

LEXICAL COMPETITION EFFECTS IN APHASIA

by

Christina May Louise Cameron-Jones

A thesis

submitted to the Victoria University of Wellington

in fulfilment of the

requirements for the degree of

Doctor of Philosophy

in Psychology

Victoria University of Wellington

2008

## Abstract

Some aphasic patients show single word production deficits in some situations where object naming is required (e.g., they perform well when objects are presented in unrelated groups (e.g., *Cat, Fork, Bread...*), but deteriorate when the same items are presented in semantically related groups (e.g., *Cat, Cow, Dog...*)) (see Wilshire & McCarthy, 2002). We investigated whether context-sensitive single-word production impairments reflect an impaired ability to resolve lexical competition. Three groups of participants (non-fluent aphasics, fluent aphasics, and older controls) completed four tasks that manipulated lexical competition: 1) A category exemplar task, where a high competition condition involved generating items from broad categories (e.g., *Animals*: “Cat. Dog” etc.), and a low competition condition involved generating items from narrow categories (e.g., *Pets*: “Cat. Dog” etc); 2) A verb generation task, where participants were presented with objects and were required to generate related verbs. The high competition objects were related to a range of verbs (e.g., *Penny*: “Spend”/“Pay”/“Buy” etc), and the low competition objects were related to one dominant verb (e.g., *Scissors*: “Cut”); 3) A name agreement task, where a high competition condition involved naming low name agreement objects (e.g., *Artist/Painter*), and a low competition condition involved naming of high name agreement objects (e.g., *Anchor*), and; 4) A sentence completion task, where extrinsic competition was introduced via presentation of auditory distracters. The low competition distracters did not make sense (e.g., *Barry wisely chose to pay the RANGE*: “Bill”/“Cashier” etc), whereas the high competition distracters did (e.g., *Barry wisely chose to pay the FINE*: “Bill”/“Cashier” etc). Our first hypothesis was that all participants would show high competition costs in increased response latencies and/or decreased accuracy. At the group level, this hypothesis was supported in all four tasks. At the individual level, there was mixed support as some participants showed predicted effects on the verb generation, name agreement, and sentence completion tasks. The second hypothesis was that exaggerated competition costs would occur in some or all non-fluent aphasics. At the group level this hypothesis was not clearly supported on any task. At the individual level there was mixed support, with some indications that non-fluents may be more likely to show significant competition effects than fluents. The third hypothesis was that non-fluent aphasics with relatively well preserved single word production but relatively impaired sentence production may be most likely to show exaggerated lexical competition effects. There was little support for this hypothesis. It was concluded that the

data do not support the hypothesis that context-sensitive single-word production impairments are symptomatic of an impaired ability to resolve lexical competition. However, we have gained information on how heterogeneous aphasics perform on tasks that manipulate lexical competition, and we have gained some insights that may direct future research down a path towards more informative results, and increased knowledge on the complex process of speech production.

## Acknowledgements

Many people have assisted me during the completion of this thesis. Most notably, my supervisor Dr. Carolyn Wilshire has been a constant source of guidance and support, and deserves a very special thank you for her contribution. Special thanks must also go to the staff and students at the Eleanor M. Saffran Center for Cognitive Neuroscience at Temple University who welcomed me into their lab, gave me free access to their participant pool, and discussed this work on many occasions. Particular thanks go to Dr. Nadine Martin, Francine Kohen, Michelene Kalinyak-Fliszar, Kelly Bowes, and Drew Tipton. I would also like to thank my fellow students and research assistants at the Victoria University Neuropsychology Laboratory who assisted with the background testing of the participants from New Zealand. Thank you Lauren Cloutman, Carolina Lukkien, Kate Smith, Rebecca Scott, Marion Roberts, and Corinne Bareham.

Thanks also to all of the people who took part in these studies, particularly the aphasic participants who have volunteered many hours of their time during the course of this project.

I also extend my gratitude to the Victoria University School of Psychology, the Victoria University Faculty of Science, and the Royal Society of New Zealand Marsden Fund, for the financial support that they have provided.

Finally, I would like to thank my husband Robin, who has been a great support financially and emotionally, and who has provided endless encouragement during the completion of this work.



## Contents

Abstract	ii
Acknowledgements	iv
List of Figures and Tables	vii
Introduction	1
Context Effects in Non-Fluent Aphasia	3
Language Theories	11
Language Models and Aphasia	23
Language Models and Context Effects	25
Insights from Studies of Frontal Lobe Function	33
Introduction to the Current Research	38
Experiment 1: Category Exemplar Task	43
Experiment 2: Verb Generation Task	79
Experiment 3: Name Agreement Task	115
Experiment 4: Sentence Completion Task	134
Determinants of Performance Across the Four Tasks: Unpacking the Contributions of Specific Components of Fluency	159
General Discussion	180
Appendices:	
Appendix A1: Category Exemplar Pilot Task Instructions	193
Appendix A2: Descriptions of the Tests Used for Diagnostic Purposes	196
Appendix A3: Case Descriptions of Aphasic Participants	200
Appendix A4: Complete Control Participant Results from the Category Exemplar Task	221
Appendix B1: Verb Generation Pilot Task Instructions	223
Appendix B2: List of Items Included in the Verb Generation Task by Competition Condition	224
Appendix B3: List of Items Included in Verb Generation Task by Frequency Condition	225
Appendix B4: Additional Results for the Verb Generation Task Frequency Condition Analysis	226
Appendix B5: List of Items included in the Verb Generation Task by	227

Association Strength Condition	
Appendix C1: List of Items Included in the Name Agreement Pilot, with Response Strength Ratios	228
Appendix C2: Items Included in the Name Agreement Task by Competition Condition	232
Appendix D1: Sentence Stems and Distracters Included in the Sentence Completion Task by Competition Condition, and Constraint Condition	234
Appendix D2: Additional Results for the Sentence Completion Task	236
References	237

## List of Figures and Tables

*Figures*

1.1	Fragment of the Dell, Schwartz, Martin, Saffran, and Gagnon (1997) two-stage model of lexical retrieval.	12
2.1	Category exemplar task: Mean number of correct exemplars produced by participant and competition condition.	74
3.1	Verb generation pilot: Example trial.	84
3.2	Verb generation task: Number of correct responses by participant, frequency condition, and competition condition.	101
3.3	Verb generation task: Geometric mean response latencies by participant, frequency condition, and competition condition.	104
3.4	Verb generation task: Non-fluent JHM's total accuracy scores across session by association strength condition.	107
3.5	Verb generation task: Geometric mean response latencies across session by participant and association strength condition.	108
4.1	Name agreement task: Mean response latencies by participant and competition condition, from the accuracy-sensitive latency analysis.	130
5.1	Sentence completion task: Number of correct responses made in less than thirty seconds by participant and competition condition.	141
5.2	Sentence completion task: Number of distracter errors and other errors across session by participant and competition condition.	144
5.3	Sentence completion task: Number of correct responses made in less than twenty seconds by participant and competition condition.	149
5.4	Sentence completion task: Number of correct responses made in less than ten seconds by participant and competition condition.	150
5.5	Sentence completion task: Accuracy rates by participants and competition/constraint condition.	154
A1	Category exemplar task: Complete control data from session one showing the mean number of correct exemplars produced by category and competition condition.	222

*Tables*

1.1	Summary of the Diagnostic Categories Used to Describe Aphasic Speech in the BDAE	2
1.2	Examples of Speech Errors Found in Speech Error Corpora	15
2.1	Category Exemplar Pilot: Semantic Categories and Sub-Categories Included	46
2.2	Category Exemplar Pilot: Minimum, Maximum, Mean, and Total Number of Items Generated by Category and Competition Condition	47
2.3	Participant Date of Birth and Gender Information	51
2.4	Non-Fluent Aphasic Participants' Scores/Classifications on Diagnostic Tests	53
2.5	Fluent Aphasic Participants' Scores/Classifications on Diagnostic Tests	56
2.6	Mean Scores and Standard Deviations on Diagnostic Tests by Aphasic Group with t-test Statistics	60
2.7	Spontaneous Speech Samples taken from the Cookie Theft Scene Description in the BDAE	62
2.8	Category Exemplar Task: Scoring Procedure	71
3.1	Verb Generation Pilot: List of Objects Included with Response Strength Ratios	84
3.2	Verb Generation Task: Number of Correct Responses by Participant and Competition Condition, with Repeated Measures Logistic Regression Statistics	95
3.3	Verb Generation Task: Geometric Mean Response Latencies by Participant and Competition Condition, Collapsed Across Session, with General Linear Model Statistics	97
3.4	Verb Generation Task: Minimum, Maximum, and Mean Response Strength Ratios and Association Strength Ratios, by Association Strength Condition	106

4.1	Name Agreement Pilot: Minimum, Maximum, and Mean Frequency Values and Word Lengths of Objects Selected for the Name Agreement Task by Competition Condition	123
4.2	Name Agreement Task: Number of Correct Responses by Participant and Competition Condition, with Binary Logistic Regression Statistics	126
4.3	Name Agreement Task: Geometric Mean Response Latencies by Participant and Competition Condition for Correct Responses, with General Linear Model Statistics	128
5.1	Sentence Completion Task: Geometric Mean Response Latencies by Participant and Competition Condition for Correct Response Pairs, with Paired t-test Statistics	146
6.1	Summary of Individual Results on all Experimental Tasks, Showing where Significant Effects were observed on Simple Accuracy and Latency Analyses	160
6.2	Additional Diagnostic Measures by Participant	163
6.3	Summary of Individual Results on Selected Tasks, Showing where Significant Effects were observed on Simple Accuracy and Latency Analyses, Ordered by QPA Words per Minute over BNT Percentage Correct Score	166
6.4	Summary of Individual Results on Selected Tasks, Showing Percentage Change Values between Competition Conditions on Simple Accuracy and Latency Analyses Ordered by QPA Words per Minute over BNT Percentage Correct Score, with Correlation Statistics	168
6.5	Summary of Individual Results on Selected Tasks, Showing where Significant Effects were observed on Simple Accuracy and Latency Analyses, Ordered by Open vs. Closed Class Words and Noun vs. Verb Composite Score	169

6.6	Summary of Individual Results on Selected Tasks, Showing Percentage Change Values between Competition Conditions on Simple Accuracy and Latency Analyses, Ordered by Open vs. Closed Class Words and Noun vs. Verb Composite Score, with Correlation Statistics	170
6.7	Summary of Individual Results on Selected Tasks, Showing where Significant Effects were observed on Simple Accuracy and Latency Analyses, Ordered by Nouns and Isolation vs. Nouns in Context Score	171
6.8	Summary of Individual Results on Selected Tasks, Showing Percentage Change Values between Competition Conditions on Simple Accuracy and Latency Analyses, Ordered by Nouns and Isolation vs. Nouns in Context Score, with Correlation Statistics	173
6.9	Summary of Individual Results on Selected Tasks, Showing where Significant Effects were observed on Simple Accuracy and Latency Analyses, Ordered by COWAT Letter Fluency Score over BNT Percentage Correct	174
6.10	Summary of Individual Results on Selected Tasks, Showing Percentage Change Values between Competition Conditions on Simple Accuracy and Latency Analyses, Ordered by COWAT Letter Fluency Score over BNT Percentage Correct, with Correlation Statistics	175
6.11	Summary of Individual Results on Selected Tasks, Showing where Significant Effects were observed on Simple Accuracy and Latency Analyses, Ordered by COWAT Category Fluency (Animals Only) over BNT Percentage Correct Score	176
6.12	Summary of Individual Results on Selected Tasks, Showing Percentage Change Values between Competition Conditions on Simple Accuracy and Latency Analyses, Ordered by COWAT Category Fluency (Animals Only) over BNT Percentage Correct Score, with Correlation Statistics	177

## General Introduction

After stroke or head injury, some patients present with non-fluent aphasia, producing speech that is made up of very short phrases that utilise a limited variety of grammatical constructions (see Goodglass, Kaplan & Barresi, 2001). Some of these aphasics show a remarkable, and seemingly disproportionate, preservation of the ability to produce words in isolation (e.g., non-fluent BM, see Wilshire and McCarthy (2002)). Because non-fluent aphasics have particular difficulty producing words within sentences, some researchers have suggested that their sentence level linguistic processing, concerning things like syntax and grammar, may be deficient (see Goodglass, Christiansen & Gallagher, 1994; Schwartz, Linebarger & Saffran, 1985). However, problematically for sentence level explanations of non-fluent aphasia, recent research has presented case studies on non-fluent aphasics who generally showed good single object naming, but showed deficiencies when single object naming occurred within specific contexts. Importantly, these contextual manipulations did not increase demands on sentence level processes – being as simple as presenting objects in semantically related groups (e.g., *Cat, Cow, Dog...*) rather than unrelated groups (e.g., *Cat, Fork, Bread...*) (see Jefferies, Baker, Doran, & Lambon Ralph, 2007; Schnur, Schwartz, Brecher, & Hodgson, 2006; Schwartz & Hodgson, 2002; Wilshire & McCarthy, 2002).

The finding that contextual manipulations can significantly impact single word production suggests that when an impairment of connected speech production occurs with disproportionately good single word production, the impairment may not be due to disrupted sentence level processes alone. Some researchers have suggested that another factor that may contribute to these impairments is an underlying difficulty selecting words when more than one word is activated simultaneously (Wilshire & McCarthy, 2002; Wilshire, Scott, & Stuart, 2006). Following studies by Wilshire and McCarthy, and Wilshire et al., the current work investigates the possibility that context-sensitive single word naming impairments may be symptomatic of an impaired ability to resolve lexical competition.

Before the current research is introduced in detail, a literature review is presented that includes summaries of empirical observations related to context-specific language deficits. The summarised research was completed with normals and with participants with a range of different acquired speech deficits (aphasias).

According to the Boston Diagnostic Aphasia Battery (BDAE), the diagnostic categories used to describe the different aphasias fall into two main groups: the non-fluent aphasias which are marked by impaired sentence production; and the fluent aphasias which are marked the production of phrases of normal length that utilise a variety of grammatical constructions (see Goodglass et al., 2001). The main diagnostic categories used to describe aphasic speech are detailed in Table 1.1.

Table 1.1  
*Summary of the Diagnostic Categories Used to Describe Aphasic Speech in the BDAE*

	Key Characteristics
	<i>Non-Fluent Aphasias</i>
Broca's	Speech production is limited to very short phrases Utilise a limited variety of grammatical constructions Articulation difficulties Dysprosody (flat melodic intonation) Impaired repetition of complex sentences Relatively good auditory comprehension
Transcortical Motor	Poor sentence production Difficulty initiating speech Have more success responding to highly structured questions than open-ended questions Good sentence repetition Good articulation Variable melodic intonation Few paraphasias (speech errors) Good auditory comprehension Occasional spontaneous production of grammatically correct sentences
Global	Impaired speech production, some may be unable to produce spontaneous speech Impaired speech comprehension May repeatedly produce over-learned phrases or nonsensical strings of sound, May be able to give yes/no responses to questions pertaining to particular topics, such as their family, their illness, or recent events
	<i>Fluent Aphasias</i>
Wernicke's	Comprehension difficulties Produce a high proportion of paraphasias in running speech Good articulatory agility Good phrase lengths Good use of different grammatical forms



Table 1.1  
*Summary of the Diagnostic Categories Used to Describe Aphasic Speech in the BDAE*

	Key Characteristics
Wernicke's cont.	Good melodic line Speech is often described as devoid of content, or empty sounding
Conduction	Good sentence production, although sentences may be disrupted by word finding difficulties Repetition is worse than spontaneous speech production Good comprehension Produce some phonological and formal paraphasias
Anomic	Word finding difficulties Intact repetition Mild comprehension problems Produce few paraphasias

After the empirical observations are reviewed, theory of language production is summarised with reference to how current theories can, and cannot, be used to explain speech patterns that occur after linguistic breakdown. Because of the limited ability to explain context-specific language deficits using current language theories, frontal lobe theory is also explored as it relates to issues raised in the research on context-specific language deficits, particularly issues of lexical competition and cognitive control. The introduction section concludes with brief descriptions of the experiments included in the current research.

### *Context Effects in Non-Fluent Aphasia*

As mentioned previously, context-specific language deficits have been observed in a number of non-fluent aphasics, and have motivated a number of studies (see; Freedman, Martin & Biegler, 2004; Schwartz & Hodgson, 2002; Schnur et al, 2006; Wilshire & McCarthy, 2002; Wilshire et al., 2006), including the current research. In this section, research that has focused on context-specific language production impairments is summarised.

Arguably the simplest task where a context-specific naming deficit has been reported involved naming one, two, or three pictures in a row. MP, a non-fluent aphasic with a left dorso-lateral frontal lobe lesion, a moderate object naming deficit, and a severe problem with sentence production, was asked to name groups of pictures in a single

utterance (e.g., for single items – “Cat”, for dyads – “Cat, ear”, or for Triads – “Red [square], ear, horse”). Without any cueing, MP scored 92.3 % correct on the single items, 42.3 % correct on the dyads, and only 20 % correct on the triads (Schwartz & Hodgson, 2002). MP clearly found the longer lists more difficult, indicating that she may have found it difficult to produce single words when more than one item was activated simultaneously.

Another example of a single-word production paradigm that has been used to investigate context effects is *blocked cyclic naming* (see: Jefferies et al., 2007; Schnur et al., 2006; Wilshire & McCarthy, 2002). Although blocked naming tasks have varied slightly from experiment to experiment, the basic design is similar to that of traditional object naming tests. Pictures are presented one by one, and participants attempt to name the pictures in separate, single word, utterances. The main differences between traditional object naming and blocked cyclic naming are that in blocked cyclic naming, the items are presented in semantically unrelated or semantically related groups, and each group of items is repeated a number of times before the next group of items is presented: (e.g., *Rabbit, Train, Coat, Knife, Banana, Desk, Coat, Banana, Rabbit, Knife, Desk, Train, etc.*, cf., *Plane, Lorry, Car, Bus, Bike, Train, Bike, Bus, Lorry, Plane, Car, Train, etc.*). The rationale behind the semantic blocking manipulation is that if simultaneous co-activation of items causes interference, this interference should be worse if the co-activated items are closely related. This is because as each target item is activated, so are the other members of its semantic group due to automatic spreading activation. This increase in competitor activation/interference should lead to a performance cost when items are presented in semantically related groups. The rationale behind the cyclic repetition of the items is similar, in that the repeated presentation is expected to increase the lexical activation levels for the items, and to result in a performance cost. A final typical difference between traditional object naming and blocked cyclic naming is that in blocked cyclic naming the rate of picture presentation is manipulated, and naming success at each rate is compared. For example, the pictures may be named at the participant’s own pace, or a new picture may be presented every five seconds, every two seconds, or every second. These rate manipulations have varied across experiments, but they always serve several functions. Most obviously, naming has to occur in rapid succession in the fast pace conditions, so although no sentences are produced, the speech production conditions are more like those seen in connected speech. Also, the pace conditions provide a coarse method of measuring response latencies because on any given trial the response must be produced within the given time frame, or the trial is scored as incorrect.

Wilshire and McCarthy (2002) presented a blocked cyclic naming experiment. BM, a non-fluent aphasic with good object naming but sparse and fragmented sentence production, made significantly more correct responses on unrelated groups than semantically related groups when he was naming at a fast rate (one item every two seconds). He also made significantly more correct responses at a slow rate (one item every three seconds) than a fast rate. In a second, highly similar, experiment, Wilshire and McCarthy compared BM's performance to that of a fluent control, mild anomic IG. IG had a temporal lobe lesion, and although he had some word finding difficulties his sentence production was good. Again, BM showed a significant semantic blocking effect, producing more correct responses on the unrelated sets than the semantically related sets. In contrast, IG did not show a semantic blocking effect even though he made more errors overall. IG's results suggested that context-specific naming deficits may not occur in aphasics who do not have non-fluent speech profiles – although many more patients need to be tested before any clear conclusions can be made (but see also Biegler, Crowther, & Martin, 2006; McCarthy & Kartsounis, 2000).

Schnur et al. (2006) completed a blocked cyclic naming task in a group setting: They tested 12 neurologically intact control participants, 7 Broca's aphasics, and 11 non-Broca's aphasics (this group included anomic aphasics, Wernicke's aphasics, and conduction aphasics). The control group was not influenced by semantic blocking condition or presentation rate, but the patient groups were significantly more accurate on the unrelated groups than the semantically related groups, and they were significantly more accurate when items were presented slowly than when they were presented quickly. Further, a significant interaction of patient group and semantic condition was observed: The Broca's aphasics showed a larger and more consistent semantic blocking effect than the non-Broca's aphasics. However, one patient from each patient group had scores that went against the general trend: One Broca's patient scored higher in the semantically related condition than the unrelated condition, and one non-Broca's patient showed a semantic interference effect that was well within the range shown by the Broca's aphasics. Another interesting pattern observed was that for some Broca's aphasics, it took several naming cycles for the semantic blocking effect to appear, (e.g., BT, DD, NQ; see also NY (Jefferies et al., 2007)), whereas other Broca's aphasics showed a consistent semantic blocking effect across all naming cycles (e.g., CT; see also PG (Jefferies, et al.); and BM (Wilshire & McCarthy, 2002)). This may indicate that some Broca's aphasics are more sensitive to semantic blocking effects than others, or that these effects may occur as a result

of different underlying impairments. However, more research is required before any clear conclusions can be made.

Another task that is highly similar to blocked cyclic naming is *blocked naming*. The difference between these tasks is that in the latter the groups of items are not repeated. Schwartz and Hodgson (2002) compared MP's results on a blocked naming task to her results on a *picture description task*, where she was asked to describe scenes containing the items from one of the item groups used in the blocked naming task (i.e., a group of animals from the blocked naming task all appeared together in a picture of a farmyard in the picture description task). On the blocked naming test, MP's naming accuracy was not influenced by a semantic grouping manipulation, however she was significantly more accurate on self-paced trials than slow trials, and on slow trials than fast trials. Consistently, the semantic relatedness manipulation did not influence MP's naming accuracy in the picture description task. Interestingly though, she named significantly more items in the self-paced condition of the blocked naming task than the picture description task even though she did not need to produce sentence-like utterances to be given credit for naming the items in the latter.

Schwartz and Hodgson (2002) also used a *cued-dyad naming* task to trigger MP's context-specific naming deficit. In this task, two empty boxes were presented on a computer screen, after which two pictures were presented inside the boxes. MP was required to name one of the pictures which was identified by a cue (the picture's border became bolded), and was given five seconds to complete each trial. Two different manipulations were included in this experiment: 1) cue timing, and 2) picture relatedness. In the cue timing conditions, the delays between the presentation of the cues and the pictures were modified. In the pre-cue conditions the cue occurred before the pictures appeared (1500 ms, 1000 ms or 500 ms earlier). In the simultaneous-cue conditions the cue occurred at the same time or after the pictures appeared (0 ms, 500 ms, 1000 ms, or 1500 ms later). Schwartz and Hodgson found that MP showed a highly significant cuing condition effect, making more correct responses in the pre-cue conditions than the simultaneous-cue conditions. They suggest that this occurred because the early cues identified which picture to name before the pictures were presented allowing MP to focus on one picture, whereas in the post cue conditions MP would prepare both object names during the lag between the picture presentation and the cue presentation, leading to the decreased naming success. In the picture relatedness manipulation the objects were semantically related (e.g., *bed – chair*), phonologically related (e.g., *sun – nun*),

semantically and phonologically related (e.g., *truck – train*), or unrelated (e.g., *shoe – kite*). Overall the difference in MP's performance between the relatedness conditions approached significance, but she showed no significant difference between the items that were semantically related and the items that were not. However, she did do significantly better when the items were phonologically related than when they were not.

Freedman et al., (2004) completed a *picture naming* task that triggered context-specific naming deficits. Their participants included two aphasics with good single picture naming and poor sentence production, who showed clear dissociations between preserved semantic knowledge and impaired semantic short term memory (ML and GR). For comparison, they also tested an aphasic with fluent speech and excellent comprehension but who makes occasional phonemic errors and has poor sentence repetition ability and a phonological short term memory (STM) deficit (EA), and a group of 12 neurologically intact control participants. In this task, either one picture was presented and the participant was asked to name that picture, or two pictures were presented and the participant was required to name both objects (e.g., “*nose and hat*”). Under the two-picture condition the items were either unrelated, or semantically related. Once the participant started responding, the pictures were removed, and the participants had to rely on their STM to complete the utterance. All of the participants were most accurate in the single picture condition but showed no accuracy differences between the semantic pairs and the unrelated pairs. All of the participants also responded the fastest in the single picture condition. Non-fluents ML and GR and the control group had significantly faster response latencies when they were naming the unrelated pairs than semantically related pairs, although the magnitude of the control's semantic blocking effect was much smaller than that shown by non-fluents ML and GR. Fluent EA did not show a significant difference between the unrelated pairs and the semantically related pairs.

Wilshire et al. (2006) presented a case study on patient JHM, who has good single word naming but impaired sentence production after a left middle cerebral artery stroke. In a *two-picture naming task*, a picture was presented on a screen, and then a second picture replaced it. Once both pictures disappeared JHM was cued to name the pictures (e.g., she would hear a beep and see “\_\_\_\_ and \_\_\_\_”). The picture pairs were semantically related (e.g., *goat – pig*), phonologically related (e.g., *goat – ghost*), or unrelated (e.g., *goat – ball*). JHM's response accuracy was high in all three conditions (96 % correct for semantically related pairs, 95 % correct for unrelated pairs, and 91 % correct for phonologically related pairs), but her response latencies were significantly longer in the

semantically related condition than the phonologically related and unrelated conditions. She showed no significant difference between the phonological and unrelated conditions. In a *picture-word interference task*, JHM was shown a series of pictures, and was required to name the pictures. However, a semantically related, phonologically related, or unrelated verbal distracter was presented with each picture. JHM's naming latencies showed significant semantic and phonological interference (controls showed significant but smaller semantic interference, but a trend towards phonological facilitation). Once again, JHM's accuracy was high in every condition (90 – 94 % correct).

A theme that runs through much of the research discussed thus far is that when stimuli are semantically grouped, it can be harder to name than when it is not semantically grouped, possibly because of increased levels of competition. Another task that manipulates competition that may be of some relevance to research on context-specific language impairments, even though it has typically been completed with normals or frontal lobe patients, is *the Stroop (1935) task*. In this task participants are shown colour words (e.g., blue) that are printed in different ink colours, and are asked to either read the word or name the colour of the ink. Whether participants are required to name the colour of the ink or read the word, they are faster if the printed word and the ink colour are congruent. High levels of response competition appear to be a factor in the incongruence costs observed. In the incongruent condition, participants need to make a choice between the two highly salient stimulus features which illicit two critically different responses, whereas in the congruent condition, both features point to the same response. Although neurologically intact participants show competition effects on this task, it is possible that people who have difficulty resolving competition may show exaggerated effects.

Hamilton and Martin (2005) present Stroop task data from non-fluent ML, who was discussed previously on p. 7. ML showed exaggerated interference effects on this task: he showed a bigger difference between congruent and incongruent trials than a control group. Interestingly, ML and controls were also given a non-verbal spatial Stroop. In this variation, arrows pointing to the left or right were presented on the left, middle, or right of a display, and the participants had to indicate with a key press the direction the arrow was pointing. An example of a congruent trial is where a right pointing arrow was presented on the right of the screen, whereas an example of an incongruent trial is where a right pointing arrow was presented on the left of the screen. On the non-verbal spatial Stroop ML performed within the normal range. This suggests that he may be particularly susceptible to competition effects when completing tasks with linguistic components.

Another example of a task that manipulates competition that has not primarily been used in aphasia research but may be relevant to research on context-specific language deficits, is the *verb generation task*. In this task participants are shown a series of items and, for each item, they are required to say the first semantically associated verb they think of. For example, presented with the word *scissors*, a participant could say, “Cut”, or presented with the word *penny*, a participant could say “Spend”, “Pay”, “Buy”, “Flip”, or “Drop”, etc.) (see Thompson-Schill, D’Esposito, Aguirre, & Farah, 1997; Thompson-Schill, D’Esposito, & Kan, 1999; Thompson-Schill, Swick, Farah, D’Esposito, Kan, & Knight, 1998; Martin & Cheng, 2006). Items like *scissors*, which have one dominant verb associate (i.e., one dominant response), should have few competitors. However items like *penny*, which have more than one verb associate (i.e., multiple possible responses), should have more competitors.

Although it wasn’t conducted within the context of aphasia research, Thompson-Schill et al. (1998) completed a lesion study where the verb generation task was given to four participant groups: patients with left inferior frontal gyrus (LIFG) lesions; patients with left prefrontal cortex lesions that excluded the LIFG; patients with right prefrontal cortex lesions; and a group of elderly controls. Patients with lesions to the LIFG produced more errors than the other groups in the low response strength condition, and there were no significant differences between the groups in the high response strength condition. This result suggested that the LIFG may have a role in selecting between competing responses. Thompson-Schill et al. (1997) also completed an fMRI study that provided corroborating evidence, as normal participants showed increased activation in the LIFG in a low response strength condition. Because the LIFG corresponds with the anterior language area that is typically damaged in non-fluent aphasia, it appears likely that non-fluent aphasics may also be particularly susceptible to response strength/competition manipulations in this task.

In addition to the fMRI study summarised above, an increasing number of studies have been published recently that have used neuro-imaging techniques to explore the anatomical correlates of semantic selection processes, and have concluded that the LIFG has a role in competition resolution (see Chan, Liu, Yip, Fox, Gao, & Tan, 2004; Copland, de Zubicaray, McMahon, Wilson, Eastburn, & Chenery, 2003, Grindrod, Bilenko, Myers, & Blumstein, 2008; Hoenig & Scheef, 2005; Ihara, Hayakawa, Wei, Munetsuna, & Fujimaki, 2007; Lee & Dapretto, 2006; Mason & Just, 2007; Rodd, Davis, & Johnsrude, 2005; Zempleni, Renken, Hoeks, Hoogduin, & Stowe, 2007). The tasks used in the neuro-imaging research tend to differ from those used in the language research summarised

previously, in that they are usually not language production tasks, however similar themes are explored and similar conclusions have been reported. For example, Grindrod et al. (2008) conducted an event-related fMRI study where participants were asked to make a word/non-word decision about the third word of a triplet. In a condition where it was possible to interpret the meaning of an ambiguous word because of the context the word was presented in (e.g., coin – money – mint), LIFG activation increased relative to a condition where interpretation of the ambiguous word was not possible (e.g., candy – money – mint). Hence, Grindrod et al. concluded that the selection of the appropriate meaning engaged the LIFG. In another example, Ihara et al. used magnetoencephalography (MEG) to observe that left anterior inferior frontal cortex activity increased during reading of ambiguous words.

In summary, a number of studies have shown context-specific naming deficits in aphasic patients. Most of the research into these deficits falls into two categories. It either aggregates across large groups of patients, or it examines individuals that have been carefully selected, usually on the basis of their performance on tasks that may be related. Most of the cases who have shown significant effects have had good single word production, and relatively bad sentence production. However, the precise cognitive determinants of these types of effects are still poorly understood.

In the current research, we intend to investigate what these types of impairments might tell us about the processes of lexical competition and selection during speech production. We will consider some more specific explanations for the various observations in greater detail below. However, before these explanations are thoroughly explored as they relate to current theories of language, several language theories are summarised in the next section.

### *Language Theories*

Spoken language is an extremely complex and versatile process. Consider the processes involved in word generation. At the most basic level, semantic representations of what we intend to communicate must become active. These concepts must be defined as words, and the relevant phonological and prosodic information must be articulated in the correct order. If a multi-word utterance is being produced, syntactic information must also be accessed and used to order words into grammatically correct sentences.



### *Theories of Word Production*

Consider the simplest case: single word production. Many current theories of word production are based on a framework that involves a symbolic network of interacting units (e.g., Dell, 1986; Dell, Schwartz, Martin, Saffran & Gagnon, 1997; MacKay, 1987; Roelofs, 1997). Although the finer details of the proposed networks vary across models, in general, they are made up of thousands of nodes that en masse are called the *mental lexicon*. The mental lexicon contains all of the information used in language processing. Each node represents a unit of linguistic information, be it a semantic feature, a syntactic feature, or a phonological feature.

The nodes in the mental lexicon interact by way of spreading activation. Each node is connected to other related nodes. When one node becomes active, it automatically sends excitation, or as the case may be, inhibition, on to related nodes. For example, activation of the node corresponding to the word “CAT”, would trickle through related nodes until it activated the related phonological nodes, /k/, /æ/, and /t/.

The translation of semantic concepts into phonological representations is defined as lexicalisation (see Harley, 2001). Although there is debate regarding the specifics of lexicalisation, in general, it is thought to occur in at least three levels of processing: a semantic level, a lexical level, and a phonological level (for an illustrative example, see Figure 1.1). Hence, two main translation processes are modelled: The translation of semantic representations into lexical representations, and the translation of lexical representations into phonological representations.

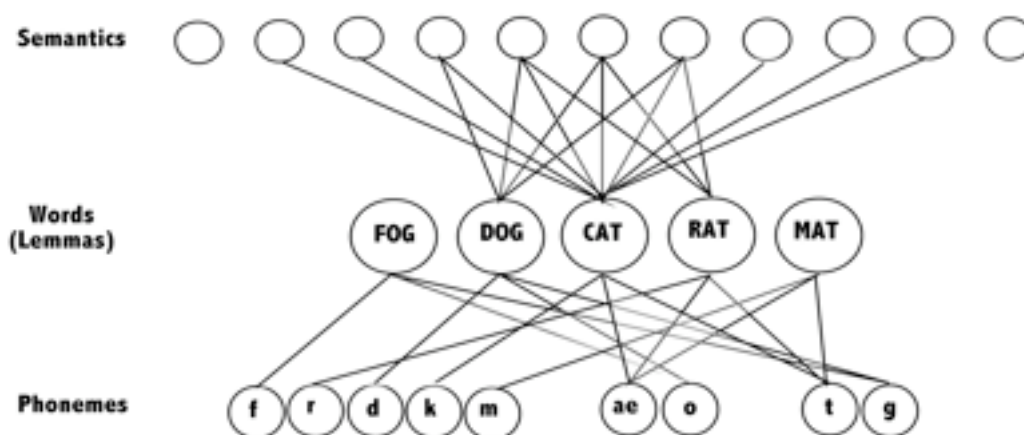


Figure 1.1. Fragment of the Dell et al. (1997) two stage model of lexical retrieval.

At the semantic level<sup>1</sup>, a semantic representation of what we intend to communicate is specified (Dell et al., 1997; Levelt, Roelofs, & Meyer, 1999). In other words, an intention to speak is formed, and the required semantic information is sourced from memory/the lexicon. Necessarily, all language models assume activation of semantic information. However, the format of the semantic level and the process of semantic conceptualisation are not detailed in some models (e.g., Levelt, Schriefers, Vorberg, Meyer, Pechmann, & Havinga, 1991). Other theorists debate how information is stored at the semantic level: Decompositional theorists argue that semantic level representations consist of a range of co-activated semantic features/nodes (e.g., ‘furry’, ‘animal’, ‘meow’,

<sup>1</sup> The semantic level is called different things in different models. For simplicity, the term *semantic level* is adopted here, but other variations include; the message level (Garrett, 1975); the conceptual stratum (Roelofs, 1997); and the conceptual level (Levelt, Schriefers, Vorberg, Meyer, Pechmann & Havinga, 1991). A variety of terms are also used to describe the lexical and phonological levels.

‘four legged’) (e.g., Caramazza, 1997; Dell et al., 1997), whereas non-decompositional theorists argue that representations at the semantic level are unitary (e.g., ‘cat’) (e.g., Roelofs, 1992; Levelt et al., 1999). There is agreement though that the semantic activations at this level are pre-verbal because, at this point, the syntactic and phonological specifications of the intended utterance can remain unspecified.

At the lexical level, each word is represented as a single unit. The syntactic features of each word/unit are specified, but not the phonological characteristics (Dell, 1986; Dell et al., 1997; Jescheniak & Levelt, 1994; Kempen & Huijbers, 1983; Levelt et al., 1991, 1999; Roelofs, 1992). These units are often referred to as *lemmas*. Activation of concepts at the semantic level flows, through spreading activation, to related lemmas at the lexical level. Returning to the example of producing the word ‘cat’ using a decompositional theory, the active semantic nodes (‘furry’, ‘animal’, ‘meow’, ‘four legged’, etc.) would all send an excitatory message down to the lemma for ‘CAT’ (excitation would also pass to other related lemmas, such as the dog lemma since dogs are also furry animals that have four legs, but the activation of the lemma ‘CAT’ should receive the most activation).

Not all language models include a lexical level. Some theorists suggest that there are direct connections from semantics to phonological level word-forms (e.g., Caramazza, 1997; Glaser & Glaser, 1989; La Heij, 1988; Seidenberg & McClelland, 1989; Starreveld & La Heij, 1995). For example, Caramazza (1997) suggests that semantic representations activate modality specific word-forms (called *lexemes*) at the phonological level, which activate the words phonological features at the same level. In a fashion though, the ‘lexeme’ selection forms an intermediary stage between semantics and phonetic specification, so the debate really concerns what processes occur at the intermediary level, rather than the existence of the intermediary level as such. The Seidenberg and McClelland (1989) parallel distributed processing (PDP) model of word recognition and naming provides another example of a model that does not include a lexical level. In this case however, the orthographic and phonological units are connected via hidden units, and it can be argued that the hidden units form an intermediary stage.

