between the symbol and its referent. Hence, poor SGE recognition scores were obtained for the Blissymbols as the participants did not have a preexisting mental representation of the symbols.

A recent study by Lutz, Briggs, & Cain (2003, p. 171) also showed a "greatly reduced generation effect for new, unfamiliar material". This study contrasted unfamiliar sentences from textbooks with familiar clichés. They concluded that the SGE can have limited effectiveness on memory for new, unfamiliar stimuli. Since the participants in this study were exposed to the Blissymbols for the first time, the Blissymbols could also be classified as new and unfamiliar stimuli hence producing similar results to the above study.

Dick & Kean (1989) and Mitchell, Hunt & Scmitt (1986) also did not find a SGE when it was tested with a group of subjects presenting with dementia. They also

argued that given the role of semantic activation in the SGE and since semantic memory is disrupted in dementia of the Alzheimer's type, a SGE could not be expected. Souliez, Pasquier, Lebert, Leconte & Petit (1996) also supported this contention. They agreed that the participant's lack of access to semantic memory or semantic activation contributed to the SGE not being found.

4.5.3 Interactions of time with symbol recognition

4.5.3.1 Results of the recognition probes

The time variable in the initial ANOVA conducted analysed the scores of the recognition probe measures conducted after training on day 1 – E1 (RPE1), day 2 – E2 (RPE2) and day 7 – E3 (RPE3). Table 4.4 shows that there was no significant interactions between time, symbols sets or treatments. Figure 4.2 shows the mean number of symbol recognised over the three recognition probes (RPE1, RPE2, RPE3).
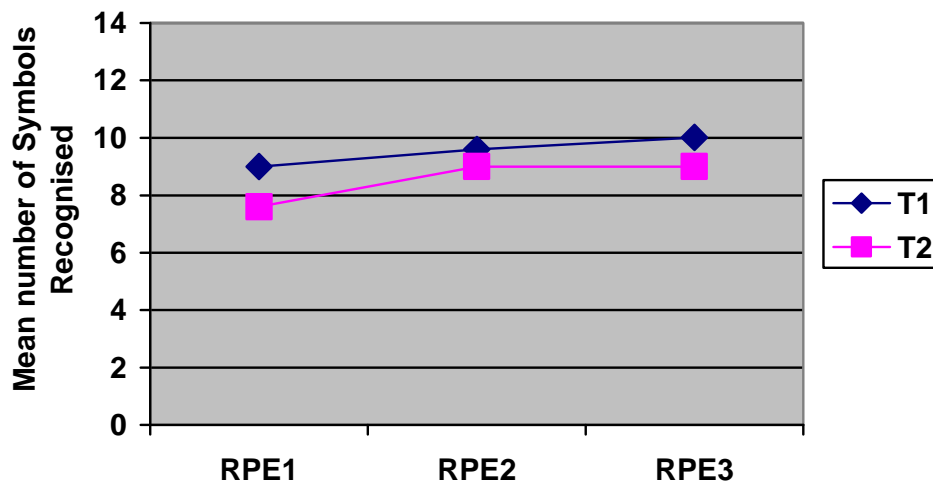

Figure 4.2 Results of the recognition probes

4.5.3.2 Results of the retention probes

A repeated analysis of variance (rANOVA) was performed in order to determine the significance of the retention probe scores. There were two retention probes conducted in order to ascertain how symbol recognition levels were being retained as a function of the two withdrawal periods and the two treatments. Two types of comparisons were made. In order to analyse the retention of recognition levels after the first withdrawal period (one day), recognition scores from recognition probe RPE1 and retention probe RTPE2 were analysed. In order to analyse the effect of the second withdrawal period (seven days) scores from recognition probe RPE2 and retention probe RTPE3 were compared.

Table 4.6 presents the rANOVA results of the comparison between recognition levels for RPE1 and RTPE2. As is evident from table 4.6, the p-values obtained were not significant and the effect sizes were small. Hence no statistically significant difference existed between these experimental sessions.   The same number of symbols recognised in RPE1 when using either treatment 1 (self-generation) or treatment 2 (non self-generation) were retained and recognised to the same level during the retention probe (RTPE2).