At the phonological level, active lemmas are translated into phonologically specified forms. For example, in Dell and colleagues’ model (Dell et al., 1997), the lemma ‘CAT’ sends excitatory messages down to the phonological nodes /k/, /æ/, and /t/. Some models include two stages of processing at the phonological level: First, morpheme selection, and second, phoneme selection (e.g., Caramazza, 1997; Roelofs, 1997). Once

the phonological components of an utterance are specified, articulation can begin (the motor movement processes involved during speech articulation are not detailed in most speech production models).

### *Theories of Sentence Production*

Despite the commonalities shared across language models, different theorists present the precise mechanisms involved in word and sentence production, particularly syntactic processing, in quite different ways. Here we summarise four popular models of language: the Garrett (1975) model of sentence production; the Levelt et al. (1999) theory of lexical access in speech production; the Dell (1986) model of sentence production; and the MacKay (1987) node structure theory of sequencing. The Dell and MacKay models have been selected for more detailed discussion because they both include mechanisms that are of direct relevance to the processes of lexical competition and selection: First, spreading activation, which is important because it accounts for how competition builds up across the lexicon; and second, controlled selection, which is important because it addresses how words may be processed when competition levels are high.

The examples discussed in the previous sub-section applied to single word production. Sentence level processes are also accounted for in most language production models, which aim to account for sentence level error patterns. For example, Garrett (1975) based his model of sentence production on speech errors that he, his friends, colleagues, and students made or heard during the course of normal conversations. Study of these errors revealed several important patterns. First, it was noted that different sized linguistic units slip, including phonemes, consonant clusters, morphemes, and whole words (see Table 1.2 for a list of illustrative examples). Second, it was noted that speech errors frequently involve the misordering of linguistic units within sentences. Further, Garrett noted that word exchanges usually preserve grammatical category and function (e.g. nouns swap with nouns, verbs swap with verbs, etc.), and words that are quite distant within the utterances can be exchanged. In contrast, sound exchanges (e.g., phoneme exchanges) occur between words with different grammatical functions, and usually occur between adjacent or close words. Garrett focused on these different linguistic category constraints, and used the observed patterns to constrain his sentence production model. For instance, because syntactic errors and phonological errors operate in dissimilar fashions, Garrett suggested that syntactic functions are assigned at one level (the functional level), and

sound forms, like morphemes and phonemes, are assigned at another level (the positional level).

Table 1.2  
*Examples of Speech Errors Found in Speech Error Corpora*

Example Number	Error Type	Example
1	Phoneme Exchange	York library → lork yibrary
2	Consonant Cluster Exchange	snow flurries → flow snurries
3	Morpheme Exchange	thinly sliced → slicely thinned
4	Word Exchange	writing a letter to my mother → writing a mother to my letter
5	Word Anticipation	the sun is in the sky → the sky is in the sky
6	Non-contextual Substitution	pass the pepper → pass the salt
7	Non-contextual Addition	the only thing I can do → the only one thing
8	Non-contextual Deletion	I just wanted to ask that → I just wanted to that
9	Word Blend	athlete/player → athler

Note: These errors were sourced from examples provided by Dell (1986).

In all, Garrett (1975) proposed that four levels of processing are completed during sentence production (see also Garrett 1976, 1982). Each level is processed independently, starting at the highest level and progressing down to the lowest level, hence his model is described as being discrete and serial. Garrett calls the highest level the message formulation level: this is where pre-linguistic semantic processes are completed. The second level is called the functional level: this is where lexical selection of open-class words occurs, and where grammatical and thematic roles are assigned. The third level is called the positional level: this is where syntactic frames are constructed, where closed-class words are selected, and where all words become phonologically specified. The bottom level is called the phonetic and articulatory level: this is where the mechanics of speech processing is completed, but this process is not clearly detailed in the model.

However, there are several error patterns that the Garrett (1975) model does not predict. One such pattern is sentence blends (e.g., “I’m making the kettle on” (example from Harley, 2001)) (Butterworth, 1981). Sentence blend errors, as well as some other error types that are not detailed here, suggest that processing of two sentences occurs in parallel, rather than being discrete.

Like Garrett (1975), Levelt et al. (1999) present a discrete serial theory of lexical access in speech production. However, unlike Garrett's model, Levelt and colleague's model also contains a post-lexical editing system that checks for differences between intended messages, and output that has been selected during the planning process before speech production. Levelt and colleagues propose that this editing system is least likely to notice errors that are semantically and phonologically similar to the intended message, meaning that errors like sentence blends are more likely to be produced than non-words for instance.

Discrete serial models, such as those by Garrett (1975) and Levelt et al. (1999), can be contrasted to interactive models. In interactive models, such as the Dell (1986) model of sentence production and the MacKay (1987) theory of sequencing, processing occurs at different levels simultaneously, so processing down-stream can influence processing upstream as well as the other way around. These models still propose that semantic processing occurs relatively early, and phonological processing occurs relatively late, but because processing is potentially simultaneous, interactive models account for more error types than discrete serial models, without the inclusion of a post-lexical editing system. Hence interactive models are arguably more parsimonious. The Dell and MacKay models are discussed next.

*The Dell (1986) Model of Sentence Production.*

Dell (1986) presents a model of sentence production that focuses in particular on the phonological encoding of speech, and attempts to account for speech error data. Dell's model has been particularly influential because it is one of few models that have been computerised, and the computer model successfully mimics most of the errors that occur in spontaneous speech.

As is typical in language production models, the Dell (1986) model converts semantically specified messages into series of sounds via levels of processing: A conceptual level (semantic), a syntactic level (lexical), and a morphological level and/or phonological level. Activation spreads within and throughout the levels via spreading activation. Once activated, nodes at all levels of the lexicon quickly spread activation on to all other related nodes. With time, these activations decay exponentially, to keep activation levels within the lexicon down.

Because this is a model of sentence production, syntactic processes are addressed, in addition to the spreading activation processes addressed in the previous section on single word production. As well as lemma selection, Dell (1986) suggests that syntactic planning

occurs at the lexical level. It is suggested that each node in the conceptual, syntactic, morphological and phonological levels is labelled with an insertion rule. These insertion rules include information stipulating the position that the units can take within sentences (i.e., the syntactic, phonological, or morphological status of the linguistic units). These insertion rules/labels are important because they form a link between the non-productive knowledge that is stored in the lexicon, and generative rules.

Generative rules are expressed in the form of frame-and-slot representations that constrain how the various nodes within each level of processing may be combined together. The generative rules are different for each level because they dictate which combinations of linguistic units are allowed at each level of language specifically. Using the sentence “This cow eats grass” as an example, Dell suggests that the generative rules at the syntactic level would form the frame, *Determiner – Noun - Present-tense verb - Noun*. Whereas, the generative rules at the phonological level for the words “This cow” would form the frame, *Initial consonant – Vowel - Final consonant – Initial consonant – Vowel*.

The slots in the frames formed by generative rules are filled by the most highly activated node that fulfils the insertion rules for that position. The selected nodes are assigned an order tag that places them in their relevant positions within the current syntactic frame. When a node is selected for insertion into the frame it receives an extra jolt of activation from the syntactic structure, called signalling activation. The signalling activation flows down to the next level, giving related nodes an additional competitive advantage.

In the Dell (1986) model, higher order levels, such as semantics, are activated before the corresponding lower order levels, such as phonological encoding. The selection of items for lower level representations can only occur after processing at higher levels has begun – because, for instance, you cannot select the phonological units of a word before its syntactic role has been established. However, it is important to note that although processing generally flows from the high level representations towards the low level representations, in practice processing can occur at all levels simultaneously. This is because during the planning process, the construction of the early portion of the sentence may be nearing completion, while the latter portion of the sentence may still be at a relatively early stage of planning: Also, while the process starts with selection of high level items that influence the selection of low level items, as the process develops, low level items send activation back up to connected nodes at high levels. In this way, low level units can influence the final selection of units at higher levels by biasing activation towards

nodes that were not the most highly activated when influenced by downwards activations alone. The final selection of items at each level is influenced by the patterns of activation in higher and lower levels.

The speech errors that the Dell (1986) theory accounts for are predicted to occur as a result of computational simultaneity, which occurs when more than one item that is labelled with the same insertion rule is activated at the same time. The theory nicely explains the categorical constraints that have been observed in speech error corpora. Computational simultaneity accounts for the manner in which interacting units in speech errors tend to be of the same linguistic category, in that items are selected from the lexicon but are mistakenly assigned to the wrong slot within the syntactic frame. To illustrate this point, please refer back to the speech errors presented in Table 1.2. Examination of example number three (*thinly sliced* → *slicely thinned*) suggests that this error has occurred during morphemic processing. The morphemic frame-and-slot model generated for the words “*thinly sliced*” would be *Word-Stem – Affix – Word-Stem – Affix*. Dell’s model predicts that the word-stem and affix slots will be filled by the most highly activated item in the lexicon that is labelled with the correct insertion rule. Errors like “*slicely thinned*” will occur when, for whatever reason, activation of the word-stem unit “*slice*” was more highly activated than the word-stem unit “*thin*”, when the first word-stem in the utterance was being selected. Examination of Example Number 6 in Table 1.2 (*pass the pepper* → *pass the salt*) indicates that errors can also be non-contextual, in that they do not always come from within the intended utterance. In this case, spreading activation throughout the lexicon has resulted in “*salt*” being more highly activated than the intended word “*pepper*” during the selection of the relevant linguistic unit. This error may have occurred even if the unit for “*pepper*” was tagged at the semantic level, as the spreading activation through the lexicon is unconstrained, and because salt and pepper are closely related, activation of one of these units would undoubtedly activate the other in turn.

To summarise the Dell (1986) model, the structure of sentences is formed by productive generative rules. These rules set out the number of words in a planned sentence, and their syntactic class. A separate structure is generated for the semantic, syntactic, morphological, and phonological linguistic items in the planned sentence. The slots in the planned sentence structures are filled by the most highly activated items in the lexicon. Higher order items are selected first – because they dictate important aspects of lower order items. However, activation at lower levels feeds back to higher levels, and this can result in the modification of higher level selections.



*The MacKay (1987) Theory of Sequencing.*

MacKay (1987) presents the *node structure theory of sequencing*, which attempts to account for the manner in which rapidly produced actions are organised for production. Language, as a highly flexible and rapidly produced behaviour, was adopted as the basis for the theory. As is typical in language models, MacKay adopts a node structure approach. In comparison to the Dell (1986) model of sentence production though, the nodes in the MacKay model have more complex and varied roles. Where syntactic processing in the Dell model involves slotting activated nodes into syntactic frames, dedicated syntactic processing nodes complete semantic planning/processing in the theory of sequencing.

In the theory of sequencing, nodes vary in two respects. First, as usual, each node contains a single, unique unit of information. But second, the type of information contained within the nodes varies. Nodes form three categories that contain three different types of information: *Content nodes* store non-productive knowledge and are selected when passive information is required; *sequence nodes* store syntactic rules and dictate the order in which other nodes in the lexicon are selected; and *timing nodes* form the mechanism which allows for the activation of sequence nodes. Each of the node types has a specific function, and will be discussed in more detail in turn.

*Content Nodes.* Content nodes form the storehouse of semantic information within the lexicon. Each content node is linked to other related nodes, by way of simple, excitatory, non-multiplicative connections, and by way of negative, quenching connections. The excitatory connections automatically become active when a node is stimulated. Hence, when it is active, the content node for “Cat” will automatically send an excitatory message to all other related content nodes, like the semantically related node for “Dog”, and the phonologically related nodes, “ ”, “ ” and “ ”. Stimulation onto content nodes is additive, in that if one node is being simultaneously stimulated by two other nodes, the net activation of that node will be greater than if the node was only being activated from one other source. Once a content node has been activated, it sends a negative quenching message to related nodes: This ensures that the overall stimulation levels within the lexicon are kept down at a controlled level.

An important feature of content nodes is that they are hierarchically organised. This hierarchy places a *sentential system* at the top, followed by a *phonological system* in the middle, and a *muscle movement system* at the bottom. In other words, the hierarchy

forms separate levels of processing. Information about individual concepts, words, phrases, and parts of sentences is stored within the sentential system, and information about parts of words is stored within the phonological system (details of the muscle movement system are not articulated within the theory).

It is also vital to note that content nodes form domains. The domains form a way of organising the nodes within both the sentential system and the phonological system. A separate domain exists for each of the different parts of speech. For instance, the words “door” and “window” both belong to the *noun domain*; the words, “prove” and “run” both belong to the *verb domain*; and the word “lift” belongs to the *noun domain* (as the name for an elevator) and the *verb domain* (as the action of raising an object). To illustrate the domains present at the phonological level, consider the word “prove”. Here the phoneme “p” belongs to the *initial stop domain*, and the phoneme “r” belongs to the *initial liquid domain*. At a higher level within the phonological system, the sound “pr” belongs to the *initial consonant group domain*.

*Sequence Nodes.* Previously it was mentioned that sequence nodes store syntactic rules. For instance, a sequence node exists for the concept ‘adjective’, and another sequence node exists for the concept ‘noun’, etc. The other feature of sequence nodes that was mentioned was that they dictate the order in which content nodes are selected. Although content nodes send excitatory signals to each other, they cannot become activated without input from sequence nodes. This is because sequence nodes send multiplicative messages to connected content nodes (this multiplicative activation is the equivalent of signalling activation in Dell (1986)). This multiplicative activation causes any active content nodes to become more highly activated – at a rate that corresponds to their activation level before stimulation from the sequence node. This means that the simple, additive stimulation that content nodes send between themselves is quickly increased to the point where activation can occur, when a related sequence node is activated.

Another important feature of sequence nodes is that they, like content nodes, are divided into domains. Three sequence node domains are identified within the theory: The first corresponds to the sentential system; the second corresponds to the phonological system; and the third corresponds to the muscle movement system. Also, sequence nodes function to organise content nodes into their domains.

The connections between sequence nodes differ from those between sequence and content nodes, and those between sequence and timing nodes. The connections between

sequence nodes are inhibitory: When activated, sequence nodes inhibit all other nodes within their domain. Hence, activation of the sequence node ‘verb’ will inhibit activation of the sequence nodes ‘noun’ and ‘adjective’, along with all other sequence nodes in the sentential system. The connections between sequence nodes have particular importance in the model, because they represent serial-order rules. If more than one sequence node receives stimulation simultaneously, serial-order rules dictate which will be activated first. For example, in English, adjectives come before nouns. If the ‘adjective’ and ‘noun’ sequence nodes are activated simultaneously, a serial-order rule dictates that ‘adjective’ should be selected first. Hence, ‘adjective’ sends an inhibitory message to ‘noun’, ensuring that ‘adjective’ is activated first.

*Timing Nodes.* Timing nodes function to activate sequence nodes, so indirectly, they determine the rate of language production (because sequence nodes must be activated before content nodes are activated). Through direct connections, timing nodes determine the rate at which sequence nodes are stimulated. When the rate of stimulation onto sequence nodes is fast, timing nodes quickly send messages back to the sequence nodes, via multiplicative connections: When the rate of stimulation onto sequence nodes is slow, timing nodes slowly send messages back to the sequence nodes. The multiplicative messages from the timing nodes give the sequence nodes enough of a boost to allow for sequence node activation.

This paragraph summarises how MacKay (1987) proposes that the different nodes interact to produce fluent, connected speech. When content nodes receive excitatory input, they stimulate all related content nodes. Each stimulated content node will in turn stimulate, usually one, related sequence node. For example, the content node for ‘extensive’ would send an excitatory message to all other semantically and phonologically related content nodes (e.g., ‘large’, ‘/I/’, etc), and the sequence node that holds the syntactic rule for the relevant part of speech (e.g., ‘adjective’). The stimulation of sequence nodes is passed onto timing nodes, which send multiplicative stimulations back to the sequence nodes at a rate that is dependent on the speed at which the sequence nodes are receiving stimulation from the content nodes. The multiplicative stimulation sent to the sequence nodes from the timing nodes allows the most highly stimulated sequence node to activate. The activated sequence node then sends a multiplicative message back to all of the content nodes in the relevant domain. The stimulation levels in the relevant content nodes raises sufficiently to allow for the activation of the most highly stimulated content nodes, and as a result, the production of rapidly sequenced behaviours.

Aphasia researchers have attempted to use models of language, such as those described above, as frameworks for understanding different kinds of aphasic speech breakdown. The models can be used to describe a lot of aphasic profiles because breakdown of each component of a model predicts a different speech deficit, plus breakdown of more than one component will lead to a different impairment than a more specific impairment. The following section starts with a general discussion of how aphasic deficits can be accounted for within language models, and is followed by a detailed consideration of how context effects may be explained.

*Language Models and Aphasia*

People with aphasia show very diverse patterns of linguistic breakdown. Many of these patterns are consistent with a deficit to one or more of the specific mechanisms proposed in models of language. The behavioural ramifications of specific impairments to the semantic, lexical, and phonological levels of processing, as well as the syntactic features of the models, are explored next, along with brief descriptions of cases who show these kinds of impairments.

Case studies on a number of aphasics who show error patterns that are consistent with inaccessible or degraded semantic level processing have been published. For example, Howard and Gatehouse (2006) presented the case of JGr, who presented with semantic errors during naming, and impaired comprehension on a range of tests including word-to-picture matching, synonym judgements, and the Pyramids and Palm Trees test (see Appendix A2 for a task description). Further, JGr's picture naming performance improved when he was given phonemic cues (e.g., /də/ for the target *dog*), whereas miscues (e.g., /kə/ for the target *dog*) lead to a decrease in correct responding (see also patients KE (Hillis, Rapp, Romani, & Caramazza, 1990), and JCU (Howard & Orchard-Lisle, 1984)). Under conditions of semantic impairment, semantic production and comprehension impairments are expected because patients are unable to retrieve full semantic representations. For example, a reduced semantic representation for cat may consist of activation of the 'four-legged', 'animal', and 'pet' nodes without the 'purr' and 'climb trees' nodes: this may lead to a patient incorrectly saying 'dog' instead of 'cat', or in the case of comprehension, confusing the two as the likelihood of a non-target item meeting the reduced semantic description increases (see; Butterworth, Howard, & McLoughlin, 1984; Howard & Gatehouse, 2006). Phonemic cues are predicted to aid performance, because they provide an additional means of limiting the number of alternatives that fit the reduced semantic specification. On the other hand, patients are expected to be particularly susceptible to phonemic miscuing, because a reduced semantic representation would forgo affective limitation of semantically related alternatives (see Howard & Gatehouse).

Case studies on a number of patients who show error patterns that are consistent with a selective lexical level impairment have also been published. For example, Howard and Gatehouse (2006) presented the case of LM, who presented with no-response errors during naming but very few semantic errors, and intact comprehension on a range of tests

including word-to-picture matching, synonym judgements, and the Pyramids and Palm Trees test (see Appendix A2 for a task description) (see also patients GM and JS (Lambon Ralph, Sage & Roberts, 2000)). Under conditions of lexical impairment a range of whole word speech production errors are expected because disrupted lexical processes may result in no lemma selection, or alternatively, disruptions to the system may result in a non-target lemma being first to reach the critical threshold. Because most of the input into the lexical level comes down from the semantic level, the models predict that some semantic errors may occur, as semantic alternatives to the target will simultaneously be receiving input from semantics. For example, active semantic nodes for the concepts '*four-legged*', '*animal*', '*pet*', '*purr*' and '*climb trees*', send activation down to the target lemma '*cat*'. However, the semantic nodes '*four-legged*', '*animal*', '*pet*' also send activation down to the non-target lemma '*dog*'. If damaged processes render the lemma '*cat*' ineffectively activated or unavailable, the lemma '*dog*' is likely to be selected in its place as a result of the shared semantic features. In interactive models, the lexical level also receives feedback activation from the phonological level, so phonological processes may also influence lemma selection (see Dell, 1986; Dell et al., 1997). The feedback from the phonological level may be directed back to phonemically related lemmas, resulting in formal or mixed errors. However, because processing at the phonological level remains intact, once a lemma is eventually selected, the phonological structure for the selected lemma should remain intact, so lexical level errors should involve whole words. An important feature of a selective lexical level impairment is that intact comprehension is expected, because the semantic level remains intact. This feature is important because it means that patients with lexical level deficits can be differentiated from those with semantic level deficits.

Case studies on a number of patients who showed error patterns that are consistent with a selective phonological level impairment have also been published. For example, Rapp and Goldrick (2000) present the case of CSS, who presented with good spoken and written comprehension, but semantic errors, phonologically similar word responses, and non-word responses in naming and repetition. Under conditions of phonological level impairment sound units may be produced in the incorrect order, or at worst, not at all. For example, consider the utterance, 'York Library': Incorrectly producing the phonemes required for this utterance may result in a phoneme exchange error, like 'Lork Yibrary'. Depending on the severity of the impairment, one or two sound units may be incorrect, resulting in a phonological error (e.g., the target 'table' may be produced as 'bable'), or a formal error (e.g., 'table' may be produced as 'cable'). If the disruption at the

phonological level is severe however, the sounds selected may bear little resemblance to those intended, resulting in production of unidentifiable non-words (neologisms), or no-response errors.

Sentence production deficits are explained somewhat differently. Sentence production requires that more than one word is processed simultaneously, and explanations of sentence production deficits tend to focus on the processes that structure co-activated words into sentences. For example, it has been suggested that agrammatic aphasics are not able to effectively map semantic information onto the syntactic frames that specify the order of word production (Schwartz et al., 1985). Alternatively, the syntactic structure may not be generated properly, resulting in unstructured speech production (Goodglass et al., 1994). Other researchers focus less on the syntactic frames, and more on the timing processes involved in sentence production. One suggestion is that sentence production deficits occur because the nodes required to produce a sentence are not produced quickly enough to be inserted into one syntactic frame (Kolk, 1987, 1995). Another alternative is that the nodes are activated quickly enough, but lose activation at an unusually fast rate, meaning that they deactivate before they are successfully integrated into the syntactic frame (Freedman et al., 2004).

#### *Language Models and Context Effects.*

Although a considerable body of work on aphasia has focused on identifying the specific level of processing impaired in particular patients based on their error patterns, relatively little study has been devoted to the more dynamic, context-specific aspects of production described in earlier sections. However, in this section we will describe three explanations that have been put forward to explain context-specific deficits: The first is that these may be due to an impairment of syntactic processing (Schwartz & Hodgson, 2002); The second is that these may be due to a problem retaining semantic or lexical information during speech planning (Freedman et al., 2004); The third is that these may be due to a difficulty resolving competition between lexical items (Wilshire & McCarthy, 2002; Wilshire et al., 2006).

*The Syntactic Hypothesis.* Many of the tasks that have been used to identify context-sensitive naming deficits could be viewed as involving the production of very simple syntactic structures. Therefore an impairment in creating a syntactic frame and/or assigning lexical items to the frame may account for at least some of the response patterns

that have been observed. The idea of a syntactic frame impairment features heavily in the Schwartz and Hodgson (2002) account of patient MP, who was described previously (see page 3). MP showed a context-specific naming deficit on a range of tasks, including picture naming, composite picture description, naming span, and cued dyad naming. However she did not show a significant effect of semantic blocking, leading to the conclusion that her context-specific deficit was not sensitive to paradigmatic competition (i.e., competition between nodes that are co-activated as a direct result of their relatedness and the spreading activation process (e.g., *pig* receives paradigmatic competition from *hog*, *sow*, *goat*, and *sheep*, etc.)).

Because of the errors that MP made, Schwartz and Hodgson (2002) suggest that her problem is not due to a lemma access deficit. Rather, they suggest that it may reflect increased interference between nodes at the phonological level. To illustrate this idea, consider the following scenario with the Dell (1986) model. Under conditions where multiple words are being produced, relevant lemmas are selected, and a syntactic frame and slot model is created. At this point in an intact system, the most highly activated lexical item tagged with the correct insertion rule is assigned its place within the frame, and receives an extra jolt of activation giving it a competitive advantage against alternative nodes. This advantage flows down to related nodes at the phonological level giving them a competitive advantage. Schwartz and Hodgson suggest that in MP's impaired lexicon, the extra jolt from the syntactic frame may not occur, resulting in increased interference between nodes at the phonological level. In other words, they suggest that MP's multiword production disadvantage is due to retrieval interference between co-activated objects. Specifically, they suggest that MP's syntactic processor is impaired, so when two or more items are co-activated at the phonological level, poorer performance is triggered.

In summary, Schwartz and Hodgson (2002) suggest that impairment to a syntactically based controlled activation mechanism at the lexical level leads to a context-specific language deficit. However, in more recent research, MP has shown a semantic blocking effect, supporting the view that she is in fact more sensitive to paradigmatic competition than controls (she is patient NO in Schnur et al., 2006). The syntactic hypothesis does not explain why this is the case, since the predicted problems occur downstream of lexical selection. It also does not account for impairments on one-word tasks where no, or at least minimal, frame insertion is required.

*The Retention Hypothesis.* Freedman et al., (2004) suggest that the lexical-semantic memory buffer used in speech planning is the same as that used in sentence comprehension



and word list recall. They suggest that context-specific naming deficits may occur as a result of an inability to retain information in this memory buffer rather than as a result of damage to lexical processes, such as spreading activation and lexical selection.

As was discussed on p. 7, Freedman et al. (2004) present data from a picture-naming task where one picture is presented (minimal memory load), or two related or unrelated pictures are presented together (higher memory load). They found that their participants showed a significant single picture response time advantage, and argue that the increased time taken to initiate two-feature utterances supports the interpretation that both items in the utterance were planned before speech was initiated: This implies co-activation of the relevant nodes, and a need to retain information online. Further, they suggest that the utterance planning process should take longer when co-activated items are closely related semantically, due to increased interference between the items, and that patients with semantic STM impairments should be particularly susceptible to this interference because their damaged semantic memory buffers should have more difficulty retaining two highly related interfering items than an intact buffer. A pattern that is consistent with this hypothesis came through in their data: Their control group showed a significant semantic blocking effect, and two non-fluent aphasics with semantic STM deficits (ML and GR) showed a much larger difference than the controls.

Freedman et al. (2004) suggest that due to their impaired semantic buffers, ML and GR found it difficult to retain the required information online during production of multiword utterances. This retention hypothesis may be accurate, however it is possible that if patients have a deficit selecting between, rather than retaining, co-activated lexical items, they will do poorly on semantic STM tests which by their nature require co-activation of items. Hence, there is an element of circularity to the retention hypothesis. Another problem for the retention hypothesis is that it does not anticipate why patients would show deficits on single word naming tasks, such as blocked cyclic naming, which do not require online retention of information (see p. 4).

*Lexical Competition and the Selection and Control Hypothesis.* The selection and control hypothesis (Wilshire & McCarthy, 2002; Wilshire et al., 2006) and the syntactic hypothesis Schwartz and Hodgson (2002), were presented in response to apparently different context-specific naming deficits. The patient presented by Schwartz and Hodgson with the syntactic hypothesis, MP, did not show exaggerated sensitivity to paradigmatic competition. However, Wilshire and colleagues' patients, BM and JHM, did (Wilshire & McCarthy; Wilshire et al.). BM (discussed in more detail previously, see p. 5), showed a

context-specific naming deficit on blocked cyclic naming tasks, where he performed more poorly when items were presented in semantically related blocks than when they were not semantically grouped. JHM (discussed in more detail previously, see p. 8), showed a deficit in a picture-word interference task and a two-picture naming task, particularly when semantically related items were presented together.

Although the selection and control hypothesis and the syntactic hypothesis were presented in response to different deficits, somewhat confusingly, they both attribute a key role to activation competition between a word and any alternatives, and suggest that damage to the same theoretical process (controlled activation) leads to the different error profiles observed in the studies. Where the accounts differ is in the nature of the controlled activation failure. In the syntactic account, controlled activation fails because syntactic processing is impaired. In the selection and control account, controlled activation fails because a modulatory biasing process upstream is impaired (the boost of activation itself may be intact, but under competitive conditions, the bias process does not direct the boost to one option over the other options). This upstream modulatory dysfunction is an important feature of the selection and control hypothesis, because it accounts for the exaggerated sensitivity to paradigmatic competition observed in the performance of BM and JHM (Wilshire & McCarthy, 2002; Wilshire, et al., 2006), that was not initially observed in patient MP (Schwartz & Hodgson, 2002) but has been observed in patient MP more recently (Schnur et al., 2006).

Wilshire and colleagues (Wilshire & McCarthy, 2002; Wilshire et al., 2006) suggest that the deficits shown by BM and JHM are best located at the lexical level, because their comprehension is relatively good, and good comprehension reflects intact processing at the semantic level. However, they argue that a simple lexical deficit does not predict their sensitivity to “word external” factors, i.e., context (see Wilshire & McCarthy, p. 180). In locating BM’s context-specific naming deficit, using the MacKay (1997) theory of sequencing, Wilshire and McCarthy focus on two processes that occur within the model: Syntagmatic competition, and paradigmatic competition. Syntagmatic competition occurs between nodes that are planned for the same utterance (e.g., the nouns *name* and *mind* would be syntagmatic competitors in the utterance, “Your name came to mind”). Because BM performs relatively well when to-be-named items are not semantically grouped, Wilshire and McCarthy suggest that his ability to resolve syntagmatic competition is relatively good, and that he is able to make lexical selections and construct syntactic frames. Paradigmatic competition occurs between nodes that are co-activated as a direct

result of their relatedness and the spreading activation process. Wilshire and McCarthy suggest that increased levels of paradigmatic competition are associated with BM's deficit.

Wilshire and McCarthy (2002) note that naming a series of semantically related items will increase levels of lexical activation within the semantic group, more so than naming a series of unrelated items. This is because with semantic blocking, each subsequent target is activated by nature of it being the target, and also as a result of being semantically related to the other items in the set, which activate each other because they are all related. As the set size increases, or with set repetition, each item within the semantic group becomes more highly activated. In fact, because of spreading activation, the target items will be activated to some level even before the relevant pictures are presented. In contrast, when the items are presented in semantically unrelated groups, the target will be activated in the normal fashion, but the spreading activation will be distributed further across the lexicon and will not escalate to such high levels because the different targets in the group will not be activating each other, rather they will be activating other non-targets with which they share a semantic relationship, which will in turn activate other non-targets. While there may be some increase in activation levels because of the repetition of targets during cyclic naming tasks, it will not rise to the same extent as when the targets share a semantic relationship and hence repeatedly activate each other, regardless of whether they are the current target or not. In other words, semantic grouping leads to increased levels of paradigmatic competition.

Consider the possibility that a high level of paradigmatic competition triggers BM's single word production deficit. What process resolves high levels of paradigmatic competition allowing for lexical selection? In Dell (1986) and MacKay (1987), syntactic functions fire additional activation towards syntactically appropriate lexical items, so that the correct nodes are selected in the correct order. Wilshire and McCarthy (2002) call this process "controlled activation" and suggest that it, not spreading activation, is used to resolve high levels of paradigmatic competition. To illustrate this, consider the Dell model of sentence production. First a semantically specified intention to speak is formed. Activation from semantic features flows down from the semantic level to corresponding lemmas at the lexical level where generative rules use the insertion labels corresponding to the most highly activated lemmas to create syntactic frames. Up until this point, all activation within the lexicon is the result of spreading activation, which causes activation to constantly and automatically reverberate, in an unconstrained fashion, throughout the lexicon. At this point though, the most activated lemma tagged with a corresponding

insertion label receives an extra boost of controlled activation from the syntactic frame, to ensure that its activation level supersedes that of other activated nodes to a level that allows it to be selected for production. The critical point to recognise here is that even if a lemma's excitation level is the highest at a given moment, the difference between its excitation level and the excitation levels of other lemmas may not be large enough for it to reach the crucial threshold without the additional controlled activation, even though its excitation level may have been high enough for it to generate a correct syntactic frame.

Also, consider the MacKay (1987) theory of sequencing. This theory includes a similar mechanism, although it is expressed quite differently. Here content nodes spread additive excitation to related content nodes, and to related sequence nodes. Active sequence nodes inhibit other sequence nodes via inhibitory connections that represent serial order rules (i.e., in English the sequence node 'adjective' inhibits 'noun' because adjectives always precede nouns, e.g., red (adjective) square (noun)). When sequence nodes, which function as the syntactic planner, are selected, they send multiplicative activation back to related content nodes. This multiplicative activation increases the level of activation in the content nodes with the relevant syntactic function to the critical level, allowing for response selection. In this model, the additive spreading activation from content nodes is the uncontrolled process, and the multiplicative activation from sequence nodes is the controlled process.

Wilshire and McCarthy (2002) surmise that when paradigmatic competition is unusually high, controlled activation may be more important than it is in situations where there is one highly dominant response. They suggest that in addition to the functions set out in current language models, the controlled activation process may be sensitive to lexicon-external factors, resulting in it being able to bias controlled activation towards one or the other option, depending on conceptual or pragmatic information say. Wilshire and McCarthy suggest that BM's problem occurs during controlled activation, with the result that he finds it difficult to resolve high levels of lexical competition. In this view, it is not strictly the lexicon that is impaired. Rather, BM is not able to bias processing towards one word, hindering lexical selection.

In summary, Wilshire and colleagues suggest that context specific deficits may occur when there is more than one highly activated item that could be used in an utterance, because controlled lexical selection processes are impaired (see Wilshire & McCarthy, 2002; Wilshire, et al., 2006). This account predicts that when there is minimal competition, such as in single word naming tasks, word production will be less impaired.

When more than one word is activated, such as when two words or a sentence is being produced, performance will be more impaired. In contrast, the syntactic account (Schwartz & Hodgson, 2002) does not account for exaggerated sensitivity to paradigmatic competition, because the context-specific naming impairment is explained by a failure downstream of lexical selection, where controlled activation fails and hence does not boost the relevant phonological nodes. Hence, although the syntactic account is possibly more parsimonious than the selection and control account, the selection and control account may more accurately account for the data patterns that have been observed in patients who show context-specific single word production deficits.

*Summary and Comment on Accounts of Context Effects.* In summary, three contrasting accounts have been suggested to explain context-specific naming deficits. The syntactic hypothesis suggests that lexical selection proceeds as usual, except the syntactic frame fails to send a jolt of controlled activation down to the phonological level, resulting in increased competition that is not sensitive to paradigmatic competition. The retention hypothesis also suggests that lexical selection is fine, but suggests that impairments occur because patients are not able to retain the information online for long enough to produce the planned utterance. The selection and control hypothesis suggests that when a high level of lexical co-activation occurs, the failure of a modulatory process that directs activation towards the best option leads to a failure in lexical selection. This hypothesis differs from the other two because it predicts that lexical selection should fail more when paradigmatic competition is high<sup>2</sup>.

The alternative accounts of context specific language production deficits described above come from a body of literature that is relatively new. To date, only a limited number of paradigms have been used to investigate this sort of deficit, and within the limited number of studies that have been completed there are individual differences between the participants tested, and differences in the experimental paradigms that have been used. These differences make drawing conclusions from looking across the literature highly tenuous. The small body of literature available on the subject also makes theoretically based conjecture speculative. It is unsurprising then, that as yet no clear consensus has

---

<sup>2</sup> It is worth noting here that some language impairments may be the result of STM or syntactic deficits, and other language impairments may result from deficits selection and/or control. However, the current research focuses on impairments that are observed during single word production tasks that minimise demands on STM and syntax. It is unlikely that deficits observed on single word production tasks are explained by syntactic or STM accounts.

been achieved as to the cause, or causes, of context-specific language production deficits. The impact that improved knowledge on this subject may have on theories of normal language production also remains unclear. Further research on this subject is required before many unanswered questions can be addressed.

One question that clearly emerges from the literature asks how language production is controlled, and how control processes relate to other processes occurring within, and outside, the lexicon. With the selection and control hypothesis, it is suggested that conceptual or pragmatic information from outside the lexicon possibly biases processing of information within the lexicon – which as an independent entity, is portrayed as being reliant on learned information and sensory inputs (see Wilshire & McCarthy, 2002). It is suggested that such modulatory processes may be particularly beneficial in situations where there are alternative ways to express an idea. In addition, it is suggested that these biasing processes may be impaired in individuals with context-sensitive naming deficits, leading to particular difficulties in situations where more than one lexical item is highly activated. Of course this question brings with it a host of related questions, for instance: Is the idea of lexicon-external factors causing context-specific language deficits parsimonious? And how would such a process work?

If a lexicon external system modulates the selection of words in speech production, as is suggested in the selection and control hypothesis, it is possible that this system may share functional similarities with other systems that are believed to play an important role in cognitive control more generally, such as those located in prefrontal cortex. If this is the case, it might be useful to apply theories of frontal lobe function to the study of language. In the next section, several of the most well established empirical examples of how the frontal lobes function and break down are briefly described. Then three theories of frontal lobe function are discussed, with a particular focus on the issue of frontal modulation, because of this theme's relevance to the current study.

### *Insights from Studies of Frontal Lobe Function*

Themes of response competition and selection are frequently addressed in studies of frontal lobe function, as patients with frontal lobe lesions frequently perform poorly on tasks that manipulate variables of this type. The following discussion focuses first on some classic empirical observations regarding frontal lobe function and the themes of selection and control. Second, some different approaches that have been taken to modelling frontal

lobe function will be introduced. Finally, three models of frontal lobe function will be summarised and the possible contributions of these models to the study of language will be discussed.

*Empirical Observations relating to Frontal Lobe Function.* A classic example of a task that is sensitive to frontal lobe dysfunction is the Stroop (1935) task (this task was described earlier; see p. 8). Frontal patients are disproportionately impaired on this task showing exaggerated response costs when incongruent stimuli are presented (see Kimberg & Farah, 1993; Perret, 1974; Stuss & Benson, 1984). This suggests that frontal patients find it particularly difficult to override competing stimulus features, and that they may have more difficulty controlling their responses than neurologically intact individuals.