Although these results do not support the superiority of the SGE, the results could have implications for patients who present with a severe aphasia. It is evident that persons with severe aphasia can learn to recognise Blissymbols when using either a self-generation type of method or a more traditional teaching method. These individuals also appear to retain the ability to recognise these Blissymbols after a one day withdrawal period suggesting the integrity of each treatment approach as a methodology for teaching these Blissymbols to persons presenting with a severe aphasia. It was seen that the level of Blissymbol recognition gained on initial training was carried over to the next training session.

Table 4.6 Comparisons between RPE1 and RTPE2		
Comparison Variable	**P-Value***	**Effect Size****
Treatment	0.840	0.194
Set	0.548	0.162
Treatment and Set	0.112	0.001

* if $p < 0.05$, then significant at a 5% level
** Effect Size: 0.01 – 0.05 small, 0.06 – 0.14 medium, >0.15 large.

The rANOVA results which compared scores obtained during recognition probe RPE2 and retention probe RTPE3 is presented in table 4.7. This helped to ascertain the retention of recognition levels following the second withdrawal period of seven days.

Table 4.7 Comparisons between RPE2 and RTPE3			
Comparison Variable	**P-Value***	**Effect Size****	**Effect size Rating**
Treatment	0.198	0.150	medium
Set	0.100	0.005	small
Treatment and Set	0.002*	0.190**	large

* if $p < 0.05$, then significant at a 5% level
** Effect Size: 0.01 – 0.05 small, 0.06 – 0.14 medium, >0.15 large.

The analysis indicated there was a significant difference between RPE2 and RTPE3 as a function of treatments and sets. This is reflected by the significant p-value (0.002) and the large effect size (0.190). In order to further analyse this significant interaction, a further rANOVA was performed. Table 4.8 below shows the results of this rANOVA which presents the mean difference between recognition levels obtained during RPE2 and RTPE3 as a function of treatments and sets. The mean difference score describes the difference in recognition levels across the two probe measures. The negative mean difference scores indicates that there were less symbols retained and recognised by day 7-E3 when compared to day 2-E2.

Table 4.8 Mean difference and standard deviations (SD) between recognition levels for RPE2 and RTPE3

Treatment	Set	Mean Difference	SD
1	1	0.7	0.9
	2	-1.5	0.7
2	1	-4.0*	1.4
	2	0.5	1.8

*largest difference in the recognition levels between time lines

This analysis indicates that the largest difference in recognition levels between RPE2 and RTPE3 was for treatment 2 (non self-generation) on set 1. This is presented in figure 4.4. This indicates there were less symbols recognised in session E3 compared to E2 when using the treatment 2 (NSG). However, for treatment 1 (SG) there was no large, significant difference in the number of symbols recognised between sessions E2 and E3. This implies that when using the SG treatment, the same number of symbols recognised in session E2 were retained and recognised to the similar level during the retention probe (RTPE3) conducted at the beginning of session E3.
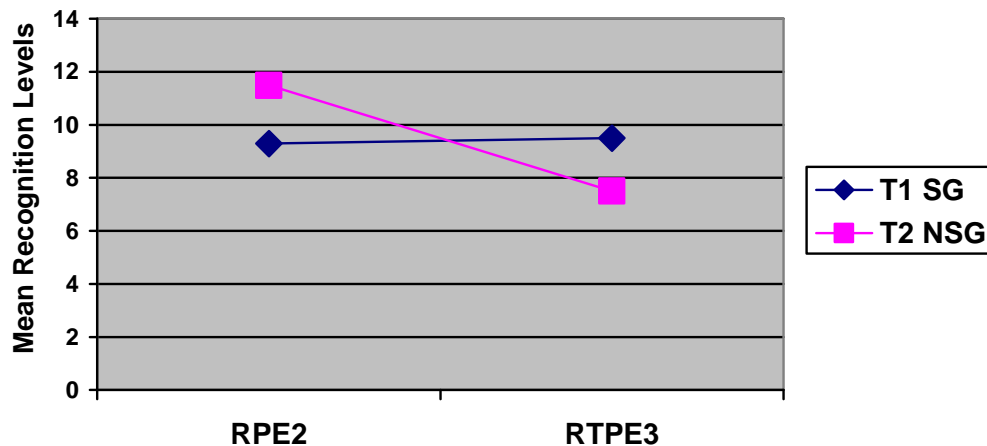

Figure 4.3 Comparison between mean recognition levels for SET 1 during RPE2 and RTPE3