Another task that is associated with frontal lobe dysfunction is the Wisconsin Card Sorting Test (Grant & Berg, 1948). In this task, participants are asked to sort cards that differ along three dimensions: the number of objects on the card, the colour of the objects, and the shape of the objects. Initially participants are required to sort the cards by colour, and are given positive feedback when they sort the cards correctly. After they sort ten consecutive cards correctly the sorting rule is changed without warning, and they are required to sort along a different dimension (e.g., they are required to sort by shape instead of colour). Importantly, the participants are not explicitly informed of the change to the sorting rule, rather the change is indicated by a change in feedback, and participants have to initiate the required modification to their response strategy. Frontal patients make an unusually high number of errors on this task as compared to patients with lesions elsewhere in the brain. This is mainly because they take a long time to cease responding according to the established rule, and to switch to another rule (Milner, 1963; 1964). This suggests that frontal patients find it particularly difficult to override highly activated pre-potent response schemas, in order to complete correct, but less established responses. They seem more reliant on past behaviours, and less able to control responding to suit changing situations.

*Introduction to Frontal Lobe Theory.* The over-arching theme addressed in theories on frontal lobe functioning concerns executive control. Executive control is a general term used to describe a range of high level cognitive tasks, including planning, abstract thinking, cognitive flexibility, rule acquisition, and attention (see Banich, 2004). In many theories of frontal cortex function, executive control is formulated as modulation that occurs as a result of bias. The general idea is that during situations where the best response is not immediately obvious, frontal processes bias processing along non-frontal neural pathways that represent potential responses. This bias favours a single response, allowing it to be

activated to a higher level than competing responses, and hence allowing for response selection (Barcelo, Suwazono, & Knight, 2000; Cohen & Servan-Schreiber, 1992; Desimone & Duncan, 1995; Kimberg & Farah, 1993, 2000; Miller & Cohen, 2001; Shallice & Burgess, 1996; Weidner, Pollmann, Müller, & von Cramon, 2002).

Three theories of frontal cortex will be summarised here. They are: 1) the Shallice and Burgess (1996) theory of prefrontal cortex function; 2) the Miller and Cohen (2001) integrative theory of prefrontal cortex function; and 3) the Botvinick, Braver, Barch, Carter, and Cohen (2001) conflict monitoring hypothesis. Then at the end of the section, we consider how these theories might be linked to models of language production.

*The Theory of Prefrontal Cortex Function (Shallice & Burgess, 1996)*

In *the theory of prefrontal cortex function*, by Shallice and Burgess (1996), a range of frontal processes converges to perform one function: the modulation of schemas that govern behaviour according to task demands. The theory distinguishes between two major kinds of behaviours: routine behaviours that are regularly invoked every time a certain set of conditions arises; and novel behaviours, where a routine response must be suppressed and replaced with an alternative behaviour, or where no pre-existing routine behaviour is available.

In routine situations, potential responses become activated in response to stimulus conditions. The potential responses compete for selection in what is known as the contention scheduling system. Because a precedent response has been experienced, the non-frontal contention scheduling system is able to automatically select the action schema that has been associated with the situation in the past as it will be the most highly activated.

In novel situations, a routine response must be suppressed and replaced with an alternative behaviour, or no pre-existing routine behaviour is available so a new behaviour must be activated. In these situations a supervisory system uses frontal processes that include, but are not limited to, problem orientation, goal setting, episodic memory retrieval, and problem solving. These processes create a response plan or strategy that is called a temporary new schema. The new schema then biases processing towards the corresponding behaviour, which is stored in the lower level contention scheduling system, and the behaviour is produced. For example, in the Stroop task, when the relatively prepotent response of reading the word is replaced with the less practiced response of naming the colour of the ink, the supervisory system would step in and bias processing towards a colour naming response so that participants would be less likely to read the word in error. In completely novel situations, episodic memory would be searched for a



potential response, and then a checking process would be completed to see if this response was adequate, or whether the process should be completed again.

*The Integrative Theory of Prefrontal Cortex (Miller & Cohen, 2001)*

In an idea that is similar to that presented by Shallice and Burgess (1996) in the theory of prefrontal cortex function, Miller and Cohen (2001) assume that behaviour requires the activation of specific relevant neural pathways, and that these activations automatically spread through multiple pathways simultaneously following activation from a particular stimulus or context. Once the most highly activated pathway is established, it is selected and the corresponding behaviour is initiated while activations along competing pathways are inhibited. However, Miller and Cohen's theory is different to Shallice and Burgess' in that the prefrontal cortex is *constantly* modulating the pattern of activation generated by the stimuli or contexts, by sending top-down (internally driven) bias signals through chosen neural pathways. However, the bias is only important in situations where routine responses do not lead to a good outcome. In a situation where there is no clear response, stimulus driven automatic activations spread through multiple pathways but no one pathway has a big enough advantage to generate its own selection: Because no one pathway has the required competitive advantage, the bias signal from prefrontal cortex that directs activation towards one pathway ensures that that pathway is activated to a higher level than the others, and provides the advantage required to produce a response. Under conditions where an automatic response is clear, the bias signal from the prefrontal cortex is still sent but has little effect, because the automatic activations already clearly correspond to a single response.

More specifically, the prefrontal cortex operates via the active maintenance of neural activity corresponding to 'internal representations of goals', which guide the modulation process. The 'internal representations of goals' bias competitive processing by guiding excitatory action potentials down neural pathways that link stimulus inputs and internal states to output representations. This neural activity provides a boost to stimulus driven neural activation along the chosen pathways, making it more likely for the chosen pathway to exceed the selection threshold. For example, in the Stroop task, the prepotent response of reading the word would trigger an excitatory action potential from the prefrontal cortex towards a word reading output representation. This bias signal would have minimal behavioural ramifications because even if it was not sent, the most dominant automatic response would presumably be to read the word. However when the word reading response is replaced with the less practiced response of naming the colour of the

ink, the frontal bias would trigger an excitatory action potential towards a colour naming output representation, allowing the prepotent reading response to be over-ridden. In this case, the behavioural ramifications of the bias signal are important, because the end result is a different behaviour.

*The Conflict Monitoring Hypothesis (Botvinick et al. 2001)*

Most theories on frontal lobe function focus on the influence that control/modulation processes have on behaviour. In contrast, the *conflict monitoring hypothesis* focuses on *how* cognitive control might operate (Botvinick et al. 2001). These researchers suggest that conflict monitoring is the first phase of cognitive control. They propose that there is a dedicated *conflict monitoring module* in the anterior cingulate cortex that measures the total amount of current activation (which is called *energy*) present among possible response representations. They suggest that this energy measure reflects levels of response conflict, because the more competing units there are, or the higher activation is within competing units, the higher overall energy levels become. In the conflict monitoring hypothesis, it is proposed that the energy measure is forwarded to a control centre, and if required, this triggers the control centre to increase its influence on task processing, allowing for stronger top-down control, which in turn leads to the selection of a response.

To illustrate the conflict monitoring and feedback processes, consider the Stroop Task (see Cohen & Huston, 1994; Stroop, 1935). When the prepotent word reading response is required, the word reading response representation will be highly active (for illustrative purposes, consider that it is active to 50 units), but activation of the unusual colour naming response representation will be minimal (say 5 units). Across both response representations, overall energy levels would be pretty low (55 units). However when the unusual colour naming response is required, the prepotent word reading response representation will be highly active (say 50 units), and activation of the unusual colour naming response representation is also high (say 50 units). Across both response representations, overall energy levels would be relatively high, and send a signal to the control centre that modulation is required. Next, the frontal control centre would send an excitatory signal to the colour naming units, raising their activation level (to say 60 units), and would send an inhibitory message towards the word reading units, decreasing their activation level (to say 15 units). These activation changes would increase the difference in activation levels between the correct and incorrect responses, allowing a critical threshold to be reached, and a response to be produced.

Overall, the conflict monitoring hypothesis appears to complement Shallice and Burgess's (1996) idea that the frontal bias process only operates when automatic responses fail to select the best response, by providing a framework describing how this process may operate.

*Summary and Comment.* In summary, these three models of frontal lobe function posit that modulation is a key function of the frontal lobes, and that modulation is important because it allows for behaviour to be directed towards the most appropriate response in situations where one dominant response is not able to be automatically generated. Hence, the themes that are addressed in the frontal lobe theories correspond with those that Wilshire and colleagues raise in relation to the selection and control hypothesis (see Wilshire & McCarthy, 2002; Wilshire et al. 2006). Potentially, the “lexicon external” system that Wilshire and McCarthy suggest may modulate the selection of words in speech production may share functional similarities with other frontal modulatory systems, such as those located in prefrontal cortex. If so, the theories that have been put forward regarding how frontal modulation operates may explain how Wilshire and colleagues' largely unspecified “lexicon external” factor may influence word selection. Further, an impairment to the hypothesised frontal modulation of linguistic information could provide a theoretically motivated explanation for why some aphasic patients show context-specific naming deficits.

### *Introduction to the Current Research*

The aim of the current research is to explore an extended version of the Wilshire and McCarthy (2002) selection and control hypothesis as a potential explanation for context-specific effects in individuals with non-fluent forms of aphasia. The selection and control hypothesis predicts that during situations where there is more than one possible linguistic response, a modulation process biases activation towards the most appropriate response, allowing the best response to be translated into action. If the selection and control process is impaired as a result of brain damage, deficits in word production should occur during situations where there is more than one highly activated potential response. On the other hand, if one response is automatically activated to a higher level than any competitors, the impairment will minimally affect production. Such situations might occur in tasks such as picture naming where there is one single dominant name associated with each picture (see also Wilshire & McCarthy, 2002; Wilshire et al., 2006). In the current

research we are going to complete a range of tasks that manipulate activation levels of alternative responses (lexical competition).

Although there are some exceptions to the following statement, in general the studies included in the current project will be analysed using mixed-effect models, which are advocated by Brysbaert (2007), and Baayen, Davidson, and Bates (2008). In this approach, each data point is entered into a statistical analysis, and participant and item are entered as repeated factors (this can be contrasted to ANOVA, where the data is collapsed across items). An important advantage of using mixed-effect models is that researchers have an increased ability to address different factors that may be influencing task performance (Baayen, et al., 2008).

As discussed previously, past evidence suggests that some non-fluent aphasics, particularly those with good single object naming and poor sentence production, are unusually susceptible to contextual factors, even in single word production tasks such as cyclic blocked naming (see Jefferies et al., 2007; Schnur et al., 2006; Wilshire & McCarthy, 2002). However, the limited research that has been completed on unimpaired individuals shows that they are also influenced by contextual factors, although the size of their context effects appear much smaller than those observed in the non-fluent aphasics (Schnur et al., 2006; Freedman et al., 2004). Hence, the prevalence of context-specific effects in language production remains unclear, not only in aphasics, but also in normal speakers.

Perhaps because of the exploratory nature of past research, most studies on lexical competition and aphasia have focused on carefully selected non-fluent aphasics with good single word naming and poor sentence production. This case study approach makes replication of past research difficult, as any attempts at replication are complicated by individual differences between patients. Since it is not always possible to replicate patient data, the current research aims to examine the performance of a range of individuals across a wide range of tasks, thereby enabling us to examine both similarities and differences across individuals and tasks. Because an important aim of this research is to test a range of individuals, both non-fluent and fluent aphasics, as well as older neurologically intact controls, will complete the key tasks.

To investigate whether context-specific language production impairments are particularly sensitive to different types of lexical competition, the current research aims to develop paradigms that primarily manipulate one of three types of competition: First,

paradigmatic competition; Second, competition resulting from underdetermined responses, and; Third, extrinsic competition.

Because paradigmatic competition occurs due to overlapping high-level semantic representations, we will include two tasks where co-activation of semantically related information is manipulated. The first is a picture-naming task in which the number of alternative names for the pictured objects varies (the name agreement task, see pp. 110-128). In the high competition condition, several alternative names are available (e.g., sofa / couch). In the low competition condition, one dominant name is used (e.g., anchor). Hence, the task should manipulate overall levels of competition from semantically related alternative names by manipulating the number of competitors available. The second is a category exemplar task (this is a category based verbal fluency task) (see also: Randolph, Braun, Goldberg & Chase 1993; Hanes, Andrewes, & Pantelis, 1995). In a high competition condition, participants are given a category and are asked to produce as many exemplars that belong to that category as possible (e.g., animals). In a low competition condition, participants are also given sub-categories that limit the number of potential responses and reduce overall competition levels (e.g., pets, water animals, farm animals, jungle animals). By dividing the categories into smaller sub-categories, the number of potential competitors is limited, and should reduce amount of competition present.

The second type of competition investigated arises when there is more than one way to respond to a situation, causing co-activation of *underdetermined* responses. To manipulate competition that arises because of underdetermined responses, we will use a verb generation task following the neuroimaging and lesion studies by Thompson-Schill and colleagues (see Thompson-Schill et al., 1997; Thompson-Schill et al., 1998). In this task, participants are shown pictures of objects, and are required to produce verbs that express what the objects do, or what people do with them. In a high competition condition, there are many associated verbs (e.g., *penny* – spend / drops / pay / buy, etc.). In the low competition condition, there is one dominant response (e.g., *chair* – sit). The current study aims to replicate past findings with new stimuli, to extend past participant pools, and to focus on the aphasic deficits that may be associated with any findings.

The third source of competition we aim to investigate comes from *extrinsic* sources, in other words, it occurs due to extraneous thought processes that are irrelevant to the current speech plan, or that are triggered by irrelevant external stimuli. Specifically, we will manipulate the strength of auditory distracter stimuli in a sentence completion task. In this task, participants will see and hear sentences that are missing the final word (e.g.,

*Barry wisely chose to pay the \_\_\_\_\_*). In the high competition condition, they will hear an example completion (e.g., *FINE*) and must choose another example. In the low competition condition, they will hear a syntactically appropriate but semantically unrelated word in place of the example completion (e.g., *RANGE*) and must choose another example. This task aims to manipulate levels of extrinsic competition by providing strong or weak auditory distracters that have to be disregarded in order to successfully complete the task.

It is hypothesised that all participants will perform better in low competition conditions than high competition conditions on all four tasks. However, it is expected that non-fluent aphasics will be influenced more consistently, and to a greater extent, than the other participants, reflecting impairments to selection and control processes. Also, based on findings that suggest that the sub-group of non-fluent aphasics who show relatively good naming but poor sentence production are particularly sensitive to manipulations of context in naming tasks (see Freedman et al., 2004; Schwartz & Hodgson, 2002; Schnur et al., 2006; Wilshire & McCarthy, 2002; Wilshire et al., 2006), we hypothesise that non-fluent aphasics who fit this profile may show particularly strong high vs. low competition condition effects on the current tasks.

In summary, the research that has been discussed in this introduction offers some support for the idea that a modulation process that operates beyond the scope of current language models may function to control language production, especially in situations where there is not one pre-specified dominant response. However, the explanations offered for the phenomena that have been observed have tended to be data driven, rather than hypothesis driven. Also, researchers have used different paradigms to investigate similar processes, which makes cross-task comparisons difficult. Further, most (although not all) of the published research has been case studies, which make it difficult to assess the precise characteristics associated with context-specific language effects.

The current research aims to:

- a) Test the selection and control hypothesis on four hypothesis driven tasks that manipulate levels of lexical competition.
- b) Include a wide range of participants including non-fluent aphasics, fluent aphasics, and matched controls.
- c) Analyse participant data on an individual basis, to avoid hiding significant patterns, but to also look for consistencies between the results of participants with similar deficit profiles.



### Category Exemplar Task Introduction

Verbal fluency measures are commonly used to examine executive functioning and word retrieval in neuropsychology. Verbal fluency measures include exemplar tasks in which participants are given a constraint on retrieval and a limited time period, to provide as many examples that fit within the given constraint as possible. For instance, in phonemic fluency tasks, participants are asked to produce as many words as possible that start with a certain letter (e.g., *F*: “Fast. Frequent. Food”, etc.). In category fluency tasks, participants are asked to produce as many examples of things that fit within a certain semantic category as possible (e.g., *Animals*: “Cat. Dog. Mouse”, etc.) (e.g., Spreen & Strauss, 1991).

Randolph et al. (1993) present a variation of a category fluency task that may prove useful in tests of the selection and control hypothesis in aphasics. In this variation, the size of the retrieval categories is manipulated using broad categories and narrow sub-categories. The broad retrieval condition is the equivalent of a standard category fluency task: Participants are given a broad category name and are asked to provide as many exemplars as possible in 60 seconds (e.g., *Animals (60 sec.)*: “Cat. Dog. Mouse”, etc.). In the narrow retrieval condition, participants are given the broad category name, then they are sequentially presented with four sub-category names and are given 15 seconds to produce exemplars from each sub-category (e.g., *Animals*: then *Animals that people keep in their home as pets (15 sec.)*: “Cat. Dog. Fish” etc., then *Animals that are found on a farm (15 sec.)*: “Sheep. Cow. Goat” etc., then *Animals that live in the jungle (15 sec.)*: “Snakes. Spiders. Gorillas” etc., then *Animals that live in water (15 sec.)*: “Whales. Badgers. Crocodiles” etc.) (p. 4). The broad vs. narrow category manipulation may be useful in our investigation of the selection and control hypothesis because in the broad condition there is a large pool of items to produce as exemplars, hence there is potential for high levels of lexical competition. In the narrow condition, the tighter constraints placed on the items that can be produced at any one time limit the total number of potential competitors to smaller groups of items, and the amount of potential lexical competition.

Many verbal fluency studies have indicated that participants apply strategies during item generation. For example, participants tend to produce items from a given category in semantically related clusters (e.g., *Items you find in a Supermarket*: “Milk. Cream. Cheese. Butter. Margarine. Bread. Muffins” etc.) Further, several papers have reported that frontal lobe damage/dysfunction interferes with such strategy use. For example,



Troyer, Moscovitch, Winocur, Alexander, and Stuss (1998) not only counted the number of exemplars produced in phonemic and category fluency tasks, but they also counted the size of the clusters produced (by counting the number of items belonging to various sub-categories that were produced successively), and the number of switches, or transitions, that were made between clusters. They observed that frontal lobe patients switched between clusters less frequently than controls, although they did not produce clusters that were significantly different in size (see also Donovan, Siegert, McDowell, and Abernethy (1999) who observed similar results with non-dementing Parkinson's disease patients). Consistent with the hypothesis that frontal patients may show deficits in strategy generation, Randolph et al. (1993) found that when sub-categories were given to provide externally rather than internally derived switches, a group of Parkinson's disease patients, a group of Huntington's disease patients, and a single patient with a frontal lesion benefited, whereas a control group did not.

The primary interest of the current research is to find out how non-fluent aphasics respond to different levels of lexical competition. In the current study, which follows that by Randolph et al. (1993), participants are asked to produce items that belong to one broad category (the high competition condition), or four smaller sub-categories (the low competition condition). According to the selection and control hypothesis, the high competition condition should induce the most competition, because more items fit within the broad categories resulting in more potential activation. This increase in activation is expected to result in a performance cost for all three participant groups: non-fluent aphasics, fluent aphasics, and older controls.

Further, it is hypothesised that because some or all non-fluent aphasics may suffer from a specific impairment to a lexical selection and control mechanism, they may show exaggerated high vs. low competition condition effects relative to the fluent aphasics and older controls: This is expected to be reflected in a group by competition condition interaction, where the non-fluents show a greater decrease in the number of exemplars produced in the high competition condition relative to the low competition condition, than the fluent aphasics and controls.

Finally, based on findings that a sub-group of non-fluent aphasics, who show good naming but poor sentence production, are particularly sensitive to manipulations of context in naming tasks (see Freedman et al., 2004; Schnur et al., 2006; Schwartz & Hodgson, 2002; Wilshire & McCarthy, 2002; Wilshire et al., 2006), we hypothesise that non-fluent aphasics who fit this profile may show particularly strong high vs. low competition

condition effects. Three non-fluents in the current sample have relatively good single word naming, scoring above 40 on the BNT, and are expected to show these exaggerated effects (see p. 64 for more details).

In the current study, we used the Randolph et al. (1993) task design as a base, but wanted to include more categories/sub-categories, and to carefully balance the number of items produced in the different conditions. Hence, the first step taken was to complete a pilot task that included an increased selection of categories/sub-categories.

### Category Exemplar Pilot

#### *Method*

##### *Participants*

Fifty first-year psychology students at Victoria University of Wellington participated in this study for course credit. All participants spoke English as their first language. Informed consent was obtained from all participants.

##### *Materials*

Ten semantic categories were selected, each of which could be further subdivided into four sub-categories that contained a large number of exemplars. Two of the categories (*Animals* and *Supermarket Items*) were taken directly from Randolph et al. (1993). The remaining eight categories were novel (see Table 2.1 for a list of the categories and sub-categories included).

Table 2.1

*Category Exemplar Pilot: Semantic Categories and Sub-categories Included*

<i>Categories</i>	<i>Sub-categories</i>			
Animals	Farm	Jungle	Pets	Water
Clothing	Footwear	Ladies'	Men's	Seasonal
Countries	African	Asian	European	Pacific
Fun Activities	Arts and Crafts	Games	Hobbies	Sports
Home Objects	Bathroom	Garage	Kitchen	Living Room
Musical Instruments	Brass	Percussion	String	Woodwind
Names	Boys'	Girls'	Surnames	Place Names
Occupations	Health and Care	Hospitality	Office Jobs	Trades
Supermarket Items	Cleaning	Drinks	Fruit and Vegetables	Meat and Seafood
Tools and Equipment	Building	Garden	Household	Office

*Design*

There were two experimental conditions: high and low competition. In the high competition condition participants were given the name of a broad category (e.g., *Animals*), and had one minute to write down as many items as possible that belonged in that category. In the low competition condition participants were first informed of the broad category (e.g., *Animals*), then they were given the name of one of the sub-categories (e.g., *Farm Animals*). They had 15 seconds to write down as many items as possible from that sub-category. Following completion of the first sub-category, the remaining three sub-categories were completed in the same way.

Each participant completed each of the ten categories, half in the high competition condition and half in the low competition condition. The high and low competition conditions were presented in blocks. There were four versions of the experiment. In versions one and three, the high competition condition was completed first and the low competition condition second. In versions two and four, the low competition condition was completed first and the high competition condition second. In versions one and two, the categories were presented in the following order: *Animals*, *Supermarket Items*, *Names*, *Fun Activities*, *Countries*, *Home Objects*, *Musical Instruments*, *Clothing*, *Occupations*, and *Tools and Equipment*. In versions three and four, the categories were presented in the reverse order. Each version was presented as an answer booklet that contained each category on a separate page.

### *Procedure*

After a brief introduction containing general information regarding the nature of the study, the participants were led through the task instructions, which were presented in a written format and read out loud by the experimenter. Separate instruction sheets were used for the high and low competition conditions (see Appendix A1 for the task instructions). The participants were given the opportunity to ask questions before the experiment progressed to the experimental trials. Commencing on the first page of the answer booklet, the participants were instructed to read the category name, which was also read aloud by the experimenter. During the low competition condition trials, the experimenter would read aloud the sub-category names at 15 second intervals. Participants were instructed when to move onto the next page in the answer booklet by the experimenter, who timed the exercise with a stopwatch. After the experimental session, a debriefing session was completed with the participants for its pedagogical value.

### *Results*

The number of correct answers given by each participant was counted for each category/sub-category. Table 2.2 shows the minimum, maximum, mean, and total number of items generated for each category and condition.

Table 2.2

*Category Exemplar Pilot: Minimum, Maximum, Mean, and Total Number of Items Generated by Category and Competition Condition*

<i>Condition</i>	<i>Category</i>	<i>Min.</i>	<i>Max.</i>	<i>Mean</i>	<i>Total</i>
High Competition	Animals	6	28	17	410
Low Competition	Farm Animals	3	7	5	132
	Jungle Animals	2	6	4	105
	Pets	4	8	6	147
	Water Animals	2	6	4	114
	<i>Total</i>				498
High Competition	Clothing	11	24	16	413
Low Competition	Footwear	3	6	4	106
	Ladies'	3	7	5	119
	Men's	1	7	5	110

Table 2.2

*Category Exemplar Pilot: Minimum, Maximum, Mean, and Total Number of Items Generated by Category and Competition Condition*

<i>Condition</i>	<i>Category</i>	<i>Min.</i>	<i>Max.</i>	<i>Mean</i>	<i>Total</i>
	Seasonal	2	5	4	88
	<i>Total</i>				423
High Competition	Countries	6	25	17	410
Low Competition	African	0	6	3	66
	Asian	0	9	4	106
	European	0	8	4	116
	Pacific	0	6	4	96
	<i>Total</i>				384
High Competition	Fun Activities	7	18	13	319
Low Competition	Arts and Crafts	2	5	3	83
	Games	0	5	3	81
	Hobbies	1	6	3	85
	Sports	3	6	5	132
	<i>Total</i>				381
High Competition	Home Objects	11	26	17	432
Low Competition	Bathroom	3	7	5	127
	Garage	2	6	4	97
	Kitchen	2	8	6	134
	Living Room	4	7	5	128
	<i>Total</i>				486
	Musical				
High Competition	Instruments	8	17	12	312
Low Competition	Brass	1	5	3	69
	Percussion	0	6	3	64
	String	2	7	4	104
	Woodwind	0	4	2	56
	<i>Total</i>				293
High Competition	Names	9	30	21	495
Low Competition	Boys Names	3	10	6	162
	Girls Names	4	9	6	168
	Place Names	0	7	4	95
	Surnames	2	7	5	117
	<i>Total</i>				542
High Competition	Occupations	8	17	12	310
Low Competition	Health and Care	3	5	4	94
	Hospitality	2	5	4	84
	Office	2	5	4	86
	Trades	2	5	4	86
	<i>Total</i>				350
High Competition	Supermarket Items	9	23	17	407

Table 2.2

*Category Exemplar Pilot: Minimum, Maximum, Mean, and Total Number of Items Generated by Category and Competition Condition*

<i>Condition</i>	<i>Category</i>	<i>Min.</i>	<i>Max.</i>	<i>Mean</i>	<i>Total</i>
Low Competition	Cleaning				
	Products	2	5	3	88
	Drinks	4	8	5	132
	Fruit and Vegetables	3	9	5	139
	Meat and Seafood	3	8	5	139
	<i>Total</i>				498
High Competition	Tools and Equipment	7	17	11	288
Low Competition	Building	3	6	4	95
	Garden	2	5	4	85
	Household	1	6	3	81
	Office	2	7	5	110
	<i>Total</i>				371

The categories that best met the following criteria were selected for use in the category exemplar task proper: 1) A high number of exemplars were consistently elicited in both conditions; 2) There was a relatively even distribution of exemplars across sub-categories in the narrow condition; 3) All participants produced at least one example for each sub-category; 4) There was no overlap in the examples given for different subcategories; and 5) The exemplars given in the high competition condition included at least one item from each of the of the sub-categories used in the low competition condition.

The following six categories met the aforementioned conditions, and were selected for use in the category exemplar task proper: *Animals*, *Supermarket Items*, *Home Objects*, *Clothing*, *Occupations*, and *Tools and Equipment*. The ‘Names’ category was rejected because in the high competition condition, participants did not generate items from all of the sub-categories included in the low competition condition (condition 5 above). Specifically, in the high competition condition, participants never generated place names. Another problem was that one participant wrote down full names (e.g., Brad Pitt) when most participants wrote down first names only, presenting a difficulty in deciding how such responses should be scored. The *Countries* category was rejected because few African countries were generated, and some participants failed to produce any exemplars for this sub-category (condition 3). The *Fun Activities* category was rejected because there was considerable overlap in the examples given for the different sub-categories (condition 4).

For instance, sports are games, games are hobbies, and hobbies can be arts and crafts. Also, there was an uneven distribution of exemplars across sub-categories, with *Sports* generating more exemplars than the other categories (condition 2). Finally, the *Musical Instruments* category was rejected because of low generation rates overall (condition 1).

## Category Exemplar Task

### *Method*

#### *Participants*

Participants included: seven non-fluent aphasics (CT, DD, DA, ECV, JHM, TB, and WS); five fluent aphasics (FS, NS, NP, STR, and XX<sup>3</sup>); and thirteen neurologically intact older controls (AK, AP, AR, BK, EK, EM, GA, IH, KR, PP, RS, SR, and TK). All participants were recruited from existing participant pools at Victoria University of Wellington and Temple University, Philadelphia, and spoke English as their first language. Informed consent was obtained from all participants. The date of birth and gender of each participant are detailed in Table 2.3. Note that Table 2.3, and the remainder of this participant section is referred to in all four experimental tasks, and includes details of participants who did not participate in the category exemplar task if they did participate in one of the other tasks.

---

<sup>3</sup> An additional fluent aphasic, PS, did not participate in the category exemplar task because he was recruited after it was completed, however his details are included in this participant section, which is also referred to for the other three experimental tasks, in which he did participate.

Table 2.3  
*Participant Date of Birth and  
 Gender Information*

	Date of Birth	Gender
Non-Fluents		
CT	20-Mar-61	F
DA	25-Feb-41	M
DD	22-May-47	M
ECV	27-May-53	F
JHM	25-Jul-60	F
TB	1-Jul-68	F
WS	4-Sep-61	M
Fluents		
FS	18-Jan-53	F
NP	27-May-39	M
NS	25-Sep-17	M
PS	28-Jan-52	M
STR	2-Oct-30	F
XX	23-Sep-43	M
Controls		
AK	11-Apr-32	M
AP	13-Nov-28	F
AR	1-Jun-38	M
BK	15-Nov-37	F
EK	29-Apr-38	F
EM	30-May-33	F
GA	3-Oct-44	M
IH	27-May-36	M
KR	12-Aug-31	F
PP	11-Oct-34	F
RS	9-Feb-23	M
SR	27-May-45	F
TK	9-May-33	M

Each aphasic participant's speech comprehension, production, repetition, and reading abilities were thoroughly assessed using a selection of speech and language tests. A description of the tests that were used for diagnostic purposes is included in Appendix A2. The non-fluent participant's scores on these tests are included in Table 2.4, and the fluent participant's scores are included in Table 2.5 (scores that are below the normal range are in bold font). The groups' mean scores and standard deviations on the tests are depicted in Table 2.6, with t-test statistics indicating where significant differences between the groups were present. A sample of each aphasic's spontaneous speech production is included in Table 2.7. Each participant's results on the diagnostic tests are discussed in



detail in the individual case summaries included in Appendix A3, and are briefly summarised in the following case descriptions.

Table 2.4  
*Non-Fluent Aphasic Participants' Scores/Classifications on Diagnostic Tests*

Test	Control mean ( <i>2 sd cut-off</i> )	CT	DA	Participant					
				DD	ECV	JHM	TB	WS	
<i>BDAE</i>									
Diagnosis	n/a	Mixed Non- Fluent	Broca's	Broca's	Anomic	Broca's	Recovered Brocas	Broca's <sup>5</sup>	
BNT (N = 60)	52.5 ( <i>48.1</i> ) <sub>1</sub>	<b>30</b>	54/59 <sup>6</sup>	<b>19</b>	<b>48</b>	<b>45/59<sup>6</sup></b>	<b>38</b>	<b>14/59<sup>6</sup></b>	
<i>COWAT</i>									
Word Fluency (F, A and S: 3 mins)	15.9 (8.7) <sup>2</sup>	9	18	<b>0</b>	<b>6</b>	<b>5</b>	12	<b>0</b>	
Category Fluency: Animals (1 min)	16.8 (8.2) <sup>2</sup>	9	12	<b>7</b>	<b>8</b>	11	14	<b>2</b>	
Category Fluency: Fruit (1 min)	Not Available	7	6	5	9	7	11	5	
<i>Lukkien &amp; Wilshire Sentence Production Test</i>		Not Available							
Sentence Score (N = 30)		0	7	0	4	2	8	0	
Syntax Score (N = 30)		4	13	0	5	9	20	0	
Root Verb Score (N = 20)		5	14	12	12	15	13	6	
Function Word Score (N = 94)		59	71	43	76	59	80	13	
Noun Score (N = 35)		22	30	16	20	28	30	9	
<i>Lukkien &amp; Wilshire Single Noun Test (N = 35)</i>		28	27	21	25	30	31	17	
<i>PALPA</i>									
Non-Word Repetition: Subtest 8 (N = 30)	Not Available	10	4	8	17	25	27	4	

Table 2.4  
*Non-Fluent Aphasic Participants' Scores/Classifications on Diagnostic Tests*

Test	Control mean (2 <i>sd</i> <i>cut-off</i> )	CT	DA	DD	Participant		TB	WS
					ECV	JHM		
Auditory Word Repetition:								
Subtest 9								
High Imageability Words (N = 40)	39.3 (36.3)	<b>36</b>	<b>32</b>	<b>34</b>	36/39 <sup>7</sup>	40	39	<b>29</b>
Low Imageability Words (N = 40)	39.5 (37.1)	<b>31</b>	<b>25</b>	<b>21</b>	<b>32</b>	<b>35</b>	39	<b>18</b>
Imageability and Frequency								
Reading: Subtest 31								
High Imageability Words (N = 40)	39.9 (39.2)	<b>35</b>	<b>38</b>	<b>23</b>	<b>36</b>	<b>38</b>	<b>39</b>	<b>29</b>
Low Imageability Words (N = 40)	39.6 (38.16)	<b>18</b>	<b>36</b>	<b>7</b>	<b>13</b>	<b>30</b>	<b>35</b>	<b>13</b>
Spelling-Sound Regularity								
Reading: Subtest 35								
Regular Words (N = 30)	30.0 (29.6)	<b>16</b>	30	<b>8</b>	<b>11</b>	<b>28</b>	<b>27</b>	<b>16</b>
Exception Words (N = 30)	29.9 (29.1)	<b>18</b>	<b>28</b>	<b>10</b>	<b>18</b>	<b>29</b>	<b>25</b>	<b>17</b>
Non-Word Reading: Subtest 36 (N = 24)	22.9 (19.9)	<b>1</b>	<b>18</b>	<b>0</b>	<b>1</b>	<b>0/11</b> <sup>8</sup>	<b>0</b>	<b>2</b>
<i>PPVT Standard Score (III)</i>	100 (70)	71 (IIIB)	118 (IIIA)	85 (IIIB)	81 (IIIB)	87 (IIIA)	82 (IIIB)	74 (IIIA)
<i>PCB</i>								
Lexical comprehension		43						
Within Category (N = 16)	15.8 (15.2) <sub>3</sub>	Not Available	16	16	<b>13</b>	16	<b>14</b>	<b>13</b>
Between Category (N = 28)	27.7 (27.2)	Not	28	28	28	28	<b>27</b>	<b>25</b>

Table 2.4  
*Non-Fluent Aphasic Participants' Scores/Classifications on Diagnostic Tests*

Test	Control mean (2 <i>sd</i> <i>cut-off</i> )	CT	DA	DD	Participant			
					ECV	JHM	TB	WS
<hr/>								
	3	Available						
Sentence comprehension (Form A)								
Lexical distracters (N = 30)	29.7 (29.1) <sub>3</sub>	27	25	26	24	29	30	26
Reverse role distracters (N = 30)	28.8 (25.8) <sub>3</sub>	15	23	14	20	16	23	19
<i>Pyramids and Palm Trees Test: Three Pictures (N = 52)</i>	49.8 (46.5) <sub>4</sub>	48	50/51 <sup>7</sup>	48	46	51	49	47
<i>Zinseger &amp; Berndt Noun-Verb Naming Test</i>								
High Frequency Nouns (N = 30)	29.8 (28.9)	25	29	15	23	27	27	20
Low Frequency Nouns (N = 30)	29.8 (28.9)	29	26	21	25	30	30	20
Verbs (N = 30)	29.4 (27.6)	27	28	12	19	26	27	6

*Note.* Scores in **bold** font are below the normal range.

<sup>1</sup> Control data for New Zealanders from Barker-Collo (2001).

<sup>2</sup> Control data for elderly New Zealanders from Donovan, Siegert, McDowall & Abernethy (1999).

<sup>3</sup> Control data from Breedin and Saffran (1999).

<sup>4</sup> Control data from Hulleman and Humphreys (2007).

<sup>5</sup> WS scored outside the specified range repetition and auditory comprehension, but this profile most closely matches his performance.

<sup>6</sup> Item 19 (Pretzel) was not tested because it is not in common usage in NZ English.

<sup>7</sup> One item was not tested due to experimenter error.

<sup>8</sup> Testing was discontinued at item 11 at the participant's request.