These results point to the possible robustness of the memory enhancing effect caused by the self-generation treatment on the recognition of the symbols. It is possible that the participant's involvement in constructing these symbols in session E2 could have

107

led to the sustained retention of recognition levels following the withdrawal period in session E3. Interestingly, this increased long-term retention benefit for generated stimuli was also reported by Kornell & Terrace (2007) when they studied the SGE in monkeys. These researchers contrasted the SG and NSG conditions using photographs as stimuli. The subjects were presented with a touch screen which included a set of five photographs. They were required to touch the photographs according to a specific sequence. In the NSG condition the subjects performed the sequence with hints provided for the correct photograph sequence required. In the SG condition, the subjects performed the sequence without hints using trial and error. The results showed that the subjects' performance levels were better using the NSG condition during the first three days of training. However, as training continued over time, the SG condition started to show higher performance levels. The researchers concluded that "although the active generation of answers during training may result in low initial performance, it enhances long-term retention and transfer" (Kornell & Terrace, 2007, p. 685). Similar results were seen in the present study because during initial sessions E1 and E2, the SG condition did not show any advantages but as training proceeded the SGE showed some signs of improving retention of recognition levels after the longer withdrawal period.

The difference in the retention of recognition levels between E2 and E3 also provides further support for the semantic-association or lexical activation theories for the SGE. It is possible that during the repeated exposures to the Blissymbols during the two preceding training sessions (E1 and E2), the participants were starting to develop a mental representation for the Blissymbols and were beginning to attach meaning to the symbols. Hence, a SGE was starting to emerge.

Additionally the Blissymbols after repeated exposures during sessions E1 and E2, were now becoming more familiar stimuli. Since the SGE has been shown to be ineffective for non-familiar stimuli, it is probable that the increased exposure during training was also increasing the familiarity levels hence the SGE emerged.

4.6 Summary

This chapter provided a presentation of the research results and discussion. It focused on answering the central research question which determined if the self-generation effect could enhance the recognition of Blissymbols in severe aphasics when it is used as a teaching strategy. The results indicated that there was no statistically relevant difference between the participants' performance on the self-generation or non self-generation approach during recognition probe testing. However, there was a statistically relevant recognition advantage on the SG treatment as seen by the better retention of recognition levels from day 2-E2 to day 7-E3. Hence, the SG treatment showed better retention of symbol recognition over time. Previous studies have shown that the self-generation effect failed to emerge with stimuli that were new or unfamiliar. This trend was also seen in this study. The results provide support for a semantic-association theory for the SGE.

CHAPTER 5

CONCLUSIONS, EVALUATIONS AND RECOMMENDATIONS

5.1 Introduction

This chapter will provide the final conclusions of the study. It provides an evaluation of the strengths and weaknesses of the study and presents some recommendations for future research.

5.2 Conclusions

5.2.1 The symbol sets

The study aimed to construct two equivalent symbol sets that could be counter-balanced between the two treatment approaches. The ANOVA analysis performed showed that there was no statistically significant interaction between the sets and the treatments. Hence the equivalency procedure was successful.

5.2.2 Self-generation (SG) versus non self-generation (NSG)

The results from the recognition probe measures conducted indicated that there was no overall difference between the recognition levels for the Blissymbols using either SG or NSG. This supports other studies which have shown that persons with severe aphasia can be trained to recognise symbols. This is consistent with the SGE research findings which show no SGE for new, non-meaningful and unfamiliar material. The finding of similar recognition levels across treatments during the recognition probes, also provided support for a semantic-association or lexical activation theory of the SGE. Since the Blissymbols were new, non-meaningful stimuli, no semantic association was made initially. The severity of the participants' aphasia may have also hampered the process of establishing a meaning association between symbol and referent pairs.

However, the results did show that there was a statistically significant difference between recognition levels using SG and NSG following the seven day withdrawal period as determined by the retention probe measure. The SG treatment appeared to produce better retention of recognition levels over time. With repeated exposure to the Blissymbols, the stimuli became more familiar to the participants and it is possible that the semantic associations between the symbol and its referent became more apparent. Hence over time, the SG effect began to emerge.

5.3 Evaluation of the Study

The strengths of the study included the following:

- The construction of two equivalent symbol sets was a strength of this study. The statistical analysis showed no bias imposed by the sets. Hence, the sets were shown to be functionally equivalent.

- The factorial design was a strength of the study because it allowed for the analysis of the interactions of all three independent variables and its resultant influence on the dependant variable.

- The use of raters to evaluate the treatment procedures supports the treatment integrity of this study. The positive inter-rater scores also support the equivalency of the training procedures used across participants.

- The inclusion of both recognition and retention probe measures was another strength on this study. If the retention probes were omitted, the emergence of the SGE as a function of time would not have been demonstrated.

- The study included an intensive pre-experimental screening phase which decided if a prospective participant was a suitable candidate for inclusion into the study. More importantly, the Receptive/ Pointing Test conducted during screening confirmed that the participants were able to understand and recognise the target referents. If not, the recognition levels may have been largely influenced by

auditory comprehension levels. The Boston Diagnostic Aphasia Examination alone could not have provided any reliable indication of how the participants understood the symbol referents. Hence this study included both test types to ascertain receptive language skills.

- The withdrawal periods helped to isolate the emergence of the self-generation effect. There was also a strong match between the training task and the test (probe measure) tasks. The testing phase called for recognition of symbols which was in-keeping with the level of training conducted. A more difficult test task like free recall would have not been appropriate for the amount of training that occurred.

- The actual format of the probe measures was a strength as it tested recognition levels for the symbols taught in the most basic, straight forward and uncomplicated manner as possible. This was important if the SGE was to emerge as demonstrated by Gardiner (1988), Slamecka & Graf (1978) and Graf (1980).

- The randomisation of symbol presentation on the grids used during the different probe measures was also an important strength. Should the presentation have been kept the same, an exposure bias could have led to better recognition performance as a function of time.

- The customisation of the connect-the-dot pictures by correlating the number of dots used per illustration to the complexity values proposed by Fuller & Lloyd (1987) supported the standardisation of these illustrations. This was a positive aspect of this study's methodology.

The limitations of the study include the following:

- Both Broca's aphasics and Global aphasics were included in the study. Although there was an equal split between these two types of severe aphasias and they were matched across four variables (time of onset, severity, lesion site and education); homogeneity between these groups is limited. However, the inclusion of both aphasia types may also have been a strong point of this study because it may have

been possible that a sub-group of aphasic individuals who may learn better using SG could have been identified. The inclusion of both these groups was necessitated by the lack of availability of a sufficient number of participants from either one of the groups.

- The study only investigated one aspect of memory which was recognition. A free recall task could have caused the SGE to emerge earlier as it is a more difficult test of memory and the benefits of generation could have been isolated. For example, in a free recall test format the participant could be presented with symbols cards, which included all the test symbols together with foils, and the participant would be asked to select all the symbols he/she had seen during the training. During such a free recall task, which also reduces the auditory comprehension demands of asking an individual with aphasia to point to the symbol named; the SGE may have had a better chance of emerging.

- The study only investigated the SGE over three training days and a seven day withdrawal period. A longer withdrawal period might have yielded additional insights in understanding the influence of the SGE.

5.4 Recommendations for Future Research

The recommendations for future research are as follows:

- Broca's and Global aphasics were included in this study. The inclusion of participants with different types of aphasia may be important in future studies as this may help to identify sub-groups of aphasic individuals who benefit more from the application of the SGE.

- Since the issue of unfamiliarity of stimuli was discussed and evaluated, it would be interesting to see how severe aphasics who have already been exposed to Blissymbols may perform when using self-generation.

- It would be interesting to see how other types of self-generation strategies apart from the dot drawing used here, could impact on the emergence of the self-generation effect. It was mentioned that the dot drawing could have been very distracting.

- The SGE could be tested with other types of stimuli such as words, numbers and pictures.

- A study which focuses on longer training schedules and longer withdrawal periods might add important insights in relation to when the SGE may emerge during training. Increasing the number of experimental sessions and changing the length of the withdrawal periods could also help to identify the robustness of the SGE over time.

- It would also be interesting to test whether the SGE would emerge when used with other types of populations who present with cognitive deficits. This would help to further delineate the involvement of semantic activation.

- Replication of this study with normal adult learners may help to isolate factors that contribute to the SGE and its effect on the learning and memory.

- The two equivalent symbol sets could be used for further study. Research into the use of these symbol sets with other populations could yield interesting results.

5.5 Summary

The major conclusions and recommendations of the present study are presented in this chapter. The study's strengths and limitations are discussed. There remain certain issues that do merit further investigation via future research.