Table 2.5  
*Fluent Aphasic Participants' Scores/Classifications on Diagnostic Tests*

Table 2.5  
*Fluent Aphasic Participants' Scores/Classifications on Diagnostic Tests*

Test	Control mean (2 <i>sd cut-off</i> )	Participant					
		FS	NP	NS	PS	STR	XX
Non-Word Repetition: Subtest 8 (N = 30)	Not Available	8	12	0	24	26	7
Auditory Word Repetition: Subtest 9							
High Imageability Words (N = 40)	39.3 (36.3)	<b>28</b>	<b>36</b>	<b>16</b>	39	<b>34</b>	<b>28</b>
Low Imageability Words (N = 40)	39.5 (37.1)	<b>14</b>	<b>32</b>	<b>7</b>	40	<b>29</b>	<b>17</b>
Imageability and Frequency Reading: Subtest 31							
High Imageability Words (N = 40)	39.9 (39.2)	<b>34</b>	<b>38</b>	<b>39</b>	40	<b>37</b>	<b>33</b>
Low Imageability Words (N = 40)	39.6 (38.16)	<b>26</b>	<b>37</b>	39	39	<b>32</b>	<b>16</b>
Spelling-Sound Regularity Reading: Subtest 35							
Regular Words (N = 30)	30.0 (29.6)	<b>13</b>	<b>29</b>	30	30	<b>28</b>	<b>22</b>
Exception Words (N = 30)	29.9 (29.1)	<b>11</b>	<b>29</b>	<b>26</b>	<b>29</b>	<b>28</b>	<b>14</b>
Non-Word Reading: Subtest 36 (N = 24)	22.9 (19.9)	<b>10</b>	<b>19</b>	21	<b>19</b>	21	<b>5</b>
<i>PPVT Standard Score</i>	100 (70)	<b>54 (IIIB)</b>	84	81	95	98	<b>57 (IIIB)</b>
<i>PCB</i>							
Lexical comprehension							

Table 2.5  
*Fluent Aphasic Participants' Scores/Classifications on Diagnostic Tests*

Test	Control mean (2 sd cut-off)	Participant					
		FS	NP	NS	PS	STR	XX
Within Category (N = 16)	15.8 (15.2) <sup>3</sup>	<b>15</b>	16	16	<b>15</b>	16	<b>15</b>
Between Category (N = 28)	27.7 (27.2) <sup>3</sup>	<b>27</b>	28	<b>19</b>	<b>27</b>	<b>27</b>	<b>26</b>
Sentence comprehension (Form A)							
Lexical distracters (N = 30)	29.7 (29.1) <sup>3</sup>	<b>26</b>	<b>29</b> <b>(Form B)</b>	<b>22 (Form B)</b>	30	<b>29</b>	<b>27</b>
Reverse role distracters (N = 30)	28.8 (25.8) <sup>3</sup>	<b>19</b>	30 (Form B)	<b>16 (Form B)</b>	27	<b>22</b>	<b>19</b>
<i>Pyramids and Palm Trees Test:</i> <i>Three Pictures (N = 52)</i> <i>Zingeser &amp; Berndt Noun-Verb Naming Test</i>	49.8 (46.5) <sup>4</sup>	<b>45</b>	50	51	51	52	49
High Frequency Nouns (N = 30)	29.8 (28.9)	<b>24</b>	<b>25</b>	<b>26</b>	<b>28</b>	<b>22</b>	<b>23</b>
Low Frequency Nouns (N = 30)	29.8 (28.9)	<b>21</b>	<b>22</b>	<b>24</b>	<b>28</b>	<b>19</b>	<b>26</b>
Verbs (N = 30)	29.4 (27.6)	<b>12</b>	<b>25</b>	<b>20</b>	<b>20</b>	<b>24</b>	<b>22</b>

*Note.* Scores in **bold** font are below the normal range.

<sup>1</sup> Control data for New Zealanders from Barker-Collo (2001).

<sup>2</sup> Control data for elderly New Zealanders from Donovan et al. (1999).

Table 2.5  
*Fluent Aphasic Participants' Scores/Classifications on Diagnostic Tests*

Test	Control mean (2 <i>sd cut-off</i> )	FS	NP	Participant			
				NS	PS	STR	XX

<sup>3</sup> Control data from Breedin and Saffran (1999).

<sup>4</sup> Control data from Hulleman and Humphreys (2007).

<sup>5</sup> ER scored lower than the specified range for paraphasias expected in Wernicke's aphasia, but this profile most closely matches her performance.

<sup>6</sup> Item 19 (Pretzel) was not tested because it is not in common usage in New Zealand English.



Table 2.6  
*Mean Scores and Standard Deviations on Diagnostic Tests by Aphasic Group with t-test Statistics*

Test	Non-Fluent		Fluent		<i>t</i>	<i>df</i>	<i>p</i>
	Mean	SD	Mean	SD			
<i>BDAE</i>							
BNT (% Correct)	60	25	60	17	0.01	11	> .05
<i>COWAT</i>							
Word Fluency (F, A and S: 3 mins)	7	6	23	7	4.03	9	< .01
Category Fluency: Animals (1 min)	9	4	11	4	1.12	11	> .05
Category Fluency: Fruit (1 min)	7	2	7	2	0.02	10	> .05
<i>Lukkien &amp; Wilshire Sentence Production Test</i>							
Sentence Score (N = 30)	3	3	5	5	0.8	9	> .05
Syntax Score (N = 30)	7	7	13	8	1.37	10	> .05
Root Verb Score (N = 20)	11	4	13	4	0.66	10	> .05
Function Word Score (N = 94)	57	23	70	14	1.25	10	> .05
Noun Score (N = 35)	22	8	23	5	0.23	10	> .05
<i>Lukkien &amp; Wilshire Single Noun Test (N = 35)</i>	26	5	26	4	0.1	11	> .05
<i>PALPA</i>							
Non-Word Repetition: Subtest 8 (N = 30)	14	10	13	10	0.13	10	> .05
Auditory Word Repetition: Subtest 9							> .05
High Imageability Words (% Correct)	88	10	75	21	1.4	7	> .05
Low Imageability Words (% Correct)	72	19	58	31	0.95	8	> .05
Imageability and Frequency Reading: Subtest 31							
High Imageability Words (N = 40)	34	6	37	3	1.13	9	> .05
Low Imageability Words (N = 40)	22	12	32	9	1.69	11	> .05
Spelling-Sound Regularity Reading: Subtest 35							
Regular Words (N = 30)	19	9	25	7	1.36	11	> .05
Exception Words (N = 30)	21	7	23	8	0.5	10	> .05
Non-Word Reading: Subtest 36 (N = 24)	3	7	16	7	3.43	11	< .01
<i>PPVT Standard Score</i>	85	15	78	19	0.76	10	> .05
<i>PCB</i>							

Table 2.6

*Mean Scores and Standard Deviations on Diagnostic Tests by Aphasic Group with t-test Statistics*

Test	Non-Fluent		Fluent		<i>t</i>	<i>df</i>	<i>p</i>
	Mean	SD	Mean	SD			
Lexical comprehension							
Within Category (N = 16)	15	2	16	1	1.27	6	> .05
Between Category (N = 28)	27	1	26	3	1.15	6	> .05
Sentence comprehension (Form A)							
Lexical distracters (N = 30)	27	2	27	3	0.31	9	> .05
Reverse role distracters (N = 30)	19	4	22	5	1.39	9	> .05
<i>Pyramids and Palm Trees Test: Three Pictures (% Correct)</i>	93	4	96	5	0.88	9	> .05
<i>Zinseger &amp; Berndt Noun-Verb Naming Test</i>							
High Frequency Nouns (N = 30)	24	5	25	2	0.47	9	> .05
Low Frequency Nouns (N = 30)	26	4	23	3	1.22	11	> .05
Verbs (N = 30)	21	9	21	5	0.06	9	> .05

Table 2.7

*Spontaneous Speech Samples taken from the Cookie Theft Scene Description in the BDAE*

Non-Fluents	
CT	Well um...   the mother is um   pouring out on the seat   um...   and the um   the girl and the boy is...   some cookies...   and um...   um...   ah...   and souping down on the cookie   and its falling....   mm...   [Truncated]
DA	/ɒl/ be...   /θ/ the ah   the ah   it's the ah   the the the ah   children's mother   is um oh   present present   is is ah /f f/ washing   the plates   /w w/ while she   she ah   wash /θɪŋθ/ the ah   the ah plates the   the um   the ah   oh basin   wash /besɪn/   is   overflow /fɒl/ flooring   [Truncated]
DD	House   ah...   cookies   fell   ah...   um   ah   um   house ah   ah   women no men   women   girls   ah sink   spill   pour   ah   women   brush   sink   ah sinks sinks   um...   uh doors door doors doors doors   [Truncated]
ECV	The um...   the little girl is um...   trying to get the ...   the um...   stuff from the little boy   but the little boy is falling down.   um...   and ah   the the lady is...   doing the dishes, but the sink is um falling.   I mean, it's running over   um...   that's -   I mean everything else is the same
JHM	Um...   the woman   ah /dedrum/ /debrun/   ah she um not um.../ʌ/ /ʌndʌstɪd/   um ...   water...   water um dripping   /ʌ/ /o/ /n/ the floor   and um a /ʌn/ she doesn't notice...   /drunɪŋ/   um the children   ah /wo/...   little boy...   girl...   um.../s/.../st/ stealing   ah cookies   [Truncated]
TB	The boy is handing the little girl a cookie   he's falling off the...   the the stool   and the lady is washing dishes and it's running over   um   she's got a plate in her hand   and...   they don't have no no drain board   [Truncated]
WS	Now, um...   Him I Gotta Go cookies...   and um...   Him I Gotta Go water...   um...   K saucers...   plates...   Him I Gotta Go   ... taps....   Him I Gotta Go....   /b/ bush....   trees and Him I Gotta Go ... {long pause}   Whoopsie!   Him I Gotta Go   Him I Gotta Go   [Truncated]
Fluents	
FS	Alright   we're in the   these are   they are three people   they are people in the kitchen   and the lady is ...   they got the   she can /hiədɪd/   hurt herself because   she's gonna fall, because the waters here   they go she does I dunno what she's talking about   [Truncated]
NP	First of all he...   he's falling off the...   the...   um, stool...   he he he's keep he's helping himself the /bɪs/   the biscuits that he's passing down, but it be...   /s/ /gʌlʌp/ the...   the stool   and...   uh mum has...   forgot to turn the   water off the tap and so it's...   [Truncated]

Table 2.7

*Spontaneous Speech Samples taken from the Cookie Theft Scene Description in the BDAE*

---

NS	What I see on this picture is a man at the   a young lady at the kitchen she's doing the...   /krʊkəri/ crockery supposed to be doing those, but the...   she's been /l-/ /lʌst/ of something a rather because all the water from the tank is overflowing is running over out into the /skɪn/ into the...   floor   [Truncated]
PS	Umm...   well the uh...   the kid, his stool, it's going to fall over...   uh the lady there, she's got the sink...   overflowing   (Experimenter: Yeah) umm...   I'm not too sure about that, but I mean the- the window's /o/ it looks like it's open, but I'm not too sure about that   [Truncated]
STR	Right, this, the, /ʌ/ they are   doing the dishes and the mother is doing the dishes and /thɪzə/ the son and the /ʌ/ daughter is...   getting a...   biscuit up in the   up in the   biscuit in the /ʌ/   um...   up up um ...   uh...   [Truncated]
XX	Was he washing dishes [?]   water flowing in the floor from um   they they're cleaning um /lɪtə lɪdʒ/   ah the water come down the the   what's-her-name paying no attention   and the boy   is reaching up to the /kælbɪd/ and all this foam fell off the /st/ stub /w wɪd/ trying to get um...   [Truncated]

---

### *Non-Fluents*

*CT.* CT presents with grammatically simple spontaneous speech that has a slow and hesitant quality. Her speech is disrupted by frequent word-finding pauses, and occasional phonemic and semantic paraphasias. On administration of the BDAE during 2006, she did not clearly fit into any of the BDAE aphasia profiles. She presented with better melodic line, phrase length, and articulatory agility scores than a pure Broca's aphasic. However, she also produced more paraphasias than are usually observed in Broca's aphasics (1-2 per minute of conversation), and her comprehension score fell below the Broca's aphasia range. At best, CT could be described as a mixed non-fluent aphasic. CT's comprehension deficit mainly affects her sentence level comprehension, but her speech production deficit clearly influences both single word and sentence production, especially the latter. CT has a mild repetition deficit that especially affects low-imageability words and non-words. She also has a severe phonological dyslexia. Her high imagery word reading is better than her low imagery word reading.

*DA.* DA presents with non-fluent speech that is characterised by articulation difficulties, flat intonation, use of fillers, long word finding pauses, and grammatical errors that he occasionally self-corrects. On administration of the BDAE during November 2006, he was diagnosed with Broca's aphasia. DA has intact single word comprehension, but a mild sentence comprehension deficit. He has a mild single word production deficit, but severely non-fluent sentence production. He also has a repetition deficit that influences his sentence and single word repetition, and a mild phonological dyslexia.

*DD.* DD presents with severely non-fluent spontaneous speech, which primarily consists of content words and fillers. His speech appears effortful, and other people are often relied on to drive conversation. He was diagnosed with Broca's aphasia on a speech and language pathology report dated 2004. Re-administration of the BDAE during 2006/2007 confirmed that this was the best diagnosis although his comprehension score fell below the Broca's range (17/100), as did his sentence repetition score (0/10). DD's comprehension deficit mainly affects his sentence level comprehension. His comprehension of sentences that contain lexical distracters was below the normal range, but was better than his comprehension of sentences that contain reverse role distracters, where he performed below chance. He also showed a severe speech production deficit that influenced both single-word and sentence production. DD also presented with a repetition impairment, and a severe deep dyslexia.

*ECV.* ECV presents with spontaneous speech marked by frequent word finding pauses and fillers, and occasional paraphasias. However, she occasionally produces surprisingly long sentences. On a speech and language pathology report from 2001, ECV was diagnosed with transcortical motor aphasia. In 2006, when examined using the BDAE, her profile best matched that of an anomic aphasic: Her spontaneous speech was characterised by good articulation and few errors, and her sentence repetition was fair (7/10). However, her auditory comprehension score fell below the range expected of both transcortical motor and anomic aphasics (37/100). ECV has a borderline single-word comprehension deficit, and moderately impaired sentence comprehension. Despite her fair performance on the repetition component of the BDAE (7/10), she presented with a moderate word repetition deficit on the PALPA Auditory Word Repetition Test (Subtest 9) and the PALPA Non-Word Repetition Task (Subtest 8). ECV has phonological dyslexia. Her high imagery word reading is better than her low imagery word reading. Because of her predominantly sentence level comprehension deficit, her non-fluent speech characteristics, and her previous diagnosis with a non-fluent aphasia, ECV has been classified as non-fluent despite her occasional well-formed sentence production.

*JHM.* JHM presents with severely non-fluent spontaneous speech that is characterised by long word finding pauses, effortful articulation, and frequent phonemic paraphasias. JHM's language was assessed during March 2004, and she was diagnosed with Broca's aphasia. She has good comprehension of single words, but a comprehension deficit that particularly affects her understanding of grammatically complex sentences. JHM has a mild single word production deficit, but a severe sentence production deficit. She also shows a severe sentence repetition deficit, but her single word repetition is considerably better, as is her non-word repetition. JHM has a mild real word reading deficit, with a severe non-word reading deficit. Usually her reading fits the pattern of a phonological dyslexic. However, she very occasionally makes semantic errors, suggesting that she may have a borderline deep dyslexia. Although both her high and low imagery word reading are impaired, her high imagery word reading is better than her low imagery word reading.

*TB.* TB presents with spontaneous speech production that is mildly impaired, but includes frequent word finding pauses. However, she is able to hold fairly normal conversations, and to successfully work in a cafeteria part time. A speech and language pathology report from 2002 indicates that TB initially presented with Broca's aphasia. However, a subsequent report from 2004 indicated some language recovery, as she then

fitted the classification of an anomic aphasic, and examination of TB's speech using the BDAE during 2007 showed minimal evidence of aphasia. However, her performance on the more specific tests that are detailed in Table 2.4 and summarised in Appendix A3, suggests that she still has considerable language impairments, especially with sentence level processing. Taken together the tests suggest that TB is best described as a recovered Broca's aphasic. TB has good single word comprehension, but impaired sentence level comprehension that particularly affects grammatically encoded information. She also has moderate single word and sentence production deficits, and a mild phonological dyslexia. Her real word reading is mildly impaired, but she cannot read non-words.

*WS.* WS has spontaneous speech that is severely non-fluent, and marked by a common perseveration of his favourite horse's name, "Him I Gotta Go" (WS was a successful jockey and horse trainer before he retired). When examined using the BDAE, his performance did not correspond with one aphasia profile. The closest match was to Broca's aphasia, where he scored within the specified range on all measures apart from the sentence repetition component where he scored at floor, and the auditory comprehension measures, where he scored below the specified range (42.5/100). WS's single word comprehension is good, but he has a moderate sentence comprehension deficit that particularly affects his understanding of grammatically encoded information. He has a severe language production impairment that affects both sentence and single word production. WS's verb production is worse than his noun production. He also presents with a moderate repetition impairment, and a severe phonological dyslexia.

#### *Fluents*

*FS.* FS presents with fluent speech that is interrupted by word finding pauses. While her melodic line and articulatory agility is good, her speech sometimes presents with a slightly slurred quality. When examined using the BDAE, FS presented with good melodic line, phrase length, and use of grammatical forms, but poor repetition, word finding, and auditory comprehension. She is best diagnosed as a Wernicke's aphasic, although on administration of the BDAE, she made fewer paraphasias than stipulated in the Wernicke's profile. Although FS did not make enough paraphasias on the BDAE to fit the Wernicke's profile, she did make four semantic paraphasias on the Lukkien Sentence Production Test, implying that she does have a considerable impairment in this area. FS's single-word and sentence comprehension is impaired, and she has a moderate speech production deficit. Both her noun and verb production are impaired, but she has

particularly bad verb production. FS also has a severe repetition impairment, and a phonological dyslexia.

*NP.* NP presents with fluent speech that incorporates a variety of grammatical constructions, but is disrupted by frequent word finding pauses. He has a moderate single word production deficit, a moderate word repetition deficit, a severe non-word repetition deficit, and a mild reading deficit, but good comprehension. When NP's language abilities were examined using the BDAE, he showed word finding difficulties and produced some paraphasias. NP was diagnosed with anomic aphasia.

*NS.* NS presents with fluent speech that is characterised by frequent paraphasias, and he frequently has obvious difficulty following conversations. When his language was examined using the BDAE, he presented with a borderline Wernicke's/Conduction Aphasia. NS has a comprehension deficit, but good non-verbal semantic access. His single word production is moderately impaired, but he has a severe word repetition deficit, and he is completely unable to repeat non-words. NS's reading is well preserved.

*PS.* PS presents with fluent spontaneous speech, but experiences word finding difficulties. During 2006, his language abilities fitted the BDAE classification of an anomic aphasic. His comprehension was good but he had a mild word production deficit, and showed more difficulty producing verbs than nouns. His word repetition was intact, although he showed a mild non-word repetition impairment. PS's reading was good.

*STR.* STR's spontaneous speech is fluent, but she makes paraphasias and word finding pauses. She was diagnosed with anomic aphasia using the BDAE. Her single word comprehension and semantic access are good, however she exhibits a mild comprehension deficit on grammatically complex sentences. She has a moderate single-word production deficit, and a moderate single-word repetition deficit. STR also has a mild reading impairment, although she does have considerable success reading single words and non-words.

*XX.* XX presents with fluent speech with an empty quality, and many word finding pauses. Examination of his language using the BDAE indicated that he has Wernicke's Aphasia. He has moderate single-word and sentence comprehension deficits, although his non-verbal semantic access is good. XX has moderate to severe single-word and sentence production deficits, and his ability to repeat aurally presented stimuli is also impaired, although his repetition of high imageability words is better than his repetition of low imageability words. XX also has impaired reading, and like his repetition, his reading of high imageability words is better than his reading of low imageability words.



### *Brief Comment on the Aphasic Participant Groups.*

Both the non-fluent and fluent group of aphasics were made up of participants with heterogeneous impairments. In the non-fluent group, three participants had relatively intact single word production, scoring over 40 on the BNT (DA, ECV, and JHM). Because non-fluent aphasics with relatively good single word production impairments are over-represented in the literature on context-specific naming deficits, these three participants were expected to show exaggerated competition condition effects on the current experimental tasks.

### *Materials*

The following six categories, with the corresponding sub-categories (see Table 2.1 for details of the sub-categories), were selected from those trialled in the pilot study for use in the current experiment: Animals, Supermarket Items, Home Objects, Clothing, Occupations, and Tools and Equipment.

### *Design*

This study is based on the Randolph et al. (1993) category exemplar task. However, several modifications have been made to the Randolph et al. task design. First, we doubled the response periods allowed: Where Randolph et al. allowed 1 minute for broad categories and 15 seconds for sub-categories, we allowed 2 minutes and 30 seconds respectively. We did this because some aphasic participants find it difficult to produce even one item name during a 15 second interval. Second, Randolph et al. presented each participant with each category once only, controlling for condition across participants. In the pilot study it was noted that there were high levels of variation in the number of exemplars produced across category and participant, so we gave each aphasic participant each category twice, once in the high competition condition and once in the low competition condition (this was done in different testing sessions).

Four versions of the task were created to control for serial order effects, and to allow for retesting of participants. Task versions one and two presented the categories in the following order: *Animals, Tools and Equipment, Clothing, Items You Find in a Home, Items You Find in a Supermarket*, and *Occupations*. Task versions three and four presented the categories in the reverse order. Task versions one and three started with one

category in the high competition condition, followed by two categories in the low competition condition, two categories in the high competition condition, and one category in the low competition condition: Task versions two and four presented the high/low competition conditions in the reverse order.

### *Procedure*

Psyscope software on a Mac laptop computer was used to present and time the experiment. Each aphasic participant and three control participants (AP, EM, and GA) were given a single version of the task on each of two experimental sessions. If they were given version one during the first session, they were given version two during the second session. Similarly, if they were given version three during the first session, they were given version four during the second session. This ensured that any serial order effects operated on the same categories during both sessions: Also, if they received a category in the high competition condition during their first session, they would receive the same category in the low competition condition during their second session, and vice-versa. The remaining control participants completed a slightly different task design where competition condition was compared across participants: The differences in design and data analysis plus the results from these controls are reported separately in Appendix A4 (these data show the same general trends as is reported in the main text to follow, except a main effect of competition condition was not observed).

During each session, participants were asked to name as many items as possible that belong to each of the aforementioned semantic categories. In the high competition condition, the broad category names were presented both in written form on the computer screen and aurally, and participants were given two minutes to respond. In the low competition condition, participants were shown the broad category name first, followed by the name of one of the four sub-categories. Participants were given 30 seconds to respond for each sub-category. At the end of each 30-second interval, they were presented with the next sub-category, and so on, until all four sub-categories were completed. Thus, the total recall period (combined time for completion of all four subcategories) was 2 minutes in the high and low competition conditions.

At the beginning of each session, after the initial task instructions, participants were given several example/practice trials, as follows: High Competition Condition - *Countries of the World* (2 minutes); Low Competition Condition - *Countries of the World*, *Countries*

*in Europe* (30 seconds) and *Countries in the Pacific* (30 seconds); High Competition Condition – *Musical Instruments* (2 minutes); Low Competition Condition – *Names, Boy's Names* (30 seconds) and *Girl's Names* (30 seconds).

#### *Data Analysis*

Each response was coded as correct or as an error. A response was considered correct if it was a genuine exemplar of the target category, except where the item formed one of the exceptions detailed in Table 2.8.

Table 2.8  
*Category Exemplar Task: Scoring Procedure*

<i>Rule</i>	<i>Description</i>	<i>Examples</i>
Gender Variants	Where gender variants were produced, only the first item produced was given a score.	"Waiter" and "Waitress" (1 point): "Cow" and "Bull" (1 point)
Unrecognised Items	If the item produced was not recognised by the experimenter, and was not found in the dictionary, no score was given for the item.	
Synonyms	Where synonyms were produced, only the first item produced was given a score.	"Candy" and "Sweets" (1 point)
Item Repetition	Where an item was repeated within a category, only the first instance of the item was scored (repetition between related sub-categories was not allowed).	
Word Repetition	Any responses where a word was repeated were not given a point	"Dining Table" and "Coffee Table" (1 point)
Basic with Subordinate	If a basic item was named with subordinates of that item, only the subordinates were given a point.	"Apples", "Brayburn", and "Gala" (2 points)
Superordinate with Basic	If a superordinate was named, it was given a score only when no basic descriptors were produced.	"Tinned Food" (1 point): "Tinned Food" and "Tinned Tomatoes" (1 point)
Imageability	Items produced had to be specific enough to elicit a clear image of a single object to receive a point. Items ending with the words 'item', 'gear', 'objects', etc., would not receive a point.	"Rain wear" (0 points): "Delicatessen items" (0 points): "Apple" (1 point)
Speech Errors	If an error was made during item production, no point was given. These errors included phonological paraphasias, articulation errors, and circumlocutions.	"Saucepans (1 point) and Spots" (0 points): "Eye, um /oʊɑmɪtri/ eyes, um" (0 points)

It was of concern that for some of our categories, the low competition condition may be more difficult than the high competition condition, due to the sub-categories being too restrictive. To avoid any difficulty differences across conditions, we wanted to balance the number of potential exemplars available as much as possible. Hence, we decided to only use the categories in which non-brain damaged individuals produced comparable numbers of exemplars for both conditions. Preliminary paired t-tests were completed to compare the control group's rate of exemplar production in the high and low competition conditions for each category. The analyses indicated that exemplar production rates were

significantly lower in the high competition condition than the low competition condition for *Clothing* ( $t(2) = 4.330, p < .05$ ) and *Tools and Equipment* ( $t(2) = 6.614, p < .05$ ): No significant differences were observed for the remaining categories (*Animals* ( $t(2) = 3.928, p > .05$ ); *Items You Find in a Home* ( $t(2) = 0.655, p > .05$ ); *Occupations* ( $t(2) = 0.218, p > .05$ ); *Items You Find in a Supermarket* ( $t(2) = 3.179, p > .05$ )). For this reason, data from the *Clothing* and *Tools and Equipment* categories are excluded from all remaining analyses.

The number of correct responses made in the high and low competition conditions were analysed at the group and individual levels. At the group level, a general linear mixed model was performed, including the following independent variables: group, condition, and a group by condition interaction (participant and category were included as random, repeated factors): Where a variable was not significant, it was removed from the analysis. At the individual level, one-group chi square tests of independence were completed to compare the number of exemplars produced by each participant in the high and low competition conditions.

### Results

An alpha level of .05 was used for all statistical tests.

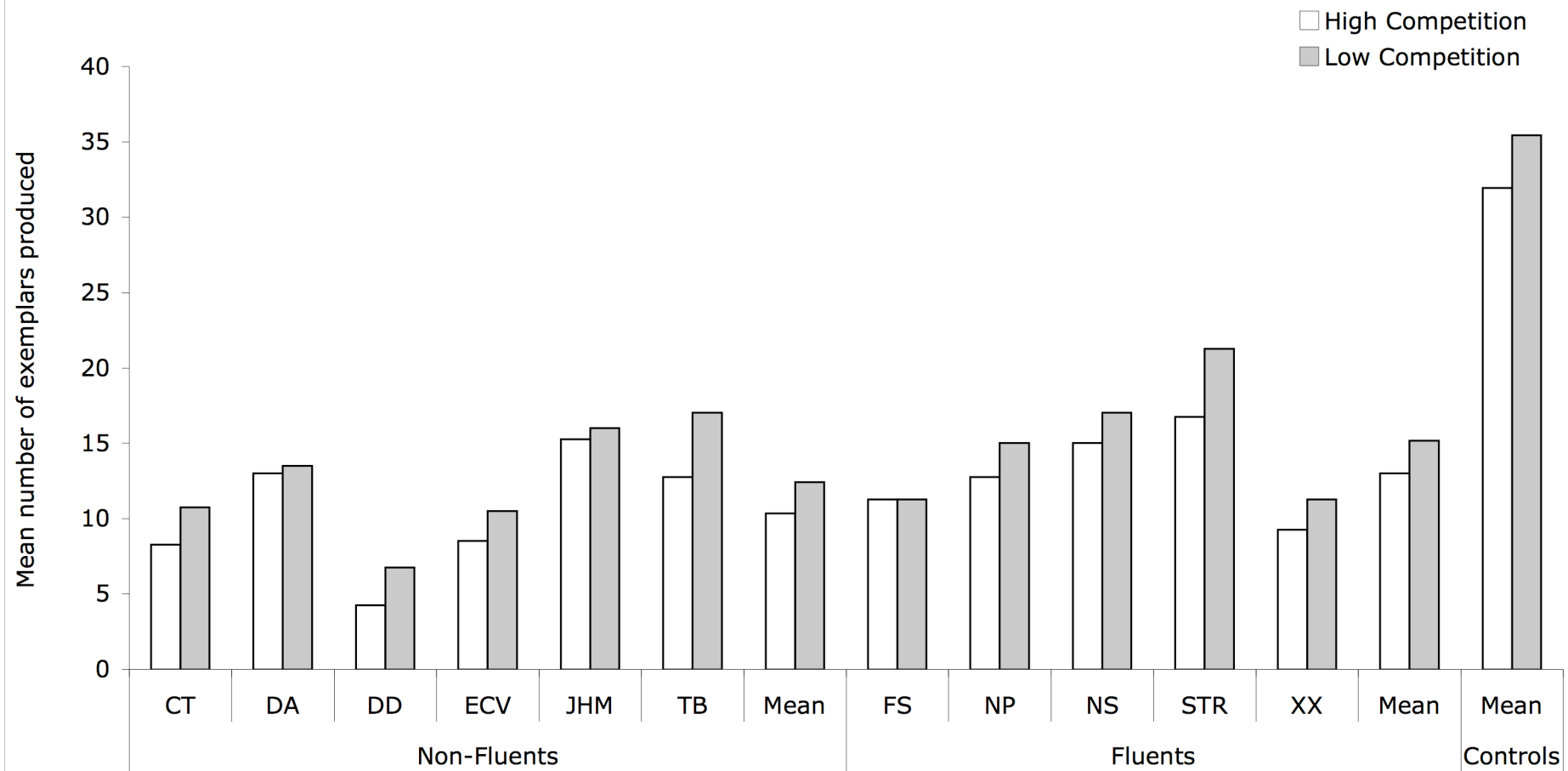
No data was available from non-fluent WS because he requested that testing was discontinued during the practice trials (he found the practice trials very difficult and frustrating).

The mean number of correct exemplars made by each participant and group in the high and low competition conditions is illustrated in Figure 2.1<sup>4</sup>. There was a strong general trend for participants to produce more exemplars in the low competition condition than the high competition condition. At the group level, this low competition condition advantage was confirmed by a highly significant main effect of condition ( $F(1, 92) = 10.40, p < .01$ ). A highly significant main effect of group was also observed ( $F(2, 92) = 43.94, p < .01$ ): However, contrasts confirmed that although the non-fluents and the fluents produced significantly fewer exemplars than the controls ( $(F(1, 92) = 81.74, p < .01)$ , and  $(F(1, 92) = 59.60, p < .01)$  respectively), there was no significant difference in the number

---

<sup>4</sup> Data from fluent NS for the occupations category was excluded from the statistical analyses because of a technical error.

of exemplars produced by the two aphasic groups ( $F(1, 92) = 1.50, p > .05$ ). Further, the expected interaction of group and competition condition was not significant ( $F(2, 90) = 0.44, p > .05$ ), suggesting that the non-fluents and the fluents did not respond in a reliably different way to the competition manipulation. However, the data trends were in the predicted direction, with the non-fluents producing a 20 % increase in the number of exemplars produced in the low competition condition, the fluents showing a 17 % increase, and the controls showing an 11 % increase.



*Figure 2.1.* Category exemplar task: Mean number of correct exemplars produced by participant and competition condition.

At the individual level, chi-square tests of independence did not reveal significant high/low competition condition effects for any of the participants, however these results may reflect a lack of statistical power (Non-fluents: CT ( $\chi^2(1) = 1.316, p > .05$ ), DD ( $\chi^2(1) = 2.273, p > .05$ ), DA ( $\chi^2(1) = 0.038, p > .05$ ), ECV ( $\chi^2(1) = 0.842, p > .05$ ), JHM ( $\chi^2(1) = 0.72, p > .05$ ), TB ( $\chi^2(1) = 2.430, p > .05$ ). Fluents: FS ( $\chi^2(1) = 0, p > .05$ ), NP ( $\chi^2(1) = 0.73, p > .05$ ), NS ( $\chi^2(1) = 0.038, p > .05$ ), STR ( $\chi^2(1) = 2.132, p > .05$ ), XX ( $\chi^2(1) = 0.78, p > .05$ )). Because no significant high/low competition condition effects were observed in the individual data, it was not possible to assess whether the individuals with good single word naming and poor sentence production were more strongly influenced by the competition manipulation than the other participants.

### Discussion

This experiment examined the category exemplar production performance of three groups of participants (non-fluent aphasics, fluent aphasics, and older controls) under two conditions. In the high competition condition, participants were given a broad category and were asked to provide as many exemplars as possible (e.g., *Animals (120 sec.)*). In the low competition condition, participants were sequentially presented with four sub-categories and were given 30 seconds to produce exemplars from each of the sub-categories (e.g., *Animals*: then *Pets (30 sec)*, then *Farm Animals (30 sec)*, then *Jungle Animals (30 sec)*, then: *Water Animals (30sec)*). We predicted that because the high competition condition included broad categories, the increased number of possible exemplars would lead to activation of a greater number of potential candidates, and consequently, higher overall levels of lexical competition. Therefore we expected that this condition might place particularly strong demands on a selection and control mechanism that functions to resolve lexical competition.

The first hypothesis was that the high competition condition would result in relatively lower levels of correct exemplar production than the low competition condition in all participant groups. This prediction was supported with the participants producing more exemplars in the low competition condition. The second, and most critical hypothesis, was that if non-fluent aphasics suffer from a deficit involving a selection and control mechanism, they should show abnormally exaggerated exemplar production costs



in the high competition condition, relative to fluent aphasics and neurologically intact controls. At the group level, although both of the aphasic groups produced fewer exemplars overall than the controls, a group by condition interaction was not observed, suggesting that the size of the high competition condition exemplar cost was not reliably different across the three groups. However, the data trends were in the predicted direction, suggesting that it may be worth investigating this in a future study, perhaps using a more powerful design and a larger number of participants. At the individual level we did not observe any support for this hypothesis, with none of the participants showing a significant condition effect. The final prediction made was that non-fluent aphasics with marked dissociations between their sentence and word production abilities would show particularly exaggerated high competition condition costs (that is DA, ECV, and JHM). The data did not support this hypothesis.

One concern that arises in this task is the potential presence of opposing competition effects. It can be argued that exemplars produced in the low competition condition are likely to be more semantically similar, and that this semantic closeness results in increased competition (proximity effects) (see the Crutch and Warrington (2004, 2005) discussion on refractory access deficits). For example, the farm animals *pig* and *cow* share more semantic features than the animals *pig* and *snake*, hence *pig* and *cow* are more difficult to distinguish, and hence, are stronger competitors. Following this logic, our low competition condition could be liable to increased competition due to increased semantic proximity between potential exemplars. Although the results of the task are consistent with our assumption that less competition occurs in the smaller sub-categories, it is possible that semantic proximity also contributes to the results. If this is the case, it would be more difficult to observe the differences that we hypothesised would occur as the semantic proximity effects would oppose the predicted competition condition effects. However, the more participants spontaneously cluster their responses in the high competition condition, the less you would expect to observe semantic proximity effects between the high and low competition conditions.

As mentioned previously, the lack of significant results at the individual level may be partially due to a lack of statistical power. Significant effects have been reported from an individual participant with a frontal lesion, PM, on the very similar task presented by Randolph et al. (1993), suggesting that it is possible to see significant differences at the individual level using a category exemplar paradigm with a highly similar task design. However, there are several potential reasons why PM might have shown a significant effect

where our participants did not. First, Randolph et al.'s, comparison was made across categories (i.e., their patient completed different categories in the broad (high competition) and narrow (low competition) conditions), whereas our comparisons were made within categories (i.e., our participants completed the same categories in the high and low competition conditions). Since we did notice that some participants showed much better exemplar production in specific categories, it is possible that category specific effects may have contributed to PM's "competition" effect: For instance, by chance he may have had better general knowledge of those categories included in the narrow condition than the broad condition. It is also possible that there was something about PM's deficit that lead to such a clear dissociation between his performances on the broad/narrow conditions. His deficit was very different to those seen in our patient groups, being the result of a gunshot wound to his anterior frontal lobes, which did not result in an aphasia/anomia. Possibly PM had difficulty planning his responses, switching between response clusters, initiating responses, or sustaining behaviour over the response period. Another possibility is that PM was exhibiting a clear cognitive fluency cost in the broad condition, and that our aphasic patients did exhibit this same cognitive fluency cost, but it was being masked by the considerable verbal fluency deficits that they also experience.

In conclusion, the data from this task do not support the selection and control hypothesis. It is possible that the task does not efficiently tap into lexical competition because too many candidate exemplars are activated in both conditions. Also, task performance may be affected by strategic factors, which might further confound the results. The possibility that selection and control deficits may be buried behind the influence of other factors is a concern, because when participants are producing category exemplars, they rely on a number of cognitive functions. For instance, participants draw upon: general knowledge (this can create noticeable gender effects for categories such as building tools and equipment); working memory (they are required to keep the task online for long periods of time); task monitoring (this is required so that they can keep track of what they have already said, so as to avoid repeating items); task switching (e.g., where strategy changes are employed within a category); and finally word retrieval, and perhaps selection and control. Because all of these factors potentially influence the results, this task, perhaps more so than the others in this series of experiments, is not a pure measure of the selection and control hypothesis. This may account for the lack of support for the hypothesis that has been observed. However, the next three experiments examine the influence of lexical competition using tasks that create more specific competition, by targeting particular

competitor words. With this more directed competition, the following tasks may more successfully identify competition differences between the groups.

### Verb Generation Task Introduction

In verb generation tasks participants are shown a series of items and, for each item, they are required to say the first semantically associated verb they think of. For example, presented with the word *Scissors*, a participant could say, “Cut”, or presented with the word *Penny*, a participant could say “Spend”, “Pay”, “Buy”, “Flip”, or “Drop”, etc.) (see Thompson-Schill et al., 1997; Thompson-Schill et al., 1998; Thompson-Schill et al., 1999; Martin & Cheng, 2006). In this paradigm, items like *Scissors*, which have one dominant verb associate (high response strength), are often compared to items like *Penny*, which have many verb associates (low response strength).

Thompson-Schill et al. (1997) used fMRI to measure activation levels in the LIFG of neurologically intact participants during verb generation. They found significantly more activity in the LIFG during low response strength trials (nouns with many verb associates) than high response strength trials (nouns with one dominant verb associate). In a second study, Thompson-Schill et al. (1998) sought corroboration of their fMRI results using a lesion study. The verb generation task was administered to four groups: patients with LIFG lesions; patients with left prefrontal cortex lesions that excluded the LIFG; patients with right prefrontal cortex lesions; and a group of elderly controls. As predicted, the patients with LIFG lesions produced more errors than the other groups in the low response strength condition, and there were no significant differences between the groups in the high response strength condition. Thompson-Schill and colleagues suggested that responses to the high response strength items reflected relatively low levels of lexical competition, because there were few potential responses to select between: Responses to the low response strength items reflected relatively high levels of lexical competition, because there were many potential responses to select between. They suggested that when *competition* levels are high, top-down modulation is required to select the best response. This modulation / selection process is hypothesised to be a function of the LIFG (Thompson-Schill et al., 1997; Thompson-Schill et al., 1998; Thompson-Schill et al., 1999). If this interpretation is correct, then the verb generation task may provide a unique window into the controlled lexical selection processes being investigated in the current research, as more control would be required during high competition conditions than low. Further, if the LIFG does play an important role in the frontal modulatory processes influencing lexical selection, and non-fluent aphasics have impaired modulatory

functioning, they may show particularly marked impairments in high competition conditions.

However, a major debate has occurred regarding whether low response strength costs really reflect increased competition between co-activated responses, as is suggested by Thompson-Schill and colleagues (e.g., Thompson-Schill et al., 1997; Thompson-Schill et al., 1998). An alternative explanation is presented by Wagner, Paré-Blagoev, Clark, and Poldrack (2001), who suggest that low response strength costs are not the result of competition *per se*, rather the cost reflects how much *controlled retrieval* is required (see also: Martin & Cheng, 2006). Where the competition view suggests that selection difficulties arise after a failure to boost lexical activation levels of the best response above activation levels of the other options available, the controlled retrieval view suggests that a failure to select an appropriate response could plausibly occur in a situation where there is only one correct answer. According to this hypothesis, when the association between an object and a verb is strong, automatic processes can select the response: When the association is weaker, a frontal controlled activation process guides retrieval, perhaps through strategy implementation.

There is some support for the controlled retrieval view. Martin and Cheng (2006) took the stimuli used by Thompson-Schill and colleagues (Thompson-Schill et al., 1997; Thompson-Schill et al., 1998), and with their own norms, calculated association strength ratings for the stimuli by dividing the frequency of the most common response by the total number of responses for that item. This gave them a measure of the strength of the association between the noun stimuli and the most common response that disregarded the amount of competition that may have occurred between co-activated alternative responses. This association strength measure can be contrasted to the response strength measure: These were also calculated by dividing the frequency of the most common response by the frequency of the second most common response, giving a measure that reflected how strongly co-activated the two most common responses were. Martin and Cheng examined both association strength and response strength and observed that the high response strength items were more highly associated than low response strength items. They concluded that response strength and association strength were confounded in the studies by Thompson-Schill and colleagues. The theoretically important implication of this observation was that competition, and selection between competing alternatives, may not account for the low response strength costs that had been observed in prior studies. Rather,

it was possible that the cost reflected controlled retrieval processes, such as strategy use, that can occur in the absence of strong lexical competition.

Martin and Cheng (2006) followed up this observation using a verb generation task where association strength was accounted for. They used three groups of items: First, low response strength items with high association strengths (e.g., *Door*: “Open” (association strength of 0.51) / “Close” (association strength of 0.45)); Second, low response strength items with low association strengths (e.g., *Map*: “Read” (association strength of 0.24) / “Find” (association strength of 0.26), “Travel” (association strength of 0.10)) and; Third, high response strength items with high association strengths (e.g., *Apple*: “Eat” (association strength of 0.58) / “Cut” (association strength of 0.05)). The results showed that their participants’ responses dissociated according to association strength, rather than response strength. Responses were significantly worse in the low response strength with low association strength condition, than the high response strength with high association strength condition, and the low response strength with high association strength condition. However, there was no difference between the low response strength with high association strength condition, and the high response strength condition with high association condition. Because the results did not show a difference between the low response strength conditions, and the high response strength condition, the authors argued that the results support the association strength hypothesis, but not the competition hypothesis. The implication of these results was that high levels of lexical co-activation and competition did not appear to instigate the response cost, rather some more effortful, and perhaps strategic, process was called on and accounted for the cost.

However, the primary interest of the current research is to find out how non-fluent aphasics respond to different levels of lexical competition. Both the competition and controlled retrieval views suggest that a frontal modulation process may be used to aid lexical selection. So, while the competition versus controlled retrieval debate does have some interesting implications in terms of the lexical processes involved in word production, such as which situations instigate use of modulatory processes (i.e. automatic processing failures, conscious strategy use, or situations of high competition), the distinction does not have clear ramifications in terms of the current research goal, and is discussed as a secondary issue.

Importantly, most existing verb generation studies, including those discussed in this introduction, have focussed on our understanding of semantic retrieval processes. Few studies have examined the relationship between the data observed and lexical selection, or

spoken language production more generally. However, the associations between non-fluent aphasia and LIFG damage, and LIFG damage and impaired selection / modulation on verb generation tasks have been highlighted as warranting further research (see Thompson-Schill, 2005; Thompson-Schill, Bedney, & Goldberg, 2005): This is a goal of the current study.

### *The Current Study*

According to the selection and control hypothesis, a frontal modulation process will be required more for low response strength items (our high competition condition) than high response strength items (our low competition condition). Here we test this hypothesis on three groups (non-fluent aphasics, fluent aphasics, and neurologically intact controls). The increased demand on the frontal modulation process in the high competition condition is expected to result in a lower accuracy and/or slower responding across groups (see Thompson-Schill et al., 1998). It is expected that relatively minimal demands will be placed on the modulation process in the low competition condition, because of the limited lexical co-activation expected in this condition.

Further, it is hypothesised that the non-fluent aphasics will show exaggerated high vs. low competition condition effects relative to the fluent aphasics and older controls, because some or all non-fluent aphasics may suffer from a specific impairment to a lexical selection and control mechanism: This is expected to be reflected in a group by competition condition interaction.

Finally, based on findings that the sub-group of non-fluent aphasics who show good naming but poor sentence production are particularly sensitive to manipulations of context in naming tasks (see Freedman et al., 2004; Schwartz & Hodgson, 2002; Schnur et al., 2006; Wilshire & McCarthy, 2002; Wilshire et al., 2006), we hypothesise that non-fluent aphasics who fit this profile may show stronger competition effects than the non-fluent aphasics that do not. Three non-fluents in the current sample have relatively good naming, scoring above 40 on the BNT, and are expected to show these exaggerated effects (see p. 64 for more details).

The procedure used is based on that used by Thompson-Schill et al. (1997), however, some modifications have been made. Rather than presenting the noun stimuli entirely in written form (a procedure which may disadvantage individuals with reading difficulties), our stimuli were presented multi-modally. Each written noun was presented

with a picture of the object, and an auditory recording of the noun (see also Martin and Cheng (2006) who used visual and auditory presentations of stimuli to adapt the task to aphasics' requirements). As is detailed in the verb generation pilot study section below, we also constructed a new set of stimuli, which were piloted on fifty first-year students at Victoria University of Wellington.

## Verb Generation Pilot Study

### *Method*

#### *Participants*

Fifty first-year students (35 female, 15 male) at Victoria University of Wellington participated in this study for course credit. All participants spoke English as their first language. Informed consent was obtained from all participants.

#### *Materials*

Two-hundred object names were sourced from the MRC Psycholinguistic Database (Coltheart, 1981)<sup>5</sup>, the Age of Acquisition set (Morrison, Chappell, & Ellis, 1997), and the Timed Picture Naming set (Székely, D'Amico, Devescovi, Federmeier, Herron, Iyer, Jacobsen, & Bates, 2003)<sup>6</sup>. A full list of the object names used is included in Table 3.1. Where possible, a picture of each object was obtained from the colour Snodgrass set (Snodgrass & Vanderwart, 1980) or the Székely et al. (2003) set. When a normed picture of an object was not available, simple clip art pictures of the objects were sourced from the internet<sup>7</sup>.

---

<sup>5</sup> The MRC Psycholinguistic database is available online at [http://www.psy.uwa.edu.au/mrcdatabase/uwa\\_mrc.htm](http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm)

<sup>6</sup> The Timed Picture Naming set is available online at <http://www.crl.ucsd.edu/~aszekely/ipnp/method.html>.

<sup>7</sup> Note: the aural presentation of the noun was omitted in the pilot study.



### Procedure

Participants were tested in groups of one to fifteen. Following a brief introduction to the area of research, the participants were told that they would be shown a series of objects, and that their job was to write down what the object does, or what we do with the object. As well as providing several examples of how the task should be done, the instructions were reiterated in a written format (see Appendix B2 for the written task instructions), and the participants were given the opportunity to ask questions before the experiment progressed. Two hundred verb generation trials were then completed in a fixed pseudo-random order. The objects (words and pictures) were presented to the participants on paper, with a space to the side of the each object where the participants wrote in their response (see Figure 3.1).

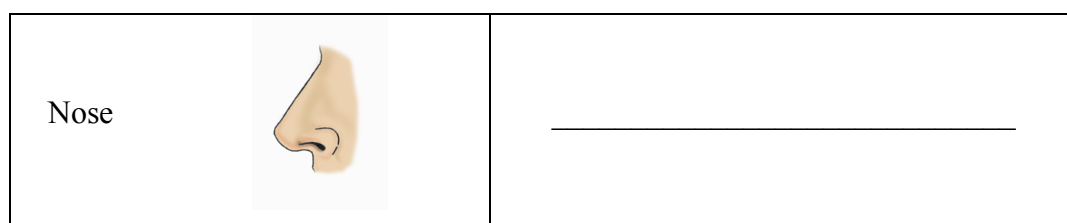


Figure 3.1. Verb generation pilot: Example trial.

### Data Analysis

The data was tidied: Obvious spelling errors were corrected, and inflectional variations of the same word root were standardised (e.g., “Swimming” and “Swims” were changed to “Swim”). Any responses that fit into the following four categories were coded as errors: 1) Non-verb responses (e.g., *Beach*: “Waves”); 2) No response trials; 3) Made up words (e.g., *Fridge*: “Coldens”), and; 4) Unintelligible responses. If a participant made errors on more than 30 trials, all of their data was excluded from further analysis, because it was evident that they did not fully understand/follow the task instructions. This criteria lead to the exclusion of data from seven participants.

The frequency of each response was calculated for each object, and response strength ratios were calculated by dividing the frequency of the most common response by the frequency of the second most common response. These response strength ratios were used to choose objects for the verb generation task proper (see the Materials section of the verb generation task proper for further details).

## *Results*

A response strength ratio was calculated for each piloted object by taking the frequency of the most common response and dividing that by the frequency of the second most common response. Please refer to Table 3.1 for a complete list of the piloted objects with response strength ratios.

Table 3.1

*Verb Generation Pilot: List of Objects Included with  
Response Strength Ratios*

	<u>Response Strength Ratio</u>
Adder	2.56
Airplane	12.67
Alligator	2.78
Anchor	1
Ant	3
Arrow	12
Axe	1.05
Baby	1.6
Bag	4.6
Ball	17.5
Balloon	1.25
Bandage	2
Barbecue	1.38
Basin	9
Basket	13
Bath	1.3
Beach	3.25
Bed	10.33
Beetle	1.78
Bell	11.67
Binoculars	2.1
Blanket	1.21
Blouse	3.83
Boat	1.38
Bomb	5.8
Book	5.83
Bracelet	2.43

Table 3.1

*Verb Generation Pilot: List of Objects Included with Response Strength Ratios*

	<u>Response Strength Ratio</u>
Bread	4.2
Bricks	3.33
Bride	2.13
Bridge	3.5
Broom	13
Bucket	1
Butter	2.7
Cactus	3.67
Cake	3.6
Camera	1
Can	11
Candle	1.38
Cane	1.78
Cannon	1.7
Canoe	2
Car	10.67
Caravan	1.8
Carnation	2.13
Caterpillar	3.67
Cattle	1.57
Chain	1.6
Chair	11.33
Chest	4
Church	13
Cider	10
Cigarette	2.89
Clock	2.86
Cloud	2.6
Clown	4.17
Cork	2.5
Cowboy	1
Crab	1.11
Crane	10
Crown	2.17
Curtain	2
Cymbals	1.63
Daisy	1.23
Dentist	1.38
Diamond	1.44

Table 3.1

*Verb Generation Pilot: List of Objects Included with Response Strength Ratios*

	<u>Response Strength Ratio</u>
Dice	6.8
Door	3.25
Drum	3.29
Duck	2.17
Eagle	2.33
Ear	2.07
Elbow	17.5
Envelope	1.1
Feet	1.73
Fence	1.2
Fire	8
Firewood	7
Fridge	1.21
Frog	1
Furnace	2.67
Ghost	5.75
Gun	3.11
Hammer	1.55
Hawk	5.25
Head	2.71
Heart	1.43
Hinge	2.67
Horn	6.6
Horse	1.08
House	7.33
Ice	1.21
Kennel	1
Kettle	19.5
Key	1.5
Kitten	2.22
Knife	4.25
Knight	1
Ladder	40
Ladle	1.5
Lamp	4.6
Lantern	4.83
Leg	1.36
Limousine	4.33
Lion	39

Table 3.1

*Verb Generation Pilot: List of Objects Included with  
Response Strength Ratios*

	<u>Response Strength Ratio</u>
Lips	12
Mallet	1.78
Match	2.89
Microscope	1.5
Mixer	5
Moon	2.57
Mop	6.2
Mosquito	7.75
Mouse	3.75
Mussel	1.4
Needle	1.83
Nose	3.44
Nun	41
Ornament	1.22
Package	1.14
Pan	2.78
Pedal	1.33
Pencil	3.78
Penny	2
Pepper	1.29
Piano	40
Picture	5.2
Pill	2.5
Pillar	1.21
Pillow	3
Pipe	20
Plug	1
Pool	13.5
Pram	4.33
Priest	17.5
Puzzle	0.5
Queen	2.11
Rabbit	3.25
Radio	1.64
Rain	2.33
Razor	1.8
Road	1.14
Rocket	1.4
Rope	5.2

Table 3.1

*Verb Generation Pilot: List of Objects Included with Response Strength Ratios*

	Response Strength Ratio
Ruler	4.43
Scales	6.4
Scissors	100 % response agreement
Seesaw	1.8
Shark	1.21
Shovel	100 % response agreement
Skirt	3.17
Sleigh	2.3
Slide	9
Snow	6.5
Soldier	1.15
Sparrow	1.5
Spoon	1.33
Statue	2
Stethoscope	6.4
Stick	1.67
Stoat	1
Stomach	5.4
Stool	3.88
Stove	19.5
Straw	6.4
Sugar	6.75
Suitcase	1.5
Sun	9.67
Sword	2.33
Syringe	1.42
Table	1
Tail	2.17
Teeth	2.5
Telephone	18
Tent	1.09
Thistle	1.09
Thread	4.25
Tiger	2.5
Tongue	1.44
Tornado	3.17
Towel	40
Tractor	3.33
Trapeze	5

Table 3.1

*Verb Generation Pilot: List of Objects Included with Response Strength Ratios*

	<u>Response Strength Ratio</u>
Tree	3.17
Trophy	1
Trout	13
Tweezers	4.17
Typewriter	2.45
Van	10
Volcano	3.75
Wallet	1.3
Watch	1.86
Weed	4.75
Wheelbarrow	1.14
Wheelchair	1.09
Whistle	18
Windmill	2.86
Wool	11
Worm	2.8
Yacht	18.5

## Verb Generation Task

### *Method*

#### *Participants*

Participants included: seven non-fluent aphasics (CT, DD, DA, ECV, JHM, TB, and WS); six fluent aphasics (FS, NP, NS, PS, STR, and XX); and thirteen neurologically intact older controls (AK, AP, AR, BK, EK, EM, GA, IH, KR, PP, RS, SR, and TK). Further details on the participants are available in the category exemplar task participant section (see pp. 49-64).

## Materials

From the 200 objects included in the verb generation pilot task, 114 were selected for use in the verb generation task proper. Following the procedure set out by Thompson-Schill et al., (1997) and Thompson-Schill et al., (1998), response strength ratios were used to divide the objects into two categories. Response strength ratios of 3.0 or less were taken to indicate low response strength, as they show that there was no strong tendency to choose the most common response over the second most common response. Objects with low response strength ratios were assigned to our high competition condition. Response strength ratios of 5.0 or more were taken to indicate high response strength, because they show that there was a strong tendency for the most common response to be chosen over the second most common response. Objects with high response strength ratios were assigned to our low competition condition.

Of the piloted objects, 107 met the high competition condition response strength ratio criteria and 57 met the low competition condition criteria. Further objects were deleted from both groups if they met one or more of the following criteria: 1) The object name was repeated in the response (e.g., *Mop*: “Mop”); 2) there was a tendency for responses to take the form of particle verbs, or; 3) two or more errors were made for that object. This left 45 objects in the low competition group. Further objects were deleted from the high competition group if the relevant frequency value was not an approximate match to a frequency value from the low competition group until the high competition group also contained 45 objects. The remaining 90 objects were selected for use in the verb generation task proper. A full list of the selected objects is included in Appendix B2.

Two checks were completed to ensure that, apart from the inherent response strength differences, the selected object groups were equivalent. First, a frequency rating for each object was obtained from the Max Planck Institute for Psycholinguistics' Celex Database (2001)<sup>8</sup>. An independent samples t-test confirmed that the high and the low competition condition objects were not significantly different in terms of word frequency ( $t(84) = 0.042, p > 0.05$ )<sup>9</sup>. Also, a word length measure was operationalised by counting the number of syllables in each word. An independent samples t-test confirmed that the high

---

<sup>8</sup> The Celex Database is available online at <http://www.mpi.nl/world/celex>.

<sup>9</sup> Frequency data was not available for two low response strength items (Feet, Pills) and two high response strength items (Lips, Scales).



and the low competition condition objects were not significantly different in terms of word length ( $t(88) = 1.224, p > 0.05$ ).

### *Procedure*

Participants were tested individually. Testing took place at participants' homes, in a quiet testing room at Victoria University, or in a quiet testing room at Temple University. Audio recordings were made of each session, and were uploaded onto computer for further processing. PsyScope software was used to present the experiment on a Mac laptop computer.

During each testing session, participants were instructed that they would be shown some pictures, and that they should say what the objects do or what we do with the objects. The instructions were given in a written format, plus the experimenter read the instructions to the participant, and explained the task in her own words, before commencing with several practice objects (these were selected from the discarded pilot stimuli, and the same objects were used for each administration of the task). The experimental objects were presented in a fixed pseudo-random order

The aphasic participants completed the task twice. The second administration of the task was completed at least two weeks after the first.

### *Data Analysis*

Each initial response was coded as correct or as an error. The only exceptions to this rule were: 1) if the first utterance was clearly not a response attempt (e.g., fillers such as "Er..." or "Um...", comments or asides, or repetitions of the noun prior to giving a verb response); or 2) if the first utterance was a fragment of a word that was subsequently produced in its entirety (e.g. "/klaI/ climb"). A response was considered correct if it consisted of a verb that was: 1) appropriate to the noun; and 2) specific to the noun (e.g., *ladder*, "Climb" was accepted, but *ladder* "Use" was not). Inflectional forms of the verbs were permitted (e.g., "Climb", "Climbed", or "Climbing"). However, auxiliary verbs (e.g., can, have, be, will, may) were not accepted as correct and nor were any responses given after a phonological or semantic cue. For correct responses, response latencies were measured by timing from the onset of the auditory stimulus until the onset of the correct response.

Response accuracy and response latency data were analysed at the group and individual levels. In order to examine whether group accuracy was differentially influenced by competition condition, a repeated measures logistic regression analysis was performed with participant and object included as repeated measures. In order to examine whether group latencies were differentially influenced by competition condition, a general linear mixed model was performed with participant and object included as repeated measures. Both group analyses included the following independent variables: group, competition condition, and a group by competition condition interaction. When group and the group by competition condition interaction were not significant, they were removed from the analyses.

For the aphasic participants, response accuracy and response latency were also analysed at the individual level in order to examine whether there were individual differences within the groups. The individual accuracy analyses also included a repeated measures logistic regression, and the individual latency analyses also included a general linear mixed model (item was included as a repeated measure in these analyses). The individual analyses included the following independent variables: competition condition, session, and a competition condition by session interaction. When the competition condition and session interaction, or the session variable, were not significant, they were removed from the analysis.

## *Results*

An alpha level of .05 was used for all statistical tests.

*Response Accuracy.* The number of correct responses made by each group is detailed in Table 3.2. Error types included phonological paraphasias, failures to respond, perseverations of previous responses, and non-verb responses. Across groups there was a general trend for more accurate responding in the low competition condition, although there were some exceptions to this trend. There was a highly significant main effect of group ( $\chi^2(2) = 181.60, p < .01$ ): The controls were more accurate than the non-fluents ( $\chi^2(1) = 131.29, p < .01$ ) and the fluents ( $\chi^2(1) = 94.46, p < .01$ ), but there was no significant difference between the aphasic groups ( $\chi^2(1) = 0.10, p = .76$ ). There was also a highly significant main effect of competition condition ( $\chi^2(1) = 7.29, p < .01$ ), with more accurate

responding in the low competition condition. However, there was a non-significant group by competition condition interaction ( $\chi^2(2) = 0.16, p < .92$ ).

Table 3.2  
*Verb Generation Task: Number of Correct Responses by Participant and Competition Condition, with Repeated Measures Logistic Regression Statistics*

	Low Competition			High Competition			Percentage Change	$\chi^2$	df	p
	Session 1	Session 2	Mean	Session 1	Session 2	Mean				
Non-Fluents										
CT	23	28	<b>26</b>	28	26	<b>27</b>	-6	0.01	1	0.93
DA	22	21	<b>22</b>	16	24	<b>20</b>	-8	0.01	1	0.90
DD	25	29	<b>27</b>	32	27	<b>30</b>	8	0.17	1	0.68
ECV	24	30	<b>27</b>	23	19	<b>21</b>	29	3.16	1	0.08
JHM	42	39	<b>41</b>	33	32	<b>33</b>	25	7.72	1	0.01
TB	35	35	<b>35</b>	34	31	<b>33</b>	8	0.59	1	0.44
WS	9	12	<b>11</b>	8	10	<b>9</b>	17	0.29	1	0.59
Group Mean			<b>27</b>			<b>26</b>	10			
Fluents										
FS	18	30	<b>24</b>	20	30	<b>25</b>	-4	0.02	1	0.89
NP	39	33	<b>36</b>	42	26	<b>34</b>	6	0.93	1	0.34
NS	18	13	<b>16</b>	15	14	<b>15</b>	7	0.15	1	0.70
PS	34	40	<b>37</b>	32	37	<b>35</b>	9	0.82	1	0.36
STR	37	38	<b>38</b>	34	36	<b>35</b>	7	2.19	1	0.14
XX	21	29	<b>25</b>	18	26	<b>22</b>	14	0.60	1	0.44
Group Mean			<b>28</b>			<b>29</b>	7			
Controls										
Group Mean	37	-	<b>37</b>	36	-	<b>36</b>	5			

The number of correct responses made by each aphasic participant is also detailed in Table 3.2. Unsurprisingly, considering the heterogeneity within the participant groups, the correct responding rates were highly variable. However, most participants were more accurate in the low competition condition. Non-fluent JHM was significantly more accurate in the low competition condition than the high competition condition. None of the other aphasics showed a significant competition condition effect, although non-fluent ECV showed a trend that approached significance ( $p = .08$ ) (see Table 3.2).

Three of the aphasic participants showed significant main effects of session. Fluent aphasics FS and XX were significantly less accurate during session one than session two (FS ( $\chi^2(1) = 14.17, p < .01$ ), XX ( $\chi^2(1) = 8.53, p < .01$ ), fluent PS showed a trend towards less accurate performance during session one than session two ( $p = .06$ ), and fluent NP was significantly *more* accurate during session one than session two ( $\chi^2(1) = 16.72, p < .01$ ). None of the other aphasics showed significant session effects ( $p > .05$ ).

None of the aphasics showed a significant session by competition condition interaction, which suggests that any accuracy differences across sessions did not influence performance across competition conditions.

*Response Latencies.* Geometric mean response times (RTs) and geometric coefficients of variance were calculated for each group and competition condition (see Table 3.3). Because the response latency distribution was highly skewed towards shorter responses, the RTs were log transformed prior to analysis. Overall, latencies from the low competition condition were significantly shorter than latencies from the high competition condition ( $F(1, 2222) = 13.84, p < .01$ ). There was also a highly significant main effect of group ( $F(2, 2222) = 50.54, p < .01$ ): The non-fluents had highly significantly longer RTs than the fluents ( $F(1, 2222) = 7.21, p < .01$ ) and the fluents had highly significantly longer RTs than the controls ( $F(1, 2222) = 37.31, p < .01$ ). However, the predicted interaction of group (non-fluents vs. controls/fluents) and competition condition was not significant ( $F(2, 2222) = 0.94, p = .39$ ). Further, there was no significant competition condition interaction when the non-fluents and fluents were compared without the controls ( $F(1, 2222) = 1.87, p < .17$ ).

Table 3.3

*Verb Generation Task: Geometric Mean Response Latencies (msec) by Participant and Competition Condition, Collapsed Across Session, with General Linear Model Statistics*

Statistics	Geometric Mean Response Latency (Geometric Coefficient of Variance)			<i>F</i>	<i>df</i>	<i>p</i>
	Low Competition	High Competition	Percentage Change			
Non-Fluents						
CT	3904 (117)	6073 (110)	56	7.45	1, 67	< .01
DA	5377 (108)	8985 (120)	67	10.03	1, 73	< .01
DD	3729 (127)	4168 (126)	12	0.34	1, 57	0.56
ECV	2824 (117)	3359 (124)	19	0.48	1, 68	0.49
JHM	3434 (111)	4092 (122)	19	2.3	1, 83	0.13
TB	3233 (50)	3494 (81)	8	0.06	1, 78	0.81
WS	3835 (98)	5060 (102)	32	4.13	1, 28	0.05
Group Geometric Mean	3384 (111)	4271 (128)	26			
Fluents						
FS	2805 (78)	2815 (89)	0	0	1, 64	0.95
NP	2345 (79)	3106 (77)	28	2.86	1, 82	0.09
NS	3240 (79)	4321 (91)	33	3.66	1, 44	0.06
PS	1868 (121)	2400 (97)	27	2.29	1, 84	0.13
STR	2524 (77)	2610 (68)	3	0.6	1, 87	0.44
XX	3565 (117)	4514 (130)	32	4.12	1, 60	0.05
Group Geometric Mean	2646 (109)	3264 (110)	23			
Controls						
Group Geometric Mean	1284 (63)	1606 (66)	25			

Geometric mean response latencies and geometric coefficients of variance for each individual aphasic participant and competition condition are shown in Table 3.3, along with statistical analysis results. There was a strong tendency for individual participant RTs to reflect the overall trend towards faster responding in the low competition condition. Three non-fluent aphasics (DA, CT, WS) and one fluent aphasic (XX) showed significantly shorter RTs in the low competition condition than the high competition condition. None of the other aphasics showed a significant competition condition effect, although two fluent aphasics (NS, NP) showed a trend towards faster RTs in the low competition condition that approached significance ( $p > .05$  but  $p < .1$ ).

Main effects of session were also analysed, but no clear trend emerged. The geometric mean RT was calculated for each aphasic participant and session. The following participants had significantly faster geometric mean RTs during the second testing session than the first: Non-fluents CT ( $F(1, 45) = 13.53, p < .01$ ), ECV ( $F(1, 25) = 15.46, p < .01$ ), and JHM ( $F(1, 60) = 7.74, p < .01$ ). The following participants had significantly slower geometric mean RTs during the second testing session than the first: Non-fluents DD ( $F(1, 23) = 9.24, p < .01$ ), and TB ( $F(1, 54) = 6.28, p < .05$ ). The following participants showed a trend towards slower responses in the second session: Fluents NP ( $p = .08$ ), and NS ( $p = .08$ ). The remaining participants showed no significant session effects: Fluents FS, PS, STR, and XX ( $p > .05$ ); and non-fluents DA and WS ( $p > .05$ ).

CT was the only participant to show a significant competition condition by session interaction ( $F(1, 34) = 4.98, p < .05$ ), with faster RTs and a smaller difference between the conditions on the second session. It is unclear why CT showed this decrease in response latency difference during the second experimental session, but it is possible that it may be partly due to decreased RTs overall (for more on this idea, see the supplementary frequency analysis section below).

These results suggest that some aphasics perform better in the low competition condition than the high competition condition. Sometimes, as in the case of JHM, this improvement is shown in increased response accuracy, without response latency differences. Sometimes this pattern is reversed, as in the case of DA, who shows highly significant differences in response latency, without a response accuracy effect. Often however, the difference is shown in both response accuracy and response latency, but to a smaller degree in each. Often these smaller differences do not meet the significance level of .05, but if you consider the accuracy and latency data together, the strong trends for improvements in both measures suggest that some reliable differences may be present. These differences in performance may be glossed over if each analysis is considered alone. Therefore, we should consider the possibility that some important trends in the data may be overlooked by the separate accuracy and latency analyses.

### *Supplementary Analyses*

*Frequency Analyses.* At best, the difference between the performance of the non-fluent and the fluent aphasics on this task can be considered one of degree, in that some non-fluent aphasics showed exceptionally large differences between the competition

conditions; some fluent aphasics showed these same trends, but the size of their differences was smaller. This trend caused a problem when interpreting the data because the largest differences between the conditions were shown by the participants who had the longest RTs overall. It is possible that the increased difference was a by-product of the increased time taken to respond (i.e., if some of the fluents had taken longer to make their responses, they may also have shown longer differences between their performances in the two competition conditions). If this was the case, you might presume that the differences would be in proportion to the overall increase in RT if the competition manipulation was having no additional influence. This does not appear to be the case, since the proportion of change increases quite markedly (see Table 3.3). However, it remains possible that those participants showing long RTs may be particularly susceptible to slight increases in difficulty, especially since their exceptionally long RTs suggest that they are already finding the task more difficult than others.

To investigate whether the larger difference shown by some participants was a by-product of them being more sensitive to increased difficulty rather than competition differences, the items were assigned to frequency conditions: The high frequency condition included the objects with the highest frequency ratings, and the low frequency condition included the objects with the lowest frequency ratings (see Appendix B3 for a list of items in the high and the low frequency conditions). It was expected that the high frequency items would be easier than the low frequency items (e.g. see Arpita, van Lieshout, & Square, 2007). If the exaggerated competition effects seen in the non-fluent aphasics were driven largely by overall RTs, then we would expect that those participants would also show exaggerated frequency condition effects. If, however, the effects were more specific, then we would not expect the non-fluent aphasics that showed exaggerated competition effects to be equally influenced by the frequency manipulation, even though the low frequency items would be more difficult.

*Response Accuracy.* Figure 3.2 depicts the accuracy data for each aphasic participant by frequency condition and competition condition. Generally the difference between the frequency conditions was smaller than that between the competition conditions. Also, frequency condition did not predict correct responding as clearly as competition condition: In the frequency condition analysis, four participants showed the anticipated response pattern, making more correct responses in the high frequency condition (non-fluents DD, ECV, and SA, and fluent STR). In contrast, in the competition condition analysis, 10 participants show the anticipated response pattern, making more



correct responses in the low competition condition (non-fluents DD, ECV, JHM, TB, and WS, and fluents NP, NS, PS, STR, and XX).

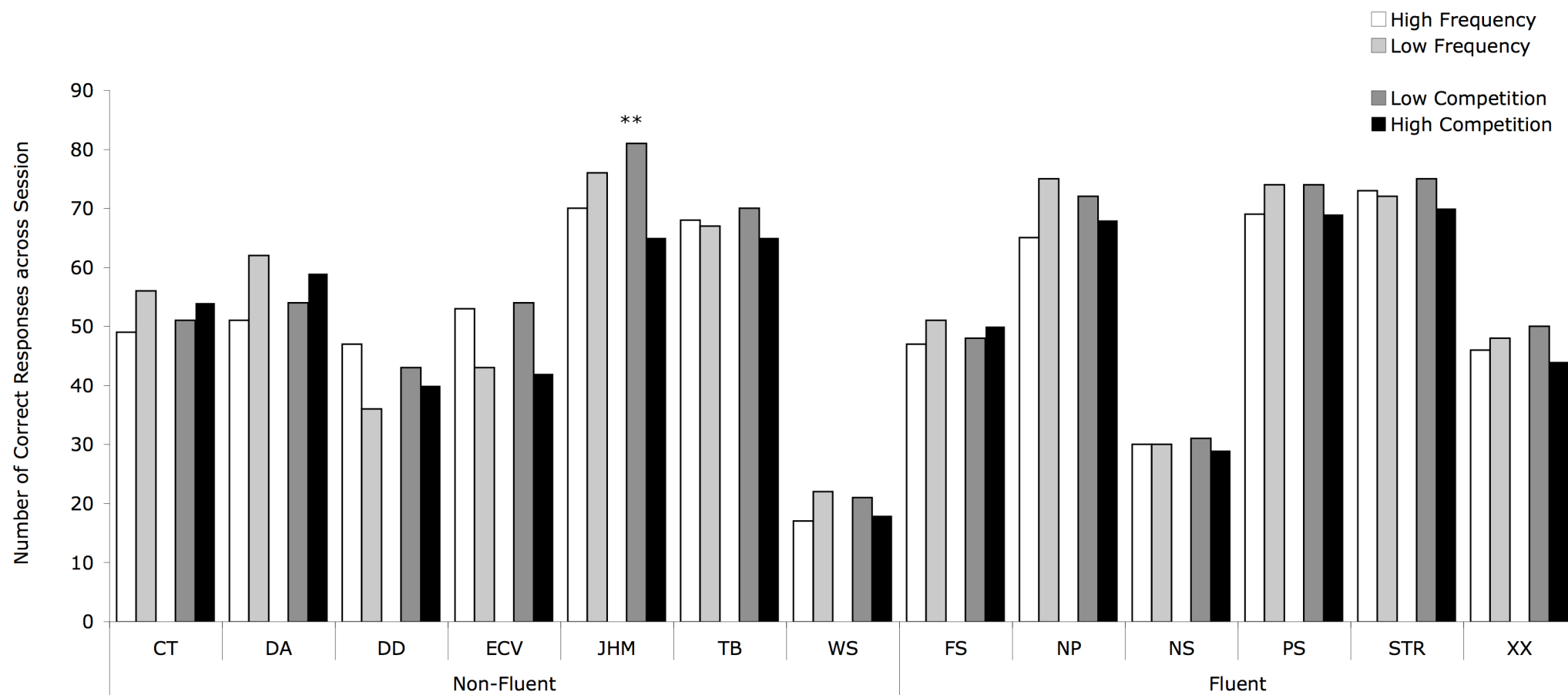


Figure 3.2. Verb generation task: Number of correct responses by participant, frequency condition, and competition condition.

A repeated measures response logistic regression analysis was completed on the data to investigate whether the frequency manipulation influenced responding in the same way as the competition manipulation (i.e., to investigate whether the participants showing the longest RTs overall would also show the largest differences between their performance on the frequency conditions). The following variables were included: frequency condition, competition condition, session, and two and three way interactions between the aforementioned variables (item was included as a repeated measure). The interactions and the session variable were removed from the model when they were non-significant. Because the frequency variable was the main focus of this analysis, it is the primary focus of this results section, and any significant results from this analysis that do not pertain to the frequency variable are detailed in Appendix B4.

Unexpectedly, nine of the aphasic participants made more correct responses in the low frequency condition than the high frequency condition. However, none of the participants showed a significant main effect of frequency condition, although non-fluent DA did show a trend towards a reverse frequency main effect, with less accurate responding in the high frequency condition ( $p = .08$ ). Several participants did show a borderline or significant interaction: Non-fluent ECV showed a significant frequency by competition condition interaction ( $\chi^2(1) = 6.03, p < .05$ ). Non-fluent DA showed a borderline frequency by competition interaction ( $\chi^2(1) = 3.82, p = .051$ ), and fluent NS showed a trend towards a frequency by competition interaction ( $p = .08$ ). Non-fluent WS showed a significant frequency by session interaction ( $\chi^2(1) = 4.30, p < .05$ ). Finally, non-fluents DD and WS showed borderline three way interactions of competition, frequency, and session (DD ( $\chi^2(1) = 3.75, p = .053$ ), WS ( $\chi^2(1) = 3.76, p = .053$ )).

As you can see in Figure 3.2, JHM, the only participant who showed a significant competition condition accuracy effect, did not show a significant frequency condition accuracy effect. This does not support the view that JHM's significant competition condition accuracy effect occurred because the high competition condition was more difficult than the low competition condition.

The unexpected insignificant trends towards reverse frequency effects observed in the accuracy data from nine aphasic participants possibly occurred because the frequency values in the low and high frequency conditions were not different enough to influence accuracy. However, in contrast, the expected trend towards more accurate responding did occur in the response latency analyses for nine of the aphasic participants (see pp. 98-100).

*Response Latency.* In a post-hoc response latency analysis, geometric means were calculated for each aphasic participant by frequency condition and competition condition (see Figure 3.3). Because the response latency distribution was highly skewed towards shorter responses, the RTs were log transformed prior to analysis. Next, a general linear mixed model was completed (with item as a repeated measure). The following independent variables were included: frequency condition, competition condition, session, and two and three way interactions between the aforementioned variables (when they were not significant, the interactions and the session variable were removed from the model). Again, because the frequency manipulation is the main focus of this analysis, it is the main focus of the following result section, and any significant results from this analysis that do not pertain to the frequency variable are detailed in Appendix B4.

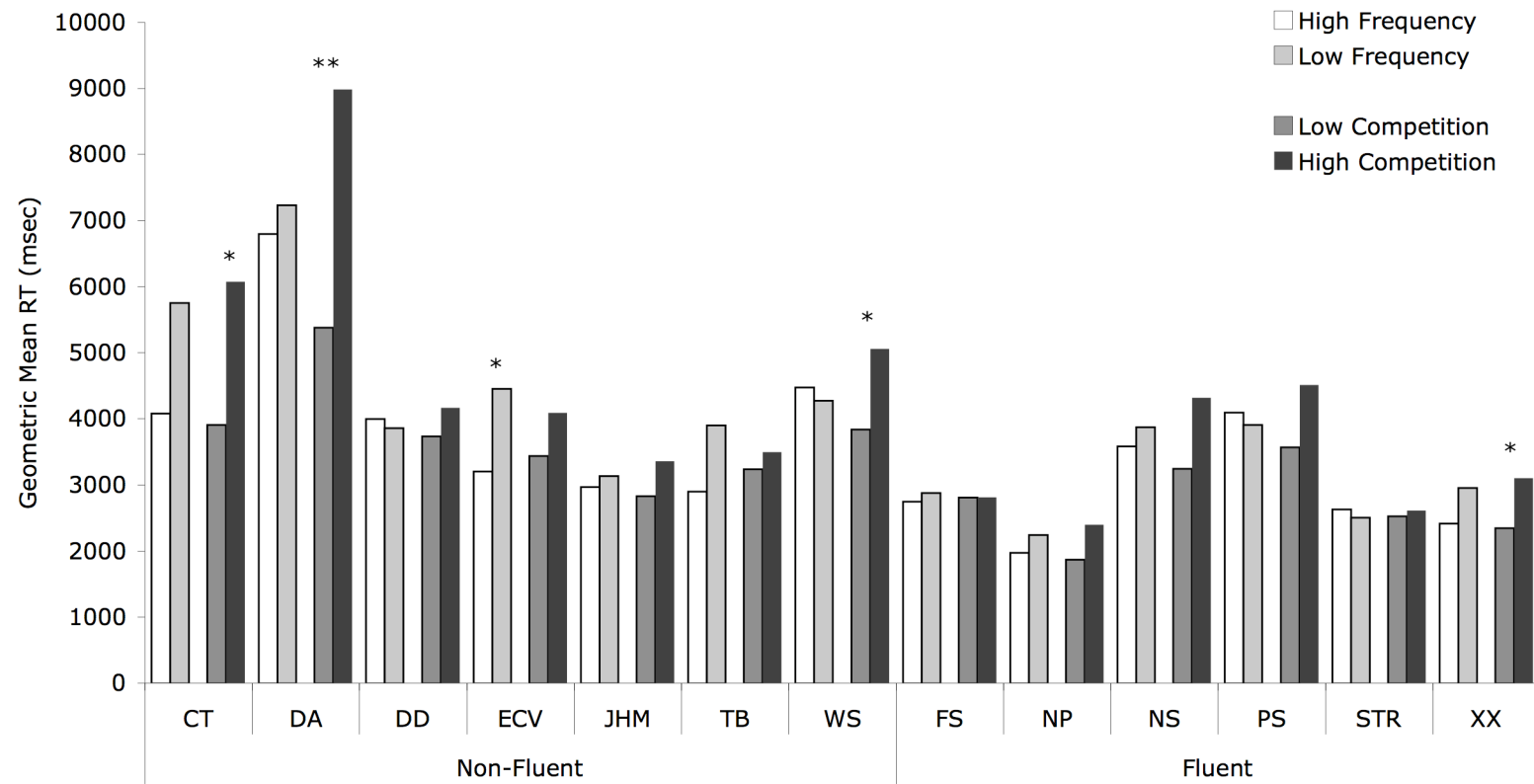


Figure 3.3. Verb generation task: Geometric mean response latencies by participant, frequency condition, and competition condition.

Non-fluent ECV showed a highly significant main effect of frequency, with faster responses in the high frequency condition ( $F(1, 68) = 7.16, p < .01$ ). Non-fluent CT showed a trend towards faster responses in the high frequency condition ( $p = .09$ ). Non-fluent WS showed a frequency by competition interaction ( $F(1, 26) = 5.54, p < .05$ ). Finally, fluent FS showed a trend towards a frequency by competition interaction ( $p = .10$ ).

As you can see in Figure 3.3, none of the participants who showed a significant frequency condition effect also showed a significant competition condition effect in their latency data, and vice versa. Hence, in general, these latency results also negate the possibility that significant competition condition latency effects may occur as a result of the high competition condition being more difficult than the low competition condition. However, ECV's significant frequency effect may be contributing to the trend towards a competition condition effect observed in her latency data. Similarly, CT showed a trend towards a significant frequency effect, which may be contributing towards her significant competition condition effect. This issue is discussed in more detail in the Discussion section.

*Association Strength Analyses.* In the Introduction, it was mentioned that the implications of results from verb generation tasks were being debated, and that two main views have been presented. Thompson-Schill and colleagues (Thompson-Schill et al. 1997; Thompson-Schill et al. 1998) presented the response strength view which suggests that competition between co-activated alternatives results in the longer RTs / lower accuracy observed in low response strength conditions. Martin and Cheng (2006) presented the association strength view which suggests that a controlled retrieval process is required in the low response strength conditions because low response strength is correlated with weak association strength, and the weaker associations between the stimuli and possible responses result in the longer RTs / lower accuracy observed in low response strength conditions.

In order to see whether association strength or response strength more accurately accounts for the differences we observed, we followed logic presented by Martin and Cheng (2006), and reassigned a selection of our stimuli to three conditions: 1) low competition; 2) high competition/high association strength; and 3) high competition/low association strength. The low competition condition contained 15 high response strength items. The high competition/high association strength condition contained 15 low response strength items with high association strength ratings (not the highest 5). Lastly, the high

competition/low association strength condition contained 15 low response strength items with low association strength ratings (not the lowest 5). A list of the items included in each of the conditions is available in Appendix B6. The mean, maximum, and minimum response strength ratios and association strength ratios for each of the conditions are depicted in Table 3.4. An independent samples t-test was completed and confirmed that the response strength differences between the two high competition conditions were not significant ( $t(28) = 0.3682, p > .05$ ).

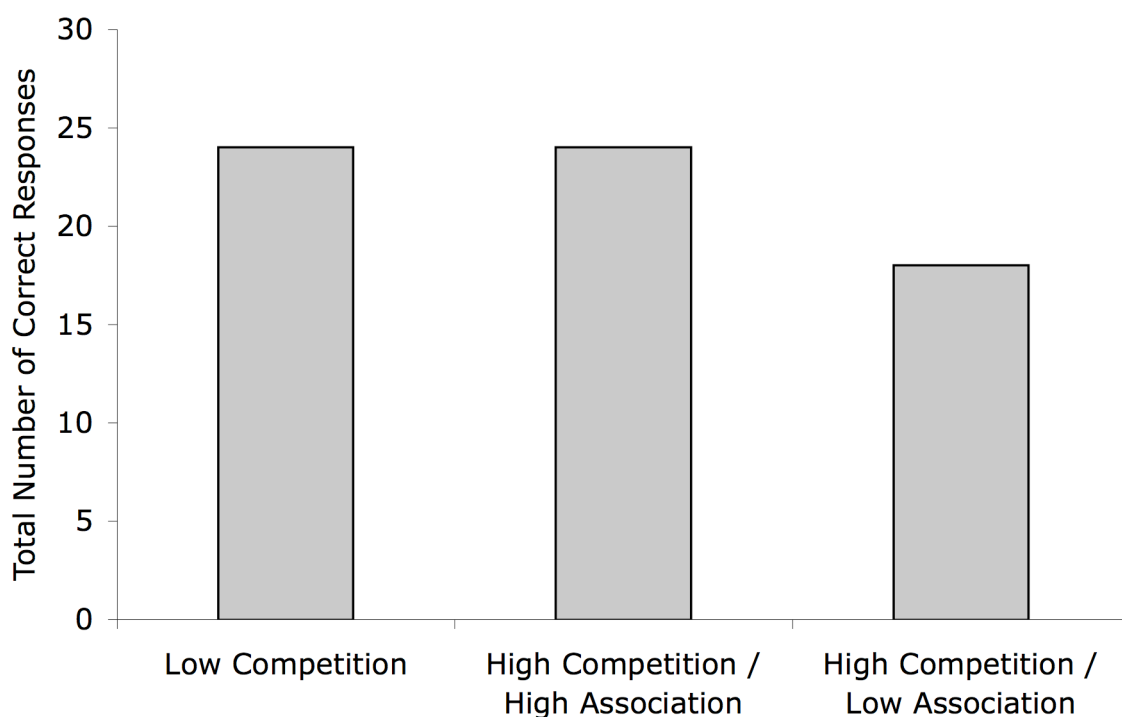
Table 3.4

*Verb Generation Task: Minimum, Maximum, and Mean Response Strength Ratios and Association Strength Ratios, by Association Strength Condition*

	Response Strength Ratio			Association Strength Ratio		
	Minimum	Maximum	Mean	Minimum	Maximum	Mean
Low Competition	5.00	13.50	8.00	0.49	0.74	0.64
High Competition, High Association Strength	1.05	2.89	1.97	0.44	0.67	0.54
High Competition, Low Association Strength	1.08	2.80	1.57	0.14	0.35	0.27

It was hypothesised that if the data dissociated between the high competition/high association and the high competition/low association conditions, rather than along the lines of competition (i.e. low competition dissociating from both high competition/high association and high competition/low association conditions) the data would support the association strength/controlled retrieval view. Because the critical difference in this analysis is between the two high competition conditions, these two conditions were compared. The data from those participants who showed significant competition effects in the main accuracy and response latency analyses was re-examined in this analysis.

JHM's average accuracy by condition is detailed in Figure 3.4. JHM's results appear to dissociate along the association strength division more than the competition division. A binomial logistic regression analysis was used to compare her accuracy scores, and she showed a strong trend towards faster responses in the high competition/high association condition than the high competition/low association condition ( $\chi^2(1) = 2.90, p = .08$ ).



*Figure 3.4.* Verb generation task: Non-fluent JHM's total accuracy scores across session by association strength condition.

Data from the following four participants was included in a latency analysis: Non-fluents' CT, DA and WS, and fluent XX. Because the RT distributions were skewed towards faster responses, the data was logged prior to analysis. The general trend in the data was that the low competition condition RTs were the fastest, followed by the high competition/high association condition RTs, then the high competition/low association condition RTs. Geometric mean RTs by participant and condition are depicted in Figure 3.5. The data was collapsed across session, and independent samples t-tests were completed on each participant's logged data to investigate whether the difference between the high competition/high association condition and the high competition/low association condition was significant. Non-fluents DA and WS both showed a significant difference (DA ( $t(30) = 2.32, p < .05$ ), WS ( $t(12) = 2.34, p < .05$ )), but both non-fluent CT and fluent XX did not show a significant difference (CT ( $t(31) = 0.02, p > .05$ ), XX ( $t(24) = .10, p > .05$ )).



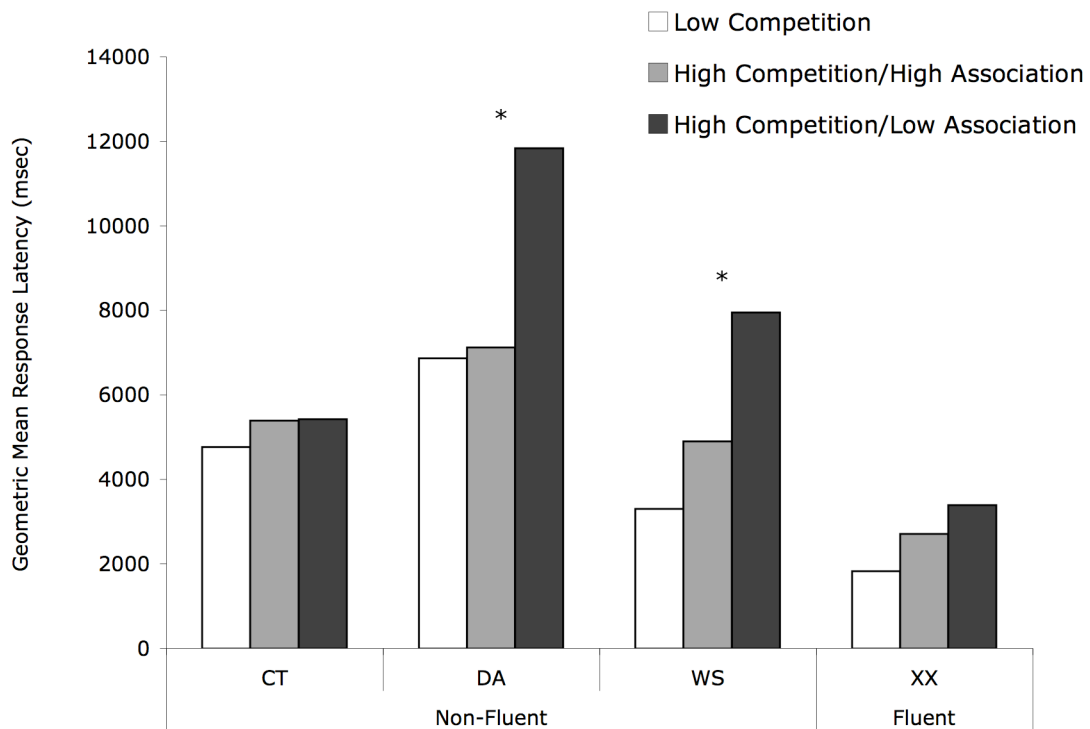


Figure 3.5. Verb generation task: Geometric mean response latencies across session by participant and association strength condition.

These accuracy and latency results suggest that the participants who showed the most convincing competition effects in the original analyses (DA in response latencies, and JHM in response accuracy) appear to be influenced by association strength more-so than competition. These results are consistent with Martin and Cheng's (2006) non-competitive account of the results. However, competition may still be a contributing factor since our high competition conditions showed a response strength ratio difference that was on the cusp of significance, and the analysis is based on few items. However, it seems that the association strength/controlled retrieval view provides the neatest explanation for the results.

### Discussion

This experiment examined verb generation performance of three groups of participants (non-fluent aphasics, fluent aphasics, and older controls) under two competition conditions. In the low competition condition, nouns were presented that consistently elicited one dominant verb (e.g., *Scissors*, “Cut”). In the high competition condition, nouns were presented that elicited a range of verbs (e.g., *Penny*, “Spend”, “Pay”, “Buy”, “Flip”, “Drop”, etc.). The reasoning behind the competition manipulation was that if a participant had a disorder affecting the hypothesised lexical selection and control mechanism, they might have more difficulty in the high competition condition, because this mechanism would be required to select between the alternative responses available.

Overall, the low competition condition was associated with higher accuracy and/or faster response latencies than the high competition condition. This was the case for all three participant groups. In individual analyses, the majority of the aphasic participants showed either significantly more accurate and faster, or a trend towards more accurate and faster, performance in both accuracy and latency in the low competition condition. Each aphasic showed results in the predicted direction on at least one of the two measures (accuracy or latency). Generally though, latency was the more sensitive measure, because many participants who showed no effect in their accuracy nonetheless revealed at least a trend in the expected direction in their latency data.

The primary hypothesis made was that if some, or all, individuals with non-fluent aphasia suffer from a specific deficit involving our hypothesised selection and control mechanism, they should show abnormally exaggerated effects of competition. Hence, we hypothesised that non-fluent aphasics would show exaggerated high competition costs in accuracy and/or latency when compared to fluent aphasics and controls. The data provided very limited support for this hypothesis. There was no significant interaction of group and competition condition in the accuracy or latency analyses. However, the trends in both the accuracy and latency data were in the predicted direction. In the group accuracy analysis and the group latency analysis, the non-fluents had the largest percentage change between the competition conditions, followed by the fluents, then the controls. However, analysis of individual participant’s data revealed considerable heterogeneity within each group. Specifically, taking into account differences in the accuracy and latency analyses, four of the non-fluent aphasics (JHM, DA, CT, and WS), showed significant competition effects, and three (DD, TB, and ECV) did not. In contrast, only one fluent aphasic (XX), showed a significant competition effect, and five (FS, STR, NP, NS, and PS) did not. If we look at

simple prevalence rates, it does appear that non-fluents are more likely than fluents to be significantly influenced by competition manipulations.

The second portion of this prediction was that those non-fluent aphasics with the most marked dissociations between their sentence and word production abilities (in this study: DA, ECV, and JHM) would be the most likely to show exaggerated competition effects. In the accuracy analysis JHM was the only aphasic participant to show a significant high competition cost. In the latency analysis DA showed the predicted significant high competition cost. Further, DA showed a percentage change between the competition conditions that was outside the normal range. However, another non-fluent participant, CT, also demonstrated these effects, even though she was not expected to show exaggerated response strength effects. As a qualifier though, in the supplementary frequency analyses CT showed a trend towards slower RTs in the high competition condition. Although this trend was not significant, it may indicate that the general increase in response latencies observed in the high competition condition could be contributing to the significant competition effect observed in her response latency data. Taking these concerns into account, and adopting a conservative approach, it does appear that at least two of the non-fluents that were predicted to show exaggerated competition effects, DA and JHM, were amongst the three aphasic participants showing the largest differences between the conditions. In conclusion, there is partial, but not complete, support for the hypothesis regarding the most susceptible individuals. However, it is still unclear what specific participant characteristics might determine these kinds of effects, and further research is required before any clear conclusion can be reached.

It is possible to speculate why some non-fluent aphasics who show this dissociation between relatively good single word production and relatively poor sentence production, might be more sensitive to response strength manipulations. One possibility is that, with their mild single word production deficits, these patients have a language profile that doesn't include another deficit that hides high competition costs. Perhaps, in order to exhibit competition effects, there needs to be some degree of successful lexical access. The patients who showed the largest high competition costs tended to be amongst the slower responders in our sample, but not amongst the least accurate. For example, DA had high accuracy rates but was very slow, and JHM had reasonably high accuracy but moderately slow RTs. These reasonably high accuracy rates may be important because they have a direct effect on the amount of statistical power available for the analyses, and increased power results in an increased likelihood of finding a significant effect where

everything else is equal. However, it is important to note that some patients with high accuracy rates did not show significant effects (e.g., non-fluent TB, and fluent STR), so this can only be a partial explanatory factor for the observed data. Nonetheless, it is possible that the other patients have the same underlying deficit, but a lack of statistical power, or some other feature of their aphasia is masking it.

The data from the current study failed to replicate data from Thompson-Schill et al. (1998) in two respects. First, across groups, the error rates in the current study appear higher than those reported by Thompson-Schill et al.. This may be due to use of a stricter scoring criterion; however, this is not clear because the earlier study did not include a clear description of how scoring was completed. This difference does not appear to be a simple consequence of patient selection, because it was also observed for the elderly controls. Another possibility is that the differences are due to the stimuli that were selected for use. It is possible that our items were more variable, our sessions took longer, or participants had less practice before starting experimental trials.

Our failure to observe a group interaction in this study contrasts sharply with the study presented by Thompson-Schill et al. (1998), who found that a LIFG lesion group showed decreased accuracy in their low response strength condition (the equivalent of our high competition condition), but controls did not. In the current study, group by condition interactions were not observed. This may be because our groups were defined using a functional procedure based on broad aphasia subtype, rather than lesion location. It is possible that the lesions of the current group, and the LIFG group selected by Thompson-Schill et al. are quite different. Brain scans of the aphasic participants may clarify this issue, as it would make lesion comparison possible, and this is another potential avenue for future research. Also, a brief examination of the individual patient data reveals that our non-fluent group in particular was extremely heterogeneous, with some individuals exhibiting greatly exaggerated competition condition effects and others exhibiting effects that appear to be within the normal range. It is possible that if all of the non-fluents selected had profiles more similar to JHM and DA a significant group interaction may have been observed. There is a need for clearer theories, so that we are able to predict which patients might show effects on the basis of their performance on diagnostic tests.

Of course, another difference between the present study and that of Thompson-Schill et al. (1998) is that two of our participant groups included people with aphasia. It is likely that the verb generation task places multiple demands on different types of language processing capacities, and therefore may be a sensitive marker for a range of different

types of language-related cognitive deficits. It is possible that any extraneous linguistic processes used during the task may influence the performance of aphasic participants more strongly than non-aphasic participants, and this may influence the magnitude of any group differences that are, or are not, observed.

In this study, unlike some previous ones based on the same paradigm, both accuracy and latency were analysed. Because significant latency differences were more prevalent than significant accuracy differences, these results suggest that the latency analyses are potentially *more* sensitive to competition differences than accuracy analyses. Because some participants showed trends towards corresponding latency and accuracy differences (e.g., NP), some showed only latency effects (e.g., DA), and some just showed only accuracy effects (e.g. JHM), it is important to look at both accuracy and latencies before concluding whether an effect is present or not. It is possible that by focussing entirely on accuracy, past research may have overlooked significant differences in some participants (e.g., Thompson-Schill et al., 1998).

The supplementary association strength analysis was completed to address the recent debate concerning the underlying basis of competition/association strength effects in the verb generation task. The distinction between the association and competition hypotheses is important because it carries implications regarding the nature of frontal modulation of linguistic processing. If the effects are due to the overall lack of availability of verb associates, then a non-competitive controlled retrieval process may be able to explain the results. On the other hand, if there is a significant additional role played by multiple strongly activated alternatives, then this would provide more explicit support for a competition view. Following Martin and Cheng (2006), three conditions were included in the supplementary analysis, allowing for comparison of association strength differences where competition was held constant. The results suggest that the competition effects shown by the two most susceptible participants (non-fluent DA in response latencies, and non-fluent JHM in response accuracy) may be underpinned by association strength rather than competition. That is, the strength of the relationship between the stimuli and the target appeared to be more important than the number of alternative targets available. This result is consistent with those presented by Martin and Cheng (2006) who found that older controls, younger controls, and non-fluent patient ML, all showed dissociations according to association strength, but not competition.

However, a concern that applies to the current study's supplementary association strength analysis and the Martin and Cheng (2006) study, is the small number of items

included in the three item conditions (15 each). However, it is difficult to select larger item groups that meet the selection criteria for these types of studies. Despite this concern, the results of the presented study, combined with those from Martin and Cheng, provide support for the view that the associative strength between the stimuli and possible responses may be primarily responsible for the effects observed in the verb generation task, as Martin and Cheng suggested. Hence, the verb generation task may not be the most appropriate task to use in order to investigate competition effects during lexical retrieval. Further tasks are needed that more explicitly manipulate selection demands without manipulating the associative relationship between the stimuli and required responses.

It is also important to note that the verb generation task has some further design limitations. One such limitation is that, as well as the competition that occurs between potential responses, additional competition may come from the stimulus, and/or non-verb associates (e.g., the noun ‘beard’ may produce competition for the verb ‘shave’ in response to the object ‘chin’) (see Martin & Byrne, 2006). The current study does not attempt to resolve any issues with extraneous competition because it is not clear how that is possible using the verb generation task. However, other tasks which rely less on associative relations between stimuli and responses have the advantage of minimising this kind of extraneous competition, as well as the advantage of avoiding the aforementioned response strength / association strength confound. Hence, several of the other tasks in this series of research may provide more “pure” tests of the selection and control hypothesis (see for example, the name agreement task).

In conclusion, more, but not all, non-fluents with good single word naming relative to poor sentence production appear to show exaggerated competition effects, relative to other non-fluents, fluents, and older controls. However, these effects are very variable within groups of participants, and they may be driven less by competition, and more by the sheer availability of response options. Hence, it is important to explore other paradigms for corroborating evidence, and to look for consistencies in performance within patients across tasks. This approach may help to identify what type or types of competition these affected individuals are most sensitive to.

## Name Agreement Task Introduction

Name agreement, a measure of the consistency with which a particular name is used to describe an object, is another measure that may be useful in investigations of lexical competition and selection. Studies have shown that during picture naming, response times (RTs) to objects with a single name (high name agreement (HNA)) tend to be faster than response times to objects with more than one name (low name agreement (LNA)) (Ellis & Morrison, 1998; Kan, Kable, Van Scoyoc, Chatterjee & Thompson-Schill, 2006; Kan & Thompson-Schill, 2004; Laiacina, Luzzatti, Zonca, Guarnaschilli, & Capitani, 2001; Shatzman & Schiller, 2004; Vitkovitch & Tyrrell, 1995). It is possible that this difference might reflect higher levels of lexical competition and higher lexical selection demands in LNA conditions, and if this is the case, name agreement tasks may be sensitive to deficits of lexical selection and control.

Kan and Thompson-Schill (2004) examined RTs in response to overt naming of line drawings of objects with HNA and LNA<sup>10</sup>, and fMRI activation in response to covert naming of the same objects. The results indicated that RTs were significantly faster in the HNA condition than the LNA condition for both covert and overt responses. When the fMRI images from the HNA and LNA conditions were compared, they showed increased activation in the LIFG during the LNA trials only (see also Kan and Thompson-Schill (2004) where these imaging results were corroborated using black and white photos instead of line drawings).

The behavioural and neural differences observed between naming of HNA and LNA items have sometimes been attributed to selection between competing alternatives (Kan and Thompson-Schill, 2004). Kan and Thompson-Schill also note that it is possible that the LNA condition requires more semantic retrieval than the HNA condition, as people retrieve more information in their search for the best answer. However, they suggest that it is implausible that the increased semantic retrieval is responsible for the differences that are observed, because they found that neural activity in the left temporal lobe was not

---

<sup>10</sup> A baseline condition was also included. In this condition distorted pictures from the HNA and LNA conditions were used. Half of these items included a superimposed rectangle. In the baseline condition participants were asked to indicate whether the rectangle was present or absent.

significantly greater during the LNA condition, as would be expected under conditions of greater semantic retrieval.

Another alternative view comes from the verb generation task literature, where it has been suggested that observed differences possibly reflect a controlled, effortful retrieval process (see Martin & Cheng, 2006; Wagner, Paré-Blagoev, Clark, and Poldrack, 2001). Kan and Thompson-Schill (2004) argue that the name agreement manipulation makes a particularly clear case in support of the selection hypothesis, because both critical conditions involve naming a single picture. They argue that unlike verb generation, naming a single picture involves relatively automatic, data-driven processes and little in the way of more controlled processes. However, this argument rests on the major assumption that picture naming is actually an “automatic” process.

The automatic processing assumption made by Kan and Thompson-Schill (2004) does not address the mechanisms of controlled selection included in the models of language production presented by MacKay (1987) and Dell (1996). Both these models include a mechanism that functions to control lexical selection, so that only a single lexical item is selected for each syntactic position in a sentence. In the MacKay (1987) theory of sequencing, the controlled selection mechanism takes the form of the multiplicative messages from the timing nodes onto the sequence nodes: These give enough of a boost to stimulation levels of sequence nodes for selection of the most highly activated sequence node. Once selected, the sequence node gives the content nodes of the corresponding syntactic type a multiplicative boost of activation: This boost also increases the difference in activation levels between active nodes of the correct syntactic type – a process that leads to selection of the most highly activated content node. In other words, the controlled process directs which class of word to produce, and through that, which word to select (for more details see pp. 19-22). In the Dell (1986) model of sentence production, the controlled selection mechanism takes the form of the signalling activation that flows down to the next level when a lexical node is selected for insertion into the syntactic frame. This gives the downstream phonological nodes that correspond to words of the selected semantic category a boost and an additional competitive advantage over phonological nodes that correspond to words of other syntactic types (for more details see pp. 16-19). In the instance of producing a LNA item (e.g., *Sofa / Couch / Settee*), both models predict that the alternative items’ shared semantic features automatically lead to co-activation of the alternative lemmas at the lexical level via spreading activation. The controlled selection process then ensures that the most highly activated lemma is selected. If the controlled



selection mechanism is impaired, both models predict a failure of some kind at the lexical level during naming of LNA items. However, these controlled selection mechanisms do operate without conscious effort, and can be contrasted to processes that require thoughtful input, such as strategy implementation, hence the relative automaticity of picture naming appears to be a valid point. Is it possible that the name agreement task is sensitive to controlled selection as it is currently described in language models, without placing any demands on the additional top-down modulation that is hypothesised in the selection and control hypothesis? If this is the case, this task might minimise some of the other extraneous, strategic factors that potentially influence performance in other tasks (e.g., verb generation), and could isolate more squarely the task of selecting amongst alternatives.

The selection and control hypothesis allows you to separate linguistic levels of processing more explicitly than the Kan and Thompson-Schill (2004) hypothesis, and emphasises the importance of selection at the lemma level, where the real bottleneck exists. Consider how the semantic and lexical levels of processing contribute to object naming: Stimuli automatically trigger activation of nodes at the semantic level that flows on to the lexical level where related lemmas are activated. It is not until activation reaches the lexical level that nodes corresponding to individual words are activated, and eventually one of the activated nodes is selected. Accordingly, activation at the semantic level will be approximately equivalent in the HNA and LNA conditions, as each picture would activate the related semantic nodes that relate to the given picture. However, at the lexical level, semantic activation would flow down to more lemmas in LNA conditions than HNA conditions, because there are more words that correspond to the given semantic concept. Because more competition at the lexical level is predicted in LNA conditions, a controlled selection process will be required to bias activation towards the best lemma. Only after the best response is activated to a level above its competitors by controlled activation, will the corresponding node be selected. In other words, lemma level competition would cause the LNA cost. In comparison, during the verb generation task, competition at the semantic level is more likely, as the verbs related to an object may be semantically distant (e.g., in response to the object *Boy*, the response could be semantically unrelated, i.e., “Play” or “Study”).

### *The Current Study*

The current study is a picture naming task that manipulates the name agreement of the stimuli to test the predictions of the selection and control hypothesis on non-fluent

aphasics, fluent aphasics, and a neurologically intact control group. According to the selection and control hypothesis, we should observe lower accuracy and/or longer RTs from all three participant groups in a high competition condition made up of LNA items. This is expected because in the high competition condition, the different possible naming alternatives will all be activated via spreading activation, resulting in lexical co-activation. In contrast, in a low competition condition made up of HNA items, only one alternative exists, so co-activation will be minimal. The increase in lexical co-activation in the high competition condition, is expected to trigger a selection and control mechanism that functions to aid lexical retrieval when automatic processes fail to identify one dominant response (this may be the result of co-activation of different alternatives, or low levels of activation overall).

Further, it is hypothesised that some or all non-fluent aphasics will show exaggerated competition condition effects relative to the fluent aphasics and older controls. This is expected to be reflected in a group by competition condition interaction.

The final hypothesis is based on findings that a sub-group of non-fluent aphasics, who show good naming but poor sentence production, are particularly sensitive to manipulations of context in naming tasks (see Freedman, Martin & Biegler, 2004; Schwartz & Hodgson, 2002; Schnur, Schwartz, Brecher, & Hodgson, 2006; Wilshire & McCarthy, 2002; Wilshire, Scott, & Stuart, 2006). We hypothesise that non-fluent aphasics who fit this profile may show stronger competition condition effects than non-fluent aphasics who do not. Three non-fluents in the current sample have relatively good single word naming, scoring above 40 on the BNT, and are expected to show these exaggerated effects (see p. 64 for more details).

The procedure used will be a simple picture naming paradigm, where pictures are presented and participants are asked to name them as quickly as possible. As is detailed in the name agreement pilot section below, we constructed a new set of stimuli, which were piloted on forty first-year students from Victoria University of Wellington. The aim of the pilot study was to obtain statistics that indicated which items had genuine HNA, or genuine LNA due to variations in the target's name rather than LNA as a result of use of incorrect alternatives (e.g., *Turtle/Tortoise*).

#### Name Agreement Pilot

## *Method*

### *Participants*

Forty first-year students at Victoria University of Wellington participated in this study for course credit. All participants spoke English as their first language. Informed consent was obtained from all participants.

### *Materials*

A set of 256 item names was gathered. This set included, but was not limited to, items that previous research has shown to have HNA or LNA (Kan and Thompson-Schill, 2004; N. Martin, personal communication, 2006; Morrison, Chappel & Ellis, 1997; Peterson and Savoy, 1998; Szekely et al. 2004). Because there were a limited amount of normed LNA items available, a number of synonyms that had not, to our knowledge, been normed previously were also included. A colour photo of each item was sourced from the Internet<sup>11</sup>. In a small minority of cases, an arrow or a circle was added to the photo of the item to clarify what part of the picture was to be named. A full list of the names of the items that were included is available in Appendix C1.

### *Procedure*

Participants were tested in groups of one to fifteen. A brief introduction to the research area was given (verbally and in written format), and then the participants were told that they would be shown some pictures on a computer, and that their task was to type in the name of each picture in the space provided. They were also told that they might be able to think of more than one name for some of the pictures but that they should type the first name that came to mind.

PsyScope software was used to present the stimuli and to record responses. First, the task instructions were reiterated in a written format, and then the participants completed the experimental trials at their own pace. The pictures were presented in a pseudo-random order, with each participant viewing the pictures in the same sequence. Each picture appeared in the middle of the screen, with a response cue underneath (e.g., Response:

---

<sup>11</sup> There was no photo for the item “*Unicorn*”, so a clip art picture was used for this item.

\_\_\_\_\_). Participants typed in their responses (which appeared on the screen after the cue as they typed), and pressed the return key to move on to the next trial.

### *Data Analysis*

The following procedures were used to tidy the data: 1) Obvious spelling errors were corrected; 2) Plural differences were disregarded; and 3) Different forms of the same compound name were standardised (e.g., *Glass House* and *Glasshouse* were considered to be the same response, as were *Merry Go Round* and *Merry-Go-Round*).

### *Results*

A response strength ratio was calculated for each item by taking the frequency of the most common response and dividing that by the frequency of the second most common response. Please refer to Appendix C1 for a complete list of the piloted items with their response strength ratios. The response strength ratios were then used to select two sets of pictures for the name agreement study: a low competition group consisting of items with high name agreement, and a high competition group consisting of items with low name agreement (details of the item selection procedure are given in the Materials section for the name agreement task proper (see pp. 116-118).

### *Name Agreement Task Proper*

### *Method*

### *Participants*

Participants included: seven non-fluent aphasics (CT, DD, DA, ECV, JHM, TB, and WS); six fluent aphasics (FS, NP, NS, PS, STR, and XX); and seven neurologically intact older controls (AP, BC, EM, GA, KR, JI, and FC). Further details on the participants are available in the category exemplar task participant section (see pp. 49-64).

## Materials

Two groups of items were selected from the 256 items included in the name agreement pilot task: A low competition group, and a high competition group. The following procedure was used to choose the items. All of the piloted items were ranked by response strength ratio (this was calculated by taking the frequency of the most common response and dividing it by frequency of the second most common response). Sixty-three items had 100% name agreement, and were selected for the low competition condition. For the high competition condition, we tried to select items that had LNA because of multiple genuine naming alternatives (e.g., *Sofa / Couch / Settee*). We tried to avoid items with LNA that occurred due to production of incorrect responses (e.g., *Alligator / Crocodile*), as it has been suggested that the name agreement costs seen for these items may occur during structural recognition, rather than during the word production process (Vitkovitch & Tyrrell, 1995). By avoiding possible differences in structural recognition, the lexical competition should be maximised. Hence, the 63 high competition items were selected by taking the items with the lowest response strength ratios, where the following exclusionary criteria were not met: 1) The low response strength ratio was the result of participants not knowing what the item was, rather than the result of a synonym or superordinate/subordinate use (e.g., *Bee / Wasp*, *Alligator / Crocodile*); 2) There was an obvious problem with the picture stimuli, such as the participants named different parts of the picture (e.g., *Shutter*: “Shutter” / “Window”). The 63 selected high competition items had a mean response strength ratio of 2.632, with a standard deviation of 1.592 (e.g., *Minister / Pastor / Preacher*, *Artist / Painter*). A full list of the items selected for both conditions is included in Appendix C2.

Checks were completed to ensure the equivalence of the high and low competition items in terms of frequency and word length. Frequency data for the most common, and second most common responses to the items was obtained from the Max Planck Institute for Psycholinguistics’ Celex Database (2001)<sup>12</sup>. Table 4.1 details the minimum, maximum and mean frequency values and syllable lengths of the selected items. An independent samples t-test confirmed that the frequency of the most common responses in the high competition group and the low competition group, were not significantly different ( $t(120) = 0.617, p > 0.05$ ). In the high competition group, many of the second most common responses were nearly as frequent as the most common response. An additional

---

<sup>12</sup> The Celex Database is available online at <http://www.mpi.nl/world/celex>.

independent samples t-test was completed, that compared the frequency of the responses in the low competition group, to the frequency of the second most common responses in the high competition group. Again, no significant difference was found ( $t(120) = 0.710, p > 0.05$ ). We also checked that the responses were equivalent in terms of word length, by counting the number of syllables in the stimulus responses. An independent samples t-test found that the number of syllables in the most common responses in the high and low competition groups were not significantly different ( $t(120) = 0.043, p > 0.05$ ). A second independent samples t-test compared the number of syllables in the responses in the low competition group, to the number of syllables in the second most common responses in the high competition group. Again, no significant difference was found ( $t(120) = 0.915, p > 0.05$ ).

Table 4.1  
*Name Agreement Pilot: Minimum, Maximum, and Mean Frequency Values and Word Lengths of Objects Selected for The Name Agreement Task by Competition Condition*

	Low Competition			High Competition		
	Minimum	Maximum	Mean	Minimum	Maximum	Mean
Frequency of Most Common Response	0	386	33	0	403	40
Frequency of Second Most Common Response	n/a	n/a	n/a	0	143	27
Syllable Count of Most Common Response	1	4	2	1	4	2
Syllable Count of Second Most Common Response	n/a	n/a	n/a	1	3	2

### *Procedure*

Participants were tested individually. Testing took place at participants' homes, in a quiet testing room at Victoria University, or in a quiet testing room at Temple University. Audio recordings were made of each session, and were uploaded onto a computer for further processing. PsyScope software was used to present the experiment on a Mac laptop computer.

During each testing session, participants were instructed that they would be shown some pictures, and that they should name the pictures. The instructions were given in a written format, plus the experimenter read the instructions to the participant, and explained the task in her own words. The trials were presented in a fixed pseudo-random order, with the high and low competition items interspersed unpredictably throughout the experiment. An auditory beep was presented simultaneously with each picture so that response latencies could be timed.

### *Data Analysis*

Each initial response was coded as correct or as an error. A response was considered correct if it was a genuine name for the item, or if 10 % or more of the participants in the pilot study offered it. The 10 % rule allowed for use of incorrect gender variants (e.g., *Rooster*: “Hen”) and superordinates and subordinates if 10 % or more of the norm sample used them also. The genuine alternative rule allowed for use of uncommon true alternative words: This meant that alternatives from American and United Kingdom English that did not pass the 10 % criteria from our New Zealand based pilot study were accepted (e.g., *Lollies*: “Candy”, *Cradle*: “Crib”, and *Jacket*: “Parka”). Incorrect use of plurals was allowed, but responses given after a phonological or semantic cue were considered incorrect. The first response given was scored, unless any of the following situations applied: 1) The first utterance was clearly not a response attempt (e.g., fillers, such as “Er...” or “Um...”, or comments or asides, such as “I don’t know”; 2) The first utterance was a fragment of a word that was subsequently produced correctly (e.g. “/ko/ coat”; 3) The patient started by correctly spelling a correct response and immediately followed this with the corresponding correct response, e.g., “B A N A N A banana”; or 4) An adjective occurred before the noun (e.g., “rain jacket”). Response latencies were measured by timing from the onset of the auditory beep stimulus until the onset of the correct response.

Response accuracy and response latency data were analysed at the group and individual levels. In order to examine whether group accuracy was differentially influenced by competition condition, a repeated measures logistic regression analysis was performed, with picture and participant included as repeated measures. In order to examine whether group latencies were differentially influenced by competition condition, a repeated measures general linear mixed model was performed, with picture and participant included as repeated measures. Both group analyses included the following independent variables: group, competition condition, and a group by competition condition interaction. When group and the group by competition condition interaction were not significant, they were removed from the analyses.

For the aphasic participants, response accuracy and response latency were also analysed at the individual level in order to examine whether there were individual differences within the groups. The individual accuracy analyses were performed using a binary logistic regression, and the individual latency analyses were performed using a



general linear mixed model. The individual analyses included the following independent variable: competition condition.

### *Results*

An alpha level of .05 was used for all statistical tests.

Outliers were removed before the data was analysed: For each participant and condition, the arithmetic mean and standard deviation was calculated, and data from any trials with RTs that were two standard deviations or more above the mean were removed.

*Response Accuracy.* The number of correct responses made by each group is detailed in Table 4.2. Error types included phonological paraphasias, failures to respond, and perseverations of previous responses. No further analysis of error types was undertaken because there was no specific target name for the items (e.g., sofa, couch and settee are all correct answers for a single trial): This makes it difficult to code errors such as phonological and semantic paraphasias.

Across groups there was a general trend for more accurate responding in the low competition condition: this resulted in a highly significant low competition condition accuracy advantage ( $\chi^2(1) = 31.92, p < .01$ ). There was also a highly significant main effect of group ( $\chi^2(2) = 212.37, p < .01$ ): The controls were significantly more accurate than the non-fluents ( $\chi^2(2) = 156.56, p < .01$ ) and the fluents ( $\chi^2(2) = 120.63, p < .01$ ), but the aphasic groups were not significantly different ( $\chi^2(2) = 0.88, p < .35$ ). However, the predicted interaction of group and competition condition was not significant ( $\chi^2(2) = 0.23, p > .05$ ) which suggests that although the aphasic groups were more impaired than the controls, they responded similarly to the competition manipulation.

Table 4.2

*Name Agreement Task: Number of Correct Responses by Participant and Competition Condition, with Binary Logistic Regression Statistics*

	Number of Correct Responses (out of 63)					
	Low Competition	High Competition	Percentage Change	<i>Wald</i> $\chi^2$	<i>df</i>	<i>p</i>
Non-Fluents						
CT	37	49	32	5.148	1	0.02
DA	44	50	14	1.494	1	0.22
DD	23	28	22	0.821	1	0.37
ECV	39	51	31	5.429	1	0.02
JHM	52	58	12	2.461	1	0.12
TB	52	51	-2	0.053	1	0.82
WS	22	34	55	4.568	1	0.03
Group Mean	38	46	19			
Fluents						
FS	30	41	37	3.860	1	0.49
NP	47	53	13	1.719	1	0.19
NS	27	42	56	7.061	1	0.01
PS	57	60	5	1.037	1	0.31
STR	43	46	7	0.334	1	0.56
XX	36	40	11	0.530	1	0.47
Group Mean	40	47	18			
Controls						
Group Mean	56	59	5			

The number of correct responses made by each participant is detailed in Table 4.2. The elderly controls' mean correct response rate was 115/126, the non-fluents' mean correct response rate was 84/126, and the fluents' mean correct response rate was 81/126. However, the aphasic participants' correct response rates varied considerably, ranging from 51 to 110, indicating that some participants found the task much easier than others. Binary logistic regression analyses confirmed that of the non-fluent aphasics, WS, CT, and ECV made significantly more correct responses in the low competition condition than the high competition condition, but DA, JHM, DD, and TB showed no significant difference. Of the fluent aphasics, NS made significantly more correct responses in the low competition condition than the high competition condition, but FS, NP, XX, STR and PS showed no significant difference (see Table 4.2). These results suggest that the additional selection demands present in the high competition condition significantly affected naming accuracy

for some participants, but these individuals were not necessarily predictable by broad aphasia subtype.

*Response Latency.* Because the RT distribution was highly skewed towards shorter responses, each RT was log transformed prior to analysis. The geometric means and geometric coefficients of variance from the trimmed data were calculated for each participant and condition (see Table 4.3). Every participant showed a trend towards faster RTs in the low competition condition. Overall, latencies from the low competition condition were significantly shorter than latencies from the high competition condition ( $F(1, 1666) = 67.63, p < .01$ ). There was also a highly significant main effect of group ( $F(2, 1666) = 10.47, p < .01$ ): RTs were not significantly different for the fluent and non-fluent groups ( $F(1, 1666) = 0.41, p > .05$ ), but the non-fluents had highly significantly longer RTs than the controls ( $F(1, 1666) = 13.40, p < .01$ ), as did the fluents ( $F(1, 1666) = 17.30, p < .01$ ). Further, there was a highly significant name agreement by group interaction ( $F(2, 1666) = 6.63, p < .01$ ). Contrasts revealed that this interaction was significant for the non-fluent and the controls when considered on their own ( $F(1, 1666) = 7.15, p < .01$ ), and also for the fluents and the controls when considered on their own ( $F(1, 1666) = 11.16, p < .01$ ). However, contrary to our prediction, this interaction was not significant for the non-fluents and fluents when considered on their own ( $F(1, 1666) = 0.48, p > .05$ ).

Table 4.3  
*Name Agreement Task: Geometric Mean Response Latencies (msec) by Participant and Competition Condition for Correct Responses, with General Linear Model Statistics*

	Geometric Mean Response Latency (Geometric Coefficient of Variation)			<i>F</i>	<i>df</i>	<i>p</i>
	Low Competition	High Competition	Percentage Change			
Non-Fluents						
CT	1297 (54)	1943 (90)	50	11.56	1, 80	< .01
DA	2081 (85)	3317 (94)	59	12.04	1, 88	< .01
DD	1057(47)	1291 (55)	22	2.70	1, 45	0.11
ECV	1426 (67)	2123 (90)	49	9.99	1, 82	< .01
JHM	1270 (49)	1833 (62)	44	16.98	1, 97	< .01
TB	1114 (77)	1495 (99)	34	5.25	1, 95	0.02
WS	1632 (66)	2963 (86)	82	16.13	1, 51	< .01
Group Geometric Mean	1389 (72)	2027 (95)	46			
Fluents						
FS	2087 (57)	2710 (89)	30	3.92	1, 66	0.05
NP	1334 (69)	2107 (80)	58	15.49	1, 89	< .01
NS	1433 (64)	2113 (66)	47	9.15	1, 61	< .01
PS	1440 (38)	2469 (84)	71	35.13	1, 111	< .01
STR	1608 (56)	2369 (86)	47	11.48	1, 86	< .01
XX	1146 (69)	1405 (52)	23	3.28	1, 70	0.07
Group Geometric Mean	1480 (62)	2171 (83)	47			
Controls						
Group Geometric Mean	971 (27)	1267 (45)	30			

At the individual level, six of the seven non-fluent aphasics had significantly faster RTs in the low competition condition (WS, DA, CT, ECV, JHM, and TB). Non-fluent DD did not show a significant low competition condition advantage. Four of the six fluent aphasics had significantly faster RTs in the low competition condition (PS, NP, NS, and STR). Fluent FS showed a borderline effect in the same direction ( $p = .0518$ ). Fluent XX did not show a significant competition effect. These results suggest that the manipulation of competition affected both non-fluent and fluent aphasics. However, not all of the non-

fluent aphasics, or the fluent aphasics, were affected, indicating that there are significant individual differences within these diagnostic categories.

*Supplementary Accuracy-Sensitive Latency Analysis.* The dependent variables in this task are the rate of correct responses, and RTs. Above, each of these variables has been analysed separately. However, it is possible that speed-accuracy tradeoffs have influenced the data. Also, selection deficits may be manifested in either incorrect response rates, or long RTs, or both. Hence, there are advantages in analysing both response latencies and response accuracy rates together. To do this, an accuracy-sensitive latency analysis was completed. In this analysis, the RTs were inversed, and null values (i.e., error/no response trials) were replaced with zeros – the equivalent of an infinite RT. For ease of interpretation, the analysed values were re-inversed, and are depicted in Figure 4.1.

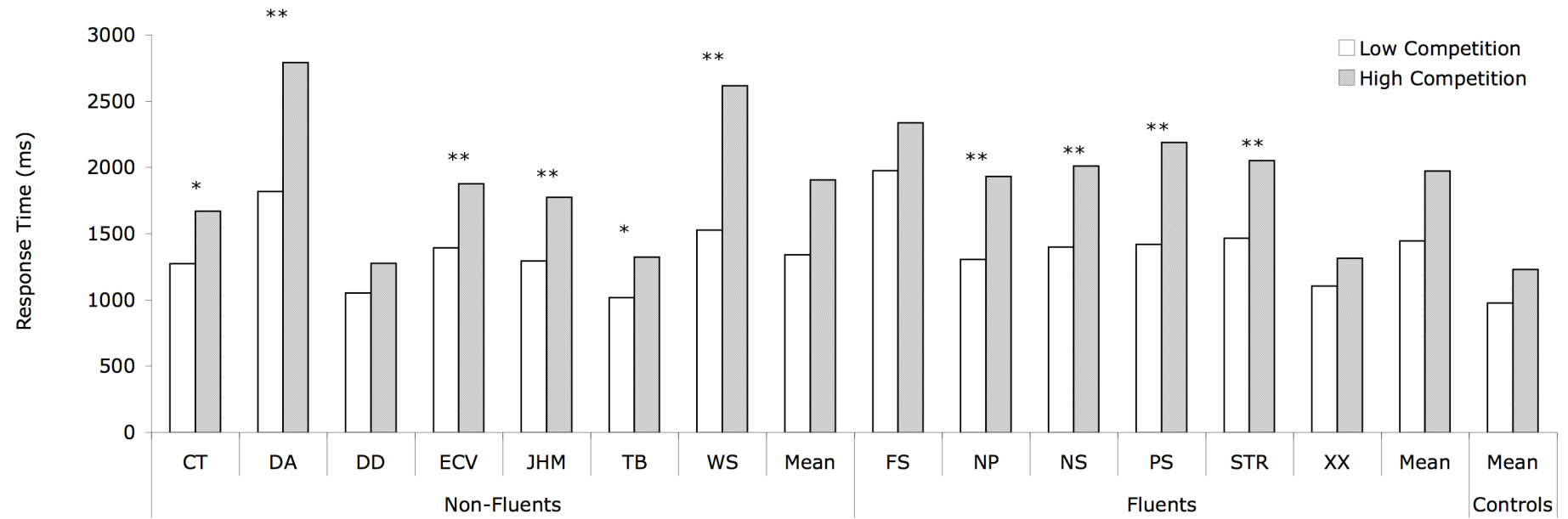


Figure 4.1. Name agreement task: Mean response latencies by participant and competition condition, from the accuracy-sensitive latency analysis.

A general linear model compared each aphasic participant's RTs from each condition. Of the non-fluent aphasics, WS, CT, JHM, DA, ECV and TB had significantly faster RTs in the low competition condition, but DD did not. Of the fluent aphasics, PS, NP, NS, and STR had significantly faster RTs in the low competition condition, but FS and XX did not. These results are the same as those observed in the conventional analysis.

As is depicted in Figure 4.1, two non-fluent aphasics, WS and DA, showed the biggest difference between scores in the high and low competition conditions. These results suggest that some non-fluent aphasics may find it particularly difficult to produce object names when alternative names are available and lexical competition increases, but that this pattern is subject to individual differences.

### *Discussion*

This experiment examined the naming performance of three groups of participants (non-fluent aphasics, fluent aphasics, and controls) for words with low name agreement (LNA)/high competition, and high name agreement (HNA)/low competition. We anticipated that in the high competition condition, relatively strong demands would be placed on a selection and control mechanism that functions to resolve lexical competition. Therefore, more demand would be placed on a lexical selection and control mechanism in the high competition condition than the low competition condition.

The first hypothesis was that the high competition condition would result in relatively higher error rates and longer response latencies across all participant groups. This was expected due to the increase in lexical competition that would occur because the object would automatically activate several alternative lemma level lexical entries. In the low competition condition, it was expected that only one lemma level lexical entry would become highly activated. This hypothesis was supported in both the group accuracy and latency analyses, and is consistent with results from past name agreement studies (see Vitkovitch & Tyrrell, 1995; Ellis & Morrison, 1998; Kan & Thompson-Schill, 2004; Kan et al., 2004; Laiacina et al., 2001).

The second hypothesis was that individuals with non-fluent aphasia should show abnormally exaggerated effects of competition condition relative to fluent aphasics and neurologically intact controls, resulting in larger low competition condition advantages in accuracy and/or latency. This was expected because if the non-fluent aphasics suffer from

a specific deficit involving our hypothesised selection and control mechanism, they should have particular difficulty selecting between the alternative names in the high competition condition. This hypothesis was tested at the group and individual levels. In the grouped accuracy analysis, the data did not support this hypothesis. Nevertheless, in the individual accuracy analyses, three of the seven non-fluents showed a significant low competition advantage (CT, ECV, and WS), but only one of the six fluents did so (NS), suggesting that at least some individuals with non-fluent aphasia may be particularly susceptible to this manipulation. In the grouped latency analysis, there was a significant interaction of group and competition condition; however, this was the result of the non-fluent and fluent aphasic groups showing larger competition effects than the control group. Critically, there was no significant interaction when the two aphasia groups were considered on their own. In the individual latency analyses, there was also no clear support for this hypothesis. Looking at the prevalence of participants showing significant competition effects in the different groups, six of the seven non-fluents showed an effect (not DD). However, four of the six fluents also showed an effect, and another showed a borderline effect (XX did not show a significant or borderline effect). It appears that the group differences that were observed between aphasics and the controls may reflect the aphasics' general language deficits, rather than selection and control deficits specifically, because the size of the effects shown by the fluent and non-fluent groups were of similar magnitude.

Further, across the traditional accuracy and latency analyses there is no clear support for the hypothesis that non-fluent aphasics who show a clear dissociation between good single-word naming and poor sentence production would show exaggerated competition effects. In the accuracy analysis, only one of the three participants that were expected to show exaggerated name agreement effects relative to the other participants did so (ECV). In the latency analysis, the three non-fluents that were predicted to show exaggerated effects showed a percentage increase in RT between the low and high competition conditions that was in the middle of the aphasic range.

The supplementary accuracy sensitive latency analysis was completed to analyse both accuracy and latency together, because it was possible that participants might show differences on one analysis or the other, or on both analyses, and these patterns may have influenced the result patterns observed in the traditional accuracy and latency analyses. In this analysis, of the four non-fluents that showed highly significant effects, three were those expected to show exaggerated effects. However, a total of eight aphasics showed highly significant differences. Of the eight aphasics that showed highly significant effects,



two non-fluent aphasics stood out because they showed larger differences between the competition conditions than the other participants. One of these participants, DA, was predicted to show exaggerated effects because he showed a clear dissociation between good single word naming and poor sentence production, but the other, WS, was not expected to show an exaggerated effect. In summary, those aphasics who showed strong competition effects include, but are not limited to non-fluent aphasics who show a dissociation between good single word production and poor sentence production. It is not clear what characteristics are associated with strong competition condition effects, but it does appear to influence both fluent and non-fluent aphasics.

In conclusion, all three groups were influenced by the competition manipulation, especially the aphasic groups. It is possible that the two groups of aphasics exhibit analogous competition effects for different reasons. Where the non-fluents may be showing strong effects because they have difficulty selecting and controlling lexical activation towards one option, the fluents may have an intact selection and control mechanism that is ineffective because the lexicon itself is damaged. If the fluents have reduced lexical activation levels overall, naming of the high competition condition pictures may be more difficult because the difference in activation levels of the alternative names may be reduced also, making the selection process more difficult, and increasing their reliance on the selection and control mechanism. This hypothesis is speculative at this point however, and needs to be investigated further before any conclusions can be reached.

## Sentence Completion Task Introduction

Another task that may be useful for investigating the processes of lexical competition and selection is sentence completion. In this task, participants are given incomplete speech samples, and are required to fill in the blanks. For example, given the cue sentence “*The death of his dog was a great \_\_\_\_\_*”, participants would be required to come up with a completion such as “shock”, or “tragedy” (see Bloom & Fischler, 1980; Lahar, Tun & Wingfield, 2004). This paradigm might be useful because the sentence stems and/or response requirements can be manipulated to vary the degree of lexical choice present during sentence completion.

One method used to vary the degree of lexical choice is to compare responses to sentence stems that are strongly constraining (e.g., *Water and sunshine help plants \_\_\_\_\_ “grow”*) to responses from sentence stems that are weakly constraining (e.g., *There was nothing wrong with the \_\_\_\_\_ “woman” / “car” / “food” / “table” / “weather”, etc.*). The logic here is similar to that presented with the verb generation task: There are few correct responses available for the strongly constraining sentence stems, whereas there are many responses available for the weakly constraining sentence stems. Hence, selection and control demands should be higher for weakly constraining sentences, because there are more options to choose between. Robinson, Shallice, and Cipolotti (2005) present a case study of patient CH, who has a progressive non-fluent aphasia. They observed that CH was more successful on strongly constraining sentences. Crowther and Martin (2007) presented data from two more participants with non-fluent aphasia. Consistent with participant CH’s data, they found that non-fluent AR was more successful on strongly constraining sentences. However, in contrast, they observed that non-fluent ML was not impaired on this task.

Another sentence completion task variation that has been used to investigate frontal lobe dysfunction is the Hayling task (Burgess & Shallice, 1996). In the Hayling task, as well as being given a typical sentence completion exercise, where a sensible sentence completion is required (e.g., *The death of his dog was a great \_\_\_\_\_ “shock”*), participants are given similar sentence stems, but are required to respond with semantically nonsensical completions (e.g., *Most cats see very well at \_\_\_\_\_ “banana” / “car” / “hill” / “cup”, etc.*). The nonsensical response requirement adds a new dimension to the task since participants need to suppress prepotent responses

that make sense. Burgess & Shallice (1996) found that, when compared to controls, patients with frontal lesions tended to produce: 1) slower responses for sensible completions; and 2) words of a closer semantic relation to the sentence frame for nonsensical completions (i.e., sensible completions, opposites of sensible completions, words that were related to the sentence subject, words that were semantically connected to the expected response, and words that vaguely fit the sentence). They concluded that when incorrect answers were automatically activated via a strong semantic relationship with the sentence stem, the frontal lesion group found it more difficult to override this activation in order to select a task appropriate word with no semantic relationship to the sentence stem. The authors also noted that the groups differed in terms of the strategies they implemented when producing nonsensical responses. It was observed that the control participants more frequently responded with different items found in the examiners office, or by systematically responding with items that were unrelated to the sentence stem but were related to each other. This is an interesting observation since conscious strategy use could be one avenue through which a frontal modulation process could operate.

More recently, the Hayling task variation has been used to investigate aphasia specifically (Robinson, Shallice, & Cipolotti, 2005). Non-fluent CH, who was discussed previously in reference to the production of sensible completions to strongly and weakly constraining sentence stems (see p. 129), was also asked to provide nonsensical completions to sentence stems that differed in constraint. When producing sensible completions, he did not show a significant RT effect, but he did respond more accurately when the response was highly constrained by the sentence. However, he did not show a constraint effect when producing nonsensical completions. This is an interesting observation, because the number of possible responses is least limited in this condition. Robinson et al. (2005) suggested that this might be due to use of a strategy through which he generated unrelated responses. For instance, if he used a strategy where he completed the sentence stem with items that were located in the exam room that caught his eye, the lexical competition that was truly influencing his results would actually have been fairly limited.

The primary interest of the current research is to find out how non-fluent aphasics respond to different levels of lexical competition. As has been discussed, past studies have done this by manipulating sentence constraint. However, it is also possible to add extrinsic competition to a sentence completion paradigm through use

of auditory distracters. In the current study we focus on this kind of extrinsic competition. Each sentence stem will be presented with a *distracter completion* that does, or does not, make sense. When participants produce their completion, they must disregard the distracter, and come up with an alternative response. In a high competition condition, the distracter will be a plausible completion of the sentence stem (e.g., *Barry wisely chose to pay the FINE*, “bill” / “cashier”, etc.). In a low competition condition, the distracter will not be related to the sentence stem (e.g., *Barry wisely chose to pay the RANGE*, “bill” / “cashier”, etc.)<sup>13</sup>. This task will be administered to three participant groups: non-fluent aphasics, fluent aphasics, and a neurologically intact control group. According to the selection and control hypothesis, the high competition condition distracters should induce more competition than the low competition condition distracters, because as well as being activated extrinsically they will be activated as a result of their semantic congruence with the sentence stem. This increase in competition is expected to result in a performance cost for all three participant groups: non-fluent aphasics, fluent aphasics, and older controls.

Further, it is hypothesised that the non-fluent aphasics will show exaggerated competition condition effects relative to the fluent aphasics and older controls because some or all non-fluents may suffer from a specific impairment to a lexical selection and control mechanism: This is expected to be reflected in a group by competition condition interaction.

Finally, based on findings that a sub-group of non-fluent aphasics, who show good naming but poor sentence production, are particularly sensitive to manipulations of context in naming tasks (see: Freedman et al., 2004; Schnur et al., 2006; Schwartz & Hodgson, 2002; Wilshire & McCarthy, 2002; Wilshire et al., 2006), we hypothesise that non-fluent aphasics who fit this profile may show stronger competition condition effects than non-fluent aphasics who do not. Three non-fluents in the current sample have relatively good single word naming, scoring above 40 on the BNT, and are expected to show these exaggerated effects (they are DA, ECV, and JHM; see pp. 49-64 for more details).

---

<sup>13</sup> The stimuli used are based on sentences sourced from the Bloom and Fischler (1980) sentence completion norms. Further details regarding the stimuli are detailed in the Materials section on p. 130.

## *Method*

### *Participants*

Participants included: seven non-fluent aphasics (CT, DD, DA, ECV, JHM, TB, and WS); five fluent aphasics (FS, NP, PS, STR, and XX); and seven neurologically intact older controls (AP, BC, EM, FC, GA, JI, and KR). Further details on the participants are available in the category exemplar task participant section (see pp. 49-64).

### *Materials*

The sentence stem stimuli were selected from the Lahar et al. (2004) sentence-final word completion norms (these are available online at <http://www.yccc.edu/faculty/lahar/norms.htm>)<sup>14</sup>. Using a standard sentence completion paradigm, Lahar et al. presented 358 participants with 119 incomplete sentences (e.g., *He mailed the letter without a \_\_\_\_\_*), and asked them to complete each sentence with a likely ending. The norms detail the different words that were used to complete each sentence, and the proportion of the participants who used each completion. The following criteria were used to select sentences from the norms for the current study: 1) The norms must indicate that at least two different responses were used; 2) To avoid very weakly constraining items, where participants could use a large number of different words to complete the sentence, the first and second most common responses had to make up at least 60% of the total responses. Using this procedure, 33 sentence stems were selected for use in this task. Although Lahar et al. did not include auditory distracters in their norming study, the high competition condition distracters included in the current task were the second most common completion of the selected sentence stems as detailed in the norms. The low competition condition distracters were syntactically correct words that were semantically unrelated to the sentence stems. A paired t-test confirmed that the distracters were frequency matched across conditions ( $t(32) = 1.971, p > .05$ ).

---

<sup>14</sup> The Lahar et al. (2004) paper provides more recent norms for the sentence completion norms originally presented by Bloom and Fischer (1980).

Details of the selected sentence stems and distracters for the high and low competition conditions are available in Appendix D1.

### *Procedure*

Each sentence stem was presented twice, once in the high competition condition and once in the low competition condition. In both conditions the sentence stem was presented visually on the computer screen (e.g., *Barry wisely chose to pay the \_\_\_\_\_*). When the sentence stem appeared on the computer screen, an audio recording of the sentence stem was played. In addition to the sentence stem, the audio recording included the auditory distracter (e.g., *Barry wisely chose to pay the FINE*). The auditory distracter did not appear visually.

Two versions of this task were created so that each participant could complete one version during each of two testing sessions. On both versions of the task each sentence stem was presented once, and in the same order. Sixteen of the items were taken from the high competition condition, and 17 from the low competition condition, or vice versa. Items from the different competition conditions were interspersed pseudo-randomly throughout version one. In version two, the alternative condition was used.

Participants were tested individually over two testing sessions that were at least two weeks apart. Testing took place at participants' homes, in a quiet testing room at Victoria University, or in a quiet testing room at Temple University. Audio recordings were made of each session, and were uploaded onto computer for further processing. PsyScope software was used to present the experiment on a Mac laptop computer. During each testing session, participants were instructed that they would see an incomplete sentence and that they should complete the sentence: Also, they would hear the sentence, but the sound file would include a completion that would sometimes make sense, but sometimes would not. They were instructed that they should not use the distracter as their completion, even if it did make sense. The instructions were given in a written format, plus the experimenter read the instructions to the participant and explained the task in her own words before commencing with several practice trials. Version one of the task was completed during the first testing session and version two was completed during the second testing session. A maximum of 30 seconds was allowed for participants to produce their response.

### *Data Analysis*

Each response was coded as correct or as an error. Initial responses were scored, although use of fillers (e.g., um, ah), one repetition of the distracter, and repetition of words from the sentence were allowed to occur before the initial response if they were not offered as a response (this was judged from prosodic cues). To be classified as correct, the response had to be a single word that made sense as a sentence completion. The use of incorrect plurals was not penalised. Morphological errors and morphological variants of the distracters were considered incorrect (e.g., *Newspaper*: “News” / “Paper”). Errors were classified as *distracter errors* if the distracter was repeated as a response, and all other error types were classified together as *other errors*. For correct responses, response latencies were measured by timing from the offset of the auditory distracter until the onset of the correct response.

Response accuracy and response latency data were analysed at the group and individual levels. In order to examine whether group accuracy was differentially influenced by competition condition across the different participant groups, a repeated measures logistic regression analysis was performed. In order to examine whether group latencies were differentially influenced by competition condition, a repeated measures general linear mixed model was performed. The regression analysis and the general linear mixed model analysis both included the following independent variables: group, competition condition, version, a two-way interaction of group and competition condition, and a three way interaction of group, competition condition, and version (participant and sentence stem were included as repeated measures). If the group variable, the version variable, or the interactions were not significant, they were removed from the analyses.

For the aphasic participants, response accuracy and response latency were also analysed at the individual level. Again, logistic regression was used to analyse response accuracy. The individual accuracy analysis included the following independent variables: competition condition, version, and a competition condition by version interaction, with sentence stem included as a repeated measure. When the competition condition and version interaction or the version variable were not significant, they were removed from the analysis. Individual latency analyses were

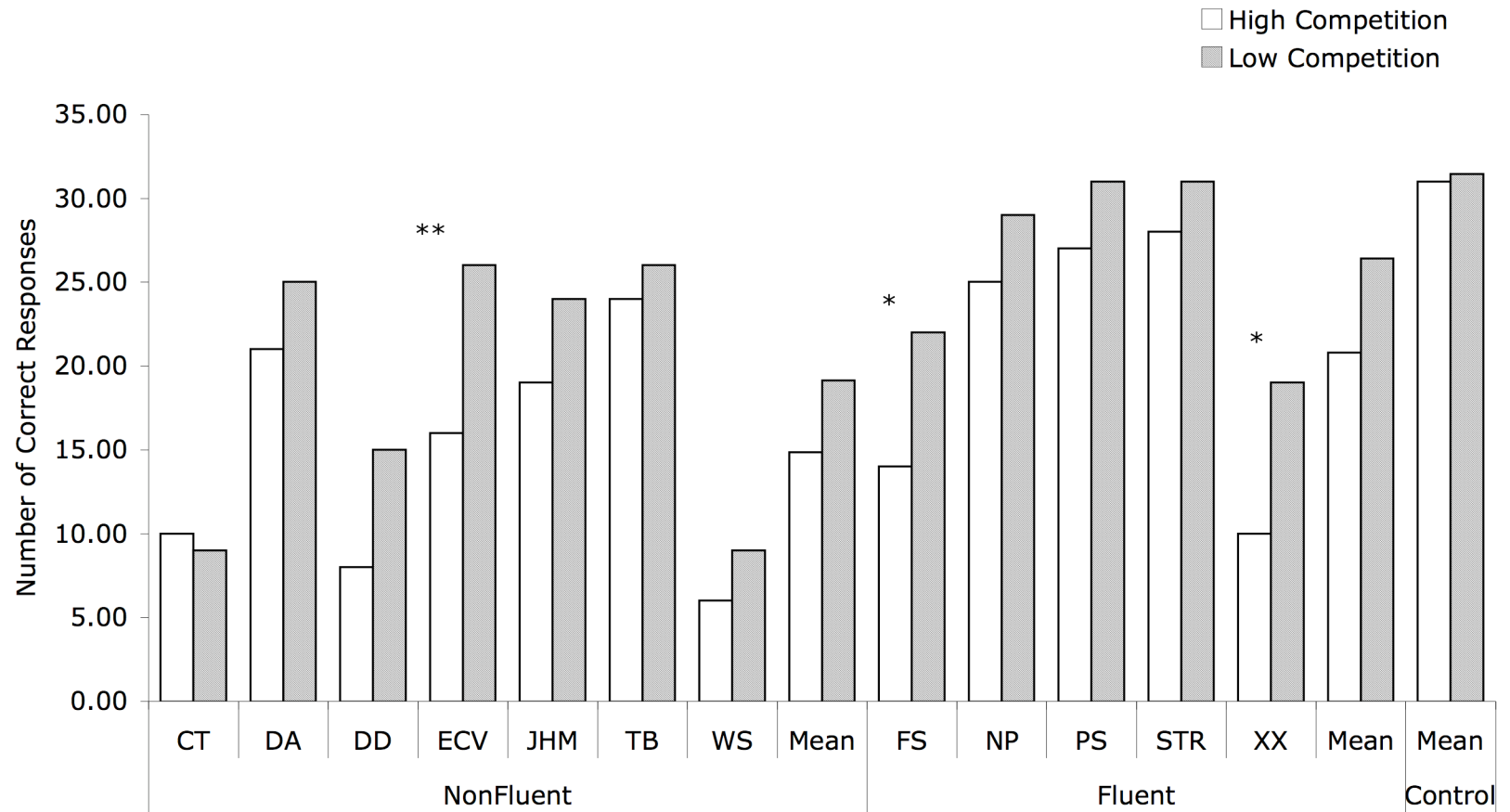
also performed: Paired t-tests were used to analyse each participant's response latencies from the high and low competition conditions.

### *Results*

An alpha level of .05 was used for all statistical tests.

*Response Accuracy.* The number of correct responses made by each participant and group in the high and low competition conditions is illustrated in Figure 5.1. Error types included: repetition of the distracter, multiple word answers, phonological paraphasias, non-word errors, attempts to spell the completion rather than say the word, production of completions that did not make sense, and failures to respond within the 30 second response period. In general, the non-fluents and the fluents showed very similar response patterns in the accuracy analysis: the non-fluents had a 22 % low competition condition accuracy advantage and the fluents had a 21 % low competition condition accuracy advantage. The controls had very even accuracy performance across competition condition, with a 1 % low competition condition advantage. The trend towards a low competition condition accuracy advantage was quite robust across aphasic participants, with only one participant showing a trend in the opposite direction (non-fluent CT).





*Figure 5.1.* Sentence completion task: Number of correct responses made in less than thirty seconds by participant and competition condition.

A repeated measures logistic regression analysis was performed on the grouped accuracy data: This revealed a highly significant main effect of condition confirming that more correct responses were made in the low competition condition than the high competition condition ( $\chi^2(1) = 25.66, p < .01$ ). There was also a highly significant main effect of group ( $\chi^2(2) = 160.72, p < .01$ ): The non-fluents made significantly fewer correct responses than the fluents ( $\chi^2(1) = 19.18, p < .01$ ), and the fluents made significantly fewer correct responses than the controls ( $\chi^2(1) = 19.86, p < .01$ ). However, against our hypothesis, there was no significant group by competition condition interaction ( $\chi^2(2) = 2.60, p > .05$ ).

In the individual accuracy analyses, one non-fluent aphasic made significantly more correct answers in the low competition condition than the high competition condition (ECV ( $\chi^2(1) = 7.14, p < .01$ )). A second non-fluent showed an effect that fell on the cusp of significance (DD ( $\chi^2(1) = 3.77, p = .05$ )). Of the fluent aphasics, FS and XX made significantly more correct answers in the low competition condition than the high competition condition (FS ( $\chi^2(1) = 4.57, p = .03$ ), XX ( $\chi^2(1) = 5.40, p = .02$ )). The remaining aphasics did not show a significant accuracy difference between the high and low competition conditions. Full statistical results are reported in Appendix D2.

The results indicate that non-fluent aphasics ECV and possibly DD, and fluent aphasics XX and FS, were less successful at completing sentences when they were presented with a high competition condition distracter than when they were presented with a low competition condition distracter. Although the other aphasics did not show a significant difference between the conditions, the general trend in the data was in the same direction as that seen in ECV, DD, XX, and FS's results (fewer correct responses in the high competition condition) suggesting that the lack of significant results may be partially due to a lack of statistical power.

*Distracter Error Analysis.* Although no clear difference in the size of the competition condition effect between the groups was observed in the accuracy analysis, it remained possible that there may be some differences in the types of errors made by the different groups. Since it was hypothesised that the non-fluent aphasics might find it more difficult to disregard the distracter than the other groups, a second

accuracy analyses was completed to compare the different types of errors made by the participant groups. The number of distracter errors was calculated relative to the total number of all other errors for each aphasic participant and each group (see Figure 5.2). The trends in the data showed that in general, the non-fluents made the most distracter errors, followed by the fluents, and the controls did not make any distracter errors. As a group, the non-fluents also had the highest proportion of distracter errors relative to the other errors they made: In the high competition condition, 41 % of their errors were distracters, and in the low competition condition 6 % of their errors were distracters. For the fluent aphasics, these figures were, 26 % and 0 % respectively.

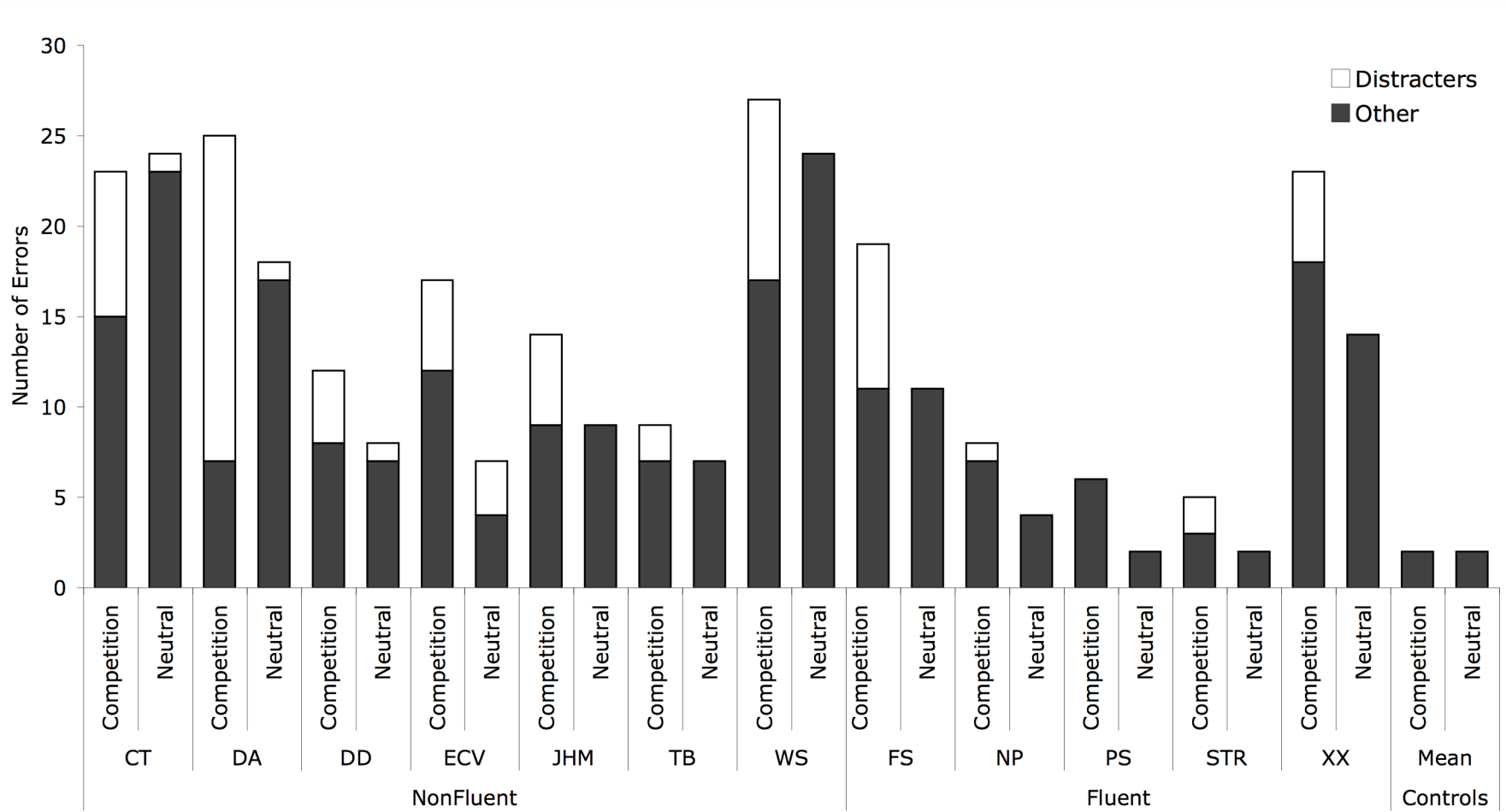


Figure 5.2. Sentence completion task: Number of distracter errors and other errors across session by participant and competition condition.

A logistic regression analysis was used to compare the fluent and non-fluent groups' relative rates of distracter errors and other errors. There was a significant main effect of group, confirming that the overall relative rate of distracter errors to other errors was highest in the non-fluent group ( $\chi^2(1) = 5.22, p < .05$ ). Also, a highly significant main effect of competition condition was observed, confirming that distracter errors occurred more frequently in the high competition condition across groups ( $\chi^2(1) = 51.57, p < .01$ ). However, an interaction of group and competition condition was not observed ( $\chi^2(1) = 1.98, p > .05$ ), suggesting that a similar increase in the number of distracters in the high competition condition was observed in both groups. It is important to note however, that within the non-fluent and the fluent groups, some participants appear more susceptible to distracter errors than others, and not all individuals in each group show a response pattern that is consistent with the group pattern. For example, non-fluent DD made a particularly high number of distracter errors in the high competition condition, and fluent PS made no distracter errors at all. Hence it appears that select participants are driving the group differences, and data from other participants is diluting the apparent severity of effects observed in those individuals who do show significant differences. These results suggest that there were qualitative differences in the types of errors made by the different groups, and provide support for the hypothesis that the non-fluent group found it more difficult to disregard the distracters than the fluent group.

*Response Latencies.* Because the RT distributions were highly skewed towards shorter responses, each RT was log transformed prior to analysis. Also, because of the paired design (each sentence appeared in both the high and low competition conditions), the participant had to get both versions of the sentence correct for their RTs for that sentence to be included in this analysis. Geometric mean RTs and geometric coefficients of variance were calculated for each participant, group, and competition condition (see Table 5.1). In general, the groups showed more marked competition condition differences in the latency data than in the accuracy data: the non-fluents had a 27 % low competition condition latency advantage, and the fluents had a 4 % low competition condition latency advantage. The controls had a 5% low competition condition latency advantage.

Table 5.1  
*Sentence Completion Task: Geometric Mean Response Latencies (msec) by Participant and Competition Condition for Correct Response Pairs, with Paired t-test Statistics*

Test Statistics	Geometric Mean Response Latency (Geometric Coefficient of Variance)					<i>t</i>	<i>df</i>	<i>p</i>
	Low Competition		High Competition		Percentage Change			
Non-Fluents								
CT	4364	(81)	10161	(122)	133	2.221	4	0.09
DA	12415	(85)	14233	(65)	15	0.779	15	0.45
DD	6271	(159)	10745	(88)	71	1.220	4	0.29
ECV	2680	(164)	3559	(138)	33	1.122	13	0.28
JHM	3498	(67)	3509	(150)	0	0.014	15	0.99
TB	2533	(95)	3659	(94)	44	2.150	19	0.04
WS	4709	(65)	3527	(59)	-25	0.734	3	0.52
Group Geometric Mean	4234	(146)	5392	(153)	27			
Fluents								
FS	3449	(284)	2359	(248)	-32	0.936	10	0.37
NP	4990	(37)	4841	(76)	-3	0.251	21	0.80
PS	125	(132)	132	(125)	6	0.223	25	0.83
STR	1426	(463)	3672	(137)	158	3.104	26	< .01
XX	1847	(263)	1672	(199)	-9	0.148	6	0.89
Group Geometric Mean	1176	(355)	1225	(144)	4			
Controls								
Group Geometric Mean	463	(375)	488	(385)	5			

To investigate the significance of any latency differences between the three groups, the data was analysed using a general linear mixed model, incorporating item and participant as random repeated factors. The expected main effect of competition condition was not significant ( $F(1, 769) = 0.01, p > .05$ ). However there was a highly significant main effect of group ( $F(2, 771) = 38.32, p < .01$ ): The non-fluents were significantly slower than the fluents ( $F(1, 771) = 3.91, p > .05$ ), who were significantly slower than the controls ( $F(1, 771) = 34.05, p < .01$ ). The main effect of version was not significant ( $F(1, 770) = 0.10, p > .05$ ). The predicted group by competition condition interaction was not significant ( $F(2, 767) = 0.76, p > .05$ ). The

three-way interaction of group, competition condition, and version was also not significant ( $F(2, 762) = 0.88, p > .05$ ).

To investigate whether individuals within the groups showed latency differences across the competition conditions, each participant's data was analysed using paired samples t-tests (see Table 5.1). Low accuracy rates did lead to a lack of statistical power for some participants. However, some significant differences were identified. Of the non-fluent aphasics, TB had significantly faster RTs in the low competition condition than the high competition condition, but CT, DA, DD, ECV, JHM, TB, and WS showed no significant difference between the conditions. Of the fluent aphasics, STR had significantly faster RTs in the low competition condition, but FS, NP and XX showed no significant difference between the conditions.

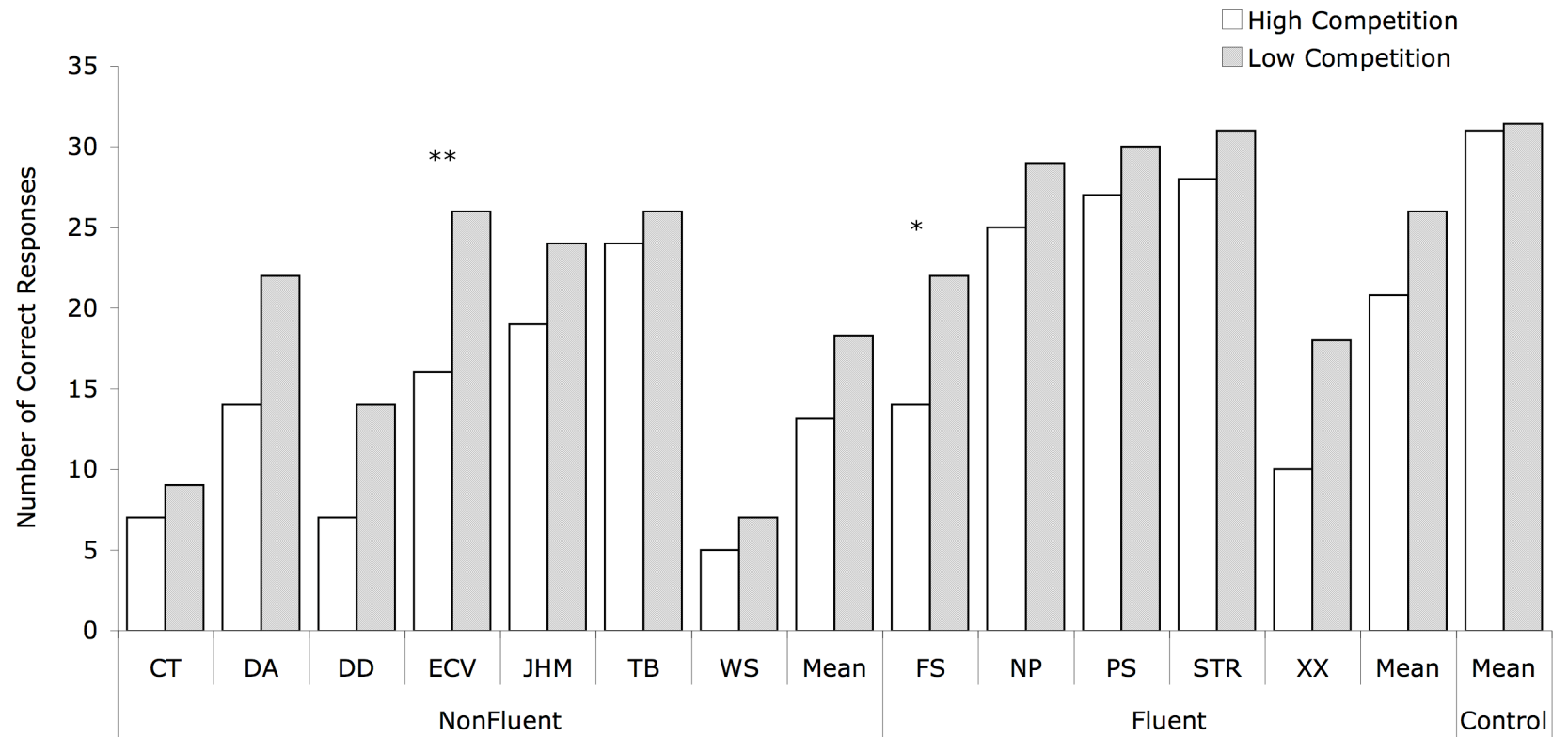
### *Supplementary Analyses*

*Accuracy Analysis using Shortened Response Periods.* The response accuracy analysis detailed in the previous subsection compared all correct initial responses to all incorrect initial responses. Where no answer was given during the generous 30-second response period, the response was considered to be incorrect. The duration of the 30-second response period was set to give slow responders a decent chance to produce a correct response, while limiting the period during which they would struggle, and often fail, to produce a response. When left unchecked the response period can run into several minutes, and might result in a response time that is potentially more a measure of determination than a measure of the process being investigated. It seemed plausible that extraneous variables, such as frustration, loss of concentration, and how determined a participant was to get the task correct, or speed-accuracy trade-offs, were more likely to influence the results during long response time trials, because participants were more likely to be stuck, or off-task. Hence, the influence of the distracters may be most evident at shorter response times, where the influence of these extraneous factors should be minimal.

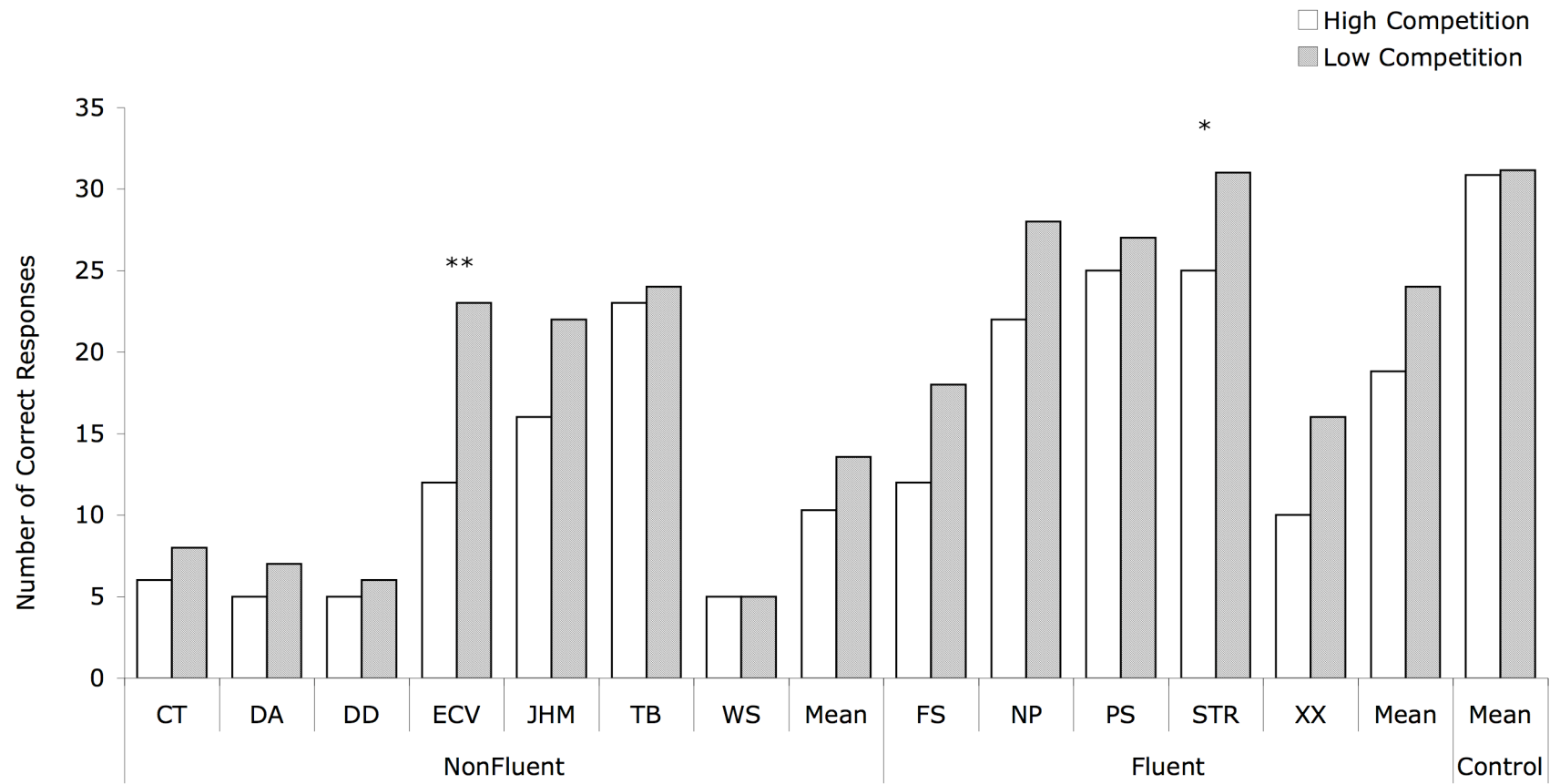
Consequently, two further exploratory analyses were completed. The first exploratory analysis compared all initial responses that were answered correctly during a 20-second response period, to those that were not (see Figure 5.3). The second compared all initial responses that were answered correctly during a 10-second response period, to those that were not (see Figure 5.4). These modifications influenced some participant's scores more than others, since relatively fast responders

made most of their correct responses in less than 10 seconds. Accurate but slower responders (e.g., non-fluent DA) were most influenced by the response period modifications. However, in all of the response accuracy analyses the general trends were very similar. In general, the number of correct responses was higher in the low competition condition, but more often than not, the difference between the low and high competition conditions was not significant at the individual participant level. With that said, in each analysis, at least one fluent and one non-fluent aphasic did show a significant difference between the competition conditions: Non-fluent ECV showed a significant difference in all analyses, including the original analysis (30 second deadline), and the two new shorter-deadline analyses; Fluent FS, who showed a significant difference in original analysis, continued to show an effect in the 20 second deadline analysis, but not in the 10 second deadline analysis. Fluent XX, who also showed a significant difference in the original analysis, failed to show significant effects on the analyses using the two shorter deadlines. Fluent STR was the only participant to reveal a significant condition effect only when a shorter deadline was used: for STR, the effect of condition was significant only in the less than 10 second analysis.





*Figure 5.3.* Sentence completion task: Number of correct responses made in less than twenty seconds by participant and competition condition.



*Figure 5.4.* Sentence completion task: Number of correct responses made in less than ten seconds by participant and competition condition.

Because the accuracy effects shown by FS and XX appear to be exacerbated with increasing RTs, the shortened response period analyses do not tell us more useful information (this pattern of responding may be related to statistical power increase associated with including more data in the analysis). However, STR only showed a significant accuracy difference on trials where she made fast response times (<10 seconds): This pattern of responding is in line with the hypothesis outlined above, that faster responses may give a cleaner measure of the competitive effects of related distracters than slower responses. If we assume that this is the case, the results suggest that her accuracy rates may have been influenced by extraneous factors, such as increasing frustration, loss of concentration, and her determination to make a response: These factors were possibly hiding the significant competition condition accuracy effect observed only in her 10 second deadline analysis.

*Sentence Constraint Analysis.* One method that has previously been used to vary the degree of lexical choice in a sentence completion task is to compare responses to sentences stems that are strongly constraining (e.g., *Water and sunshine help plants \_\_\_\_\_* “grow”) to responses to sentence stems that are weakly constraining (e.g., *There was nothing wrong with the \_\_\_\_\_* “woman” / “car” / “food” / “table” / “weather”, etc.) (Robinson et al., 2005; Crowther & Martin, 2007). Although our design did not explicitly manipulate sentence constraint, some of the sentence stems were more constraining than others. Therefore, it was possible to informally examine whether the degree of sentence stem constraint modulated any of the effects observed, at least for those participants who showed significant competition condition effects (non-fluent ECV and fluents FS, XX, and STR).

This is a potentially interesting exercise because sentence constraint and competition condition may tap into some different cognitive mechanisms. Unlike the competition condition manipulation, which is designed to manipulate the degree of lexical activation of one carefully selected competitor word, a sentence constraint manipulation would operate in a different way, because it would vary the number of alternative responses available and/or how hard it is to think of a completion. Also, because extrinsic competition is introduced with the distracter manipulation, it is possible that this manipulation taps into an inhibition mechanism more so than the constraint manipulation, because any activation from the distracters must be over-

ridden to successfully complete the task. Hence, if a constraint mechanism or a competition mechanism, or both, particularly influence participants' performance we may gain some insight regarding which of the aforementioned processes has more influence on word production.

In order to examine sentence constraint, a response strength ratio was calculated for each of the sentence stems, by taking the response frequency data from the Lahar et al. (2004) norms, and dividing the proportion of the most common response by the proportion of the second most common response. These response strength ratios were used to identify a weakly constraining group of items (with ratios of less than three), and a strongly constraining group of items (with ratios of nine or more). The constraint condition for each of the sentences is detailed in Appendix D1. Because this study was not designed to manipulate constraint, the weak constraint group included 24 items, and the strong constraint group included 9 items. Since the size of the groups is so different, the results of this analysis need to be interpreted with due caution. Nevertheless, they may show some interesting response patterns that could be followed up in future research.

Accuracy rates were calculated for each participant and competition/constraint condition (see Figure 5.5). A clear trend did not emerge in the descriptive data. However, non-fluent ECV and fluent XX showed similar trends: As well as showing a high competition condition disadvantage, they showed trends towards a weak constraint disadvantage, with a general shift down in accuracy on the weak constraint sentences relative to the strong constraint sentences. This result indicates that they may be influenced by the number of co-activated responses (i.e., spread of activation), and the level of competitor co-activation (i.e., the degree of competitor activation irrespective of spread). These results are also consistent with the possibility that the participants may be influenced by the difficulty of the search process. Fluent FS showed a different response pattern: Although she showed a high competition condition disadvantage, sentence constraint did not appear to influence her performance. This result suggests that she may be more sensitive to the degree of activation that needs to be overcome, or that she finds it particularly difficult to inhibit/overcome extrinsic competition. Because fluent STR only showed a significant distracter accuracy effect on responses that occur in less than 10 seconds, she shows little evidence of a competition condition effect in this analysis. She also shows no evidence of a constraint effect in this analysis, showing a trend in the

unpredicted direction, towards a weak constraint advantage. At this stage however, the possibilities outlined above are highly speculative, as they are based on a small amount of data from few participants. However, attempting to tease these features apart is an issue that could be addressed in future research.

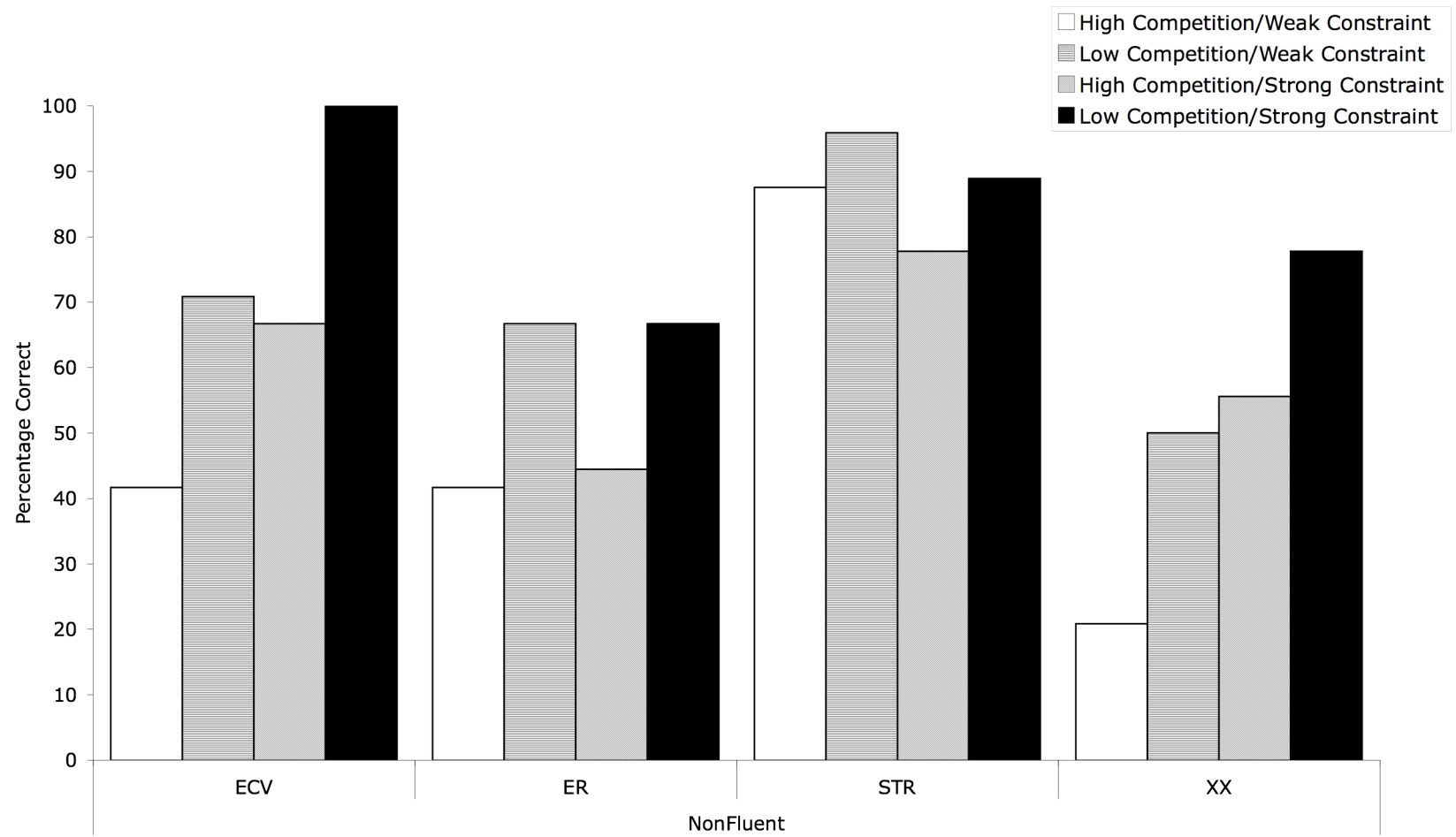


Figure 5.5. Sentence completion task: Accuracy rates by participant and competition/constraint

### *Discussion*

This experiment examined sentence completion performance of three groups of participants (non-fluent aphasics, fluent aphasics, and older controls) under two competition conditions. In the low competition condition, sentence stems were presented with an auditory distracter that did not make sense as a sentence completion (e.g., *Barry wisely chose to pay the RANGE*). In the high competition condition, sentence stems were presented with an auditory distracter that did make sense as a sentence completion (e.g., *Barry wisely chose to pay the FINE*). There were several factors that contributed to the logic behind the competition condition manipulation. First, it was assumed that in both competition conditions the presentation of the distracter would raise its lexical activation. Further, in the high competition condition only, it was expected that the sentence stems would also automatically activate the distracters by virtue of the shared semantic association. Thus, relatively strong demands would be placed on the selection and control mechanism that functions to resolve lexical competition in the high competition condition. The second step in the logic behind the manipulation was that if a participant had a disorder affecting a lexical selection and control mechanism, they would have more difficulty in the high competition condition.

The first hypothesis made was that the high competition condition would result in relatively high error rates and long response latencies than the low competition condition in all participant groups. This prediction was supported in the group accuracy analysis, but not in the group latency analysis. The failure to observe a high competition condition disadvantage in the group latency analysis was unexpected, but may be related to difficulties that were encountered identifying when participants started planning their responses across competition conditions. This issue is discussed in more detail on p. 152.

The second hypothesis made was that if some, or all, individuals with non-fluent aphasia suffer from a deficit involving a selection and control mechanism, they should show abnormally exaggerated high competition condition costs in accuracy and/or latency, relative to fluent aphasics and neurologically intact controls. This hypothesis was tested at the group and individual levels. The grouped data did not support this hypothesis because there was no significant interaction of group and

competition condition in the accuracy or latency analyses. At the individual level, non-fluent ECV and fluents FS and XX showed significant competition condition effects in the accuracy analyses. In the latency analyses, non-fluent TB and fluent STR showed significant competition condition effects. If you consider both accuracy and latency together, two non-fluents (ECV and TB) and three fluents (FS, XX, and STR) showed significant competition condition effects. Hence, if you look at the prevalence rates of significant differences across accuracy and latency it appears that, contrary to expectations, the fluents had a higher ratio of participants who showed significant low competition condition advantages in accuracy or latency than the non-fluents (3:5 and 2:7 respectively). Hence no clear support was observed for the hypothesis that non-fluents would show exaggerated competition condition effects relative to fluents and controls.

The final prediction made was that non-fluent aphasics with marked dissociations between their poor sentence production and good single word production abilities would show particularly exaggerated competition condition effects (that is DA, ECV, and JHM). In the accuracy analysis, one of these participants, non-fluent ECV, was the only aphasic participant to show a highly significant low competition condition advantage. However, non-fluents DA and JHM did not show significant differences in either accuracy or latency. The other participants to show significant differences in the individual analyses did not fit this profile (non-fluent TB, and fluents FS, XX, and STR). Adopting a conservative approach, there does not appear to be much support for this hypothesis. However, it is interesting to note that ECV showed a highly significant difference in this study but not the verb generation task, while DA and JHM showed significant differences in the verb generation task, but not the current study. This point reinforces the fact that these tasks are not pure measures of lexical competition. However, it is still unclear what specific participant characteristics might determine these kinds of effects, and what situations may cause these kinds of effects to arise. The current evidence suggests that fluent aphasics may also be susceptible to these competition condition effects, although it is possible that they may be producing similar accuracy and latency data as the result of a qualitatively distinct deficit. Although no clear differences were observed between the aphasic groups in the latency and accuracy analyses, the results from the distracter error analysis are consistent with the idea that there may be some qualitative differences in the types of errors made by the non-



fluents and the fluents. Specifically, it appeared that many of the non-fluent aphasics produced more distracter errors (that is, they repeated the distracters in incorrect completions of the sentence stem).

One concern with the current study is that participants may use a strategy where they generate a completion before they hear the distracter, and switch to another response if their completion matches the distracter. If participants were using such a strategy, then there may be accuracy and/or latency costs associated with their switch, which would only be evident when the distracter was an appropriate sentence completion. However, it should be borne in mind that even in the high competition condition, the distracter word was never the most common completion of the sentence, but rather than second most common completion. Thus, there should be relatively few occasions where the participant came up with that completion spontaneously and hence be forced to switch to another one. Also, many (if not all) of the aphasic participants had some reading difficulties, so it would seem unlikely that they would read and process the written stimulus fast enough to generate possible candidate responses prior to hearing the distracter word. Indeed, it was certainly not the case that those aphasic participants who showed the most marked competition effects were those that were the fastest or most accurate readers.

The question regarding when participants actually start planning their responses makes all of the latency analyses difficult to interpret. However, the RTs were consistently measured from the end of the distracter until the beginning of the correct response, so any individual differences in strategy use should be minimally apparent in the individual analyses, presuming the patients used the same method throughout. However, there may have been differences in the manner of completion between the groups, so the group RT results in particular should be treated with caution (this may be a contributing factor to the failure to observe a significant group latency effect).

In summary, as was discussed in the introduction, the selection and control hypothesis predicts that the high competition condition will be more difficult than the low competition condition, that non-fluents will be the most strongly influenced by competition condition effects, and that non-fluents who show clear dissociations between good single word production and poor sentence production may be the most strongly influenced by distracter effects. The current data does support the hypothesis that the competitive distracters are more difficult to ignore than neutral distracters.

Also, some individuals do appear to be more susceptible to distracter effects than others, but due to high error rates and a lack of statistical power, it is difficult to conclude that those who do not show significant effects are not affected by the competition manipulation at some level. It is not clear what participant characteristics are associated with the occurrence of significant competition condition effects in this task, but there does not appear to be a clear distinction between non-fluent and fluent aphasics. Identifying any common features in those participants who do show significant effects is a possible avenue for future research, as is further investigation of the types of errors made, because the distracter error analysis provides evidence which suggests that there may be qualitative differences in response patterns between groups.

### Determinants of Performance across the Four Tasks: Unpacking the Contributions of Specific Components of Fluency

One goal of the current research was to examine whether some participants would consistently show exaggerated lexical competition effects on the different experimental tasks we completed. Table 6.1 presents a summary of each aphasic participant's sensitivity to competition effects on each task. Any significant competition effects that were observed on traditional accuracy or latency analyses are marked with a tick (all these effects were in the predicted direction). Every aphasic but one showed a significant competition effect in accuracy and/or latency on at least one task. However, none of the participants showed a significant competition effect on more than two tasks. The aphasics who showed the highest prevalence of significant competition effects were non-fluents CT, ECV, and WS, who each showed three significant competition effects across two tasks. The aphasics who showed the lowest prevalence of significant competition effects were non-fluent DD, who did not show a significant competition effect on any of the tasks, and fluents FS, NP and PS who showed one significant competition effect each.

Table 6.1

*Summary of Individual Results on all Experimental Tasks, Showing where Significant Effects were observed on Simple Accuracy and Latency Analyses*

	Category Exemplar	Verb Generation		Name Agreement		Sentence Completion	
	Number of Exemplars	Accuracy	Latency	Accuracy	Latency	Accuracy	Latency
Non-Fluents							
CT	-	-	√	√	√	-	-
DA	-	-	√	-	√	-	-
DD	-	-	-	-	-	-	-
ECV	-	-	-	√	√	√	-
JHM	-	√	-	-	√	-	-
TB	-	-	-	-	√	-	√
WS	n/a	-	√	√	√	-	-
Fluents							
FS	-	-	-	-	-	√	-
NP	-	-	-	-	√	-	-
NS	-	-	-	√	√	n/a	n/a
PS	n/a	-	-	-	√	-	-
STR	-	-	-	-	√	-	√
XX	-	-	√	-	-	√	-

*Are specific fluency components related to lexical competition effects?*

As well as looking at the prevalence of competition effects across tasks, we were interested in whether participants who were particularly susceptible to competition effects might have characteristics in common that were more specific than the BDAE's non-fluent/fluent distinction. If so, these features could provide useful information about the specific cognitive determinants of these kinds of competition effects. Fluency, as it is formulated in the BDAE, is a heterogeneous concept based on a cluster of separate, although highly correlated, aspects of speech such as: articulatory agility, sentence intonation, phrase length, and the variety of grammatical forms used. We chose to compare scores from four more specific measures associated with the fluent/non-fluent distinction, to performance on our tasks. Each of the fluency measures we explored is summarised next.

The first measure we explored was the rate of speech production. This measure was chosen because it reflects performance on a range of fluency related skills. These include sentence level processes such as how effectively syntactic and

grammatical information is processed. However, it also reflects other aspects of fluency, such as how quickly participants can move from one idea to another (conceptual fluency), and how quickly participants can articulate words. The specific measures we chose were: 1) the number of words produced per minute on the Qualitative Production Analysis (QPA) (Saffran, Berndt, & Schwartz, 1989), which provides a raw measure of the rate of speech in a relatively unconstrained task (see Appendix A2 for a description of this test), and; 2) QPA words per minute score divided by percentage correct score from the Boston Naming Test (BNT) (Goodglass, Kaplan, & Barresi, 2001), which provides a measure of speech rate that also accounts for the overall severity of impairment (using naming as an example of a relatively rate-independent production task).

The second set of measures explored were the production rates of open and closed class words, and of nouns and verbs. This measure was chosen to show which participants had particular difficulties with elements of grammatical encoding. Also, because these scores are calculated independently of the rate, or amount, of speech produced, they could potentially pick up different patterns to those observed in measures that consider speech rate. The specific measures we chose were: 1) the proportion of open to closed class words produced on the QPA, which provides a measure based on a relatively spontaneous speech sample; 2) the proportion of open to closed class words produced on the Sentence Production Test (SPT) (Lukken, 2006), for a measure from single sentence descriptions of pictured events, and; 3) the proportion of nouns to verbs produced on the noun-verb naming test (Zingeser & Berndt, 1988).

The third measure explored was the ability to produce nouns in isolation compared to the ability to produce nouns in context. This measure was chosen because it may be particularly relevant to the evaluation of competition's role, since production of words in sentences is likely to involve controlling lexical competition from other words planned for the same utterance. This measure has also been related to the non-fluent/fluent distinction (see Williams and Canter, 1982). The specific measure we chose was the number of nouns produced correctly on the Single Noun Test (SNT) (Lukken, 2006), in which participants name single pictured objects. This score was then divided by the number of nouns produced correctly on the SPT, where

participants must name the same items within sentences that describe more complex events.

The final set of measures explored related to conceptual fluency, more specifically, the ability to select and generate lexical items that fit within certain categories. Our measures were based on the COWAT (Spreen & Strauss, 1991). These measures were chosen because they give a less constrained measure of fluency than object naming, allowing for more potential use of compensatory strategies, while still keeping syntactic demands minimal. The score also reflects high-level processes, such as how quickly participants can come up with new ideas. However, articulatory processes should have less influence here than they do on some of the other fluency measures, because participants are only required to produce short single word utterances. The specific measures chosen were: 1) COWAT letter fluency score divided by BNT percentage correct score, which was selected to give a measure of conceptual fluency that accounts for overall naming deficit severity; and 2) COWAT category fluency (Animals) divided by BNT percentage correct score, which was selected to give a measure of conceptual fluency that accounts for overall naming deficit severity without placing such high demands on phonological processing as the letter fluency task above.

Each participant's score for the aforementioned measures is detailed in Table 6.2 where available.

Table 6.2  
*Additional Diagnostic Measures by Participant*

	Non-Fluents							Fluents					
	CT	DA	DD	ECV	JHM	TB	WS	FS	NP	NS	PS	STR	XX
QPA words per minute	25.57	22.00	22.00	72.35	23.00	77.00	Not Available <sup>1</sup>	53.00	92.00	48.00	Not Available <sup>2</sup>	110.00	54.00
QPA words per minute / BNT percentage correct	0.51	0.24	0.69	0.90	0.30	1.22	Not Available <sup>1</sup>	0.78	1.75	0.73	Not Available <sup>2</sup>	2.50	1.35
QPA proportion open to closed class words	1.05	0.88	3.61	0.65	1.59	0.83	Not Available <sup>1</sup>	0.81	1.19	0.85	Not Available <sup>2</sup>	0.90	0.94
Lukkien & Wilshire SPT proportion open to closed class words	1.00	1.13	1.00	0.71	1.27	1.01	1.86	0.73	0.79	1.07	0.93	0.86	0.90
Lukkien & Wilshire proportion nouns correct on SNT / proportion correct nouns on SPT	1.27	0.90	1.31	1.25	1.07	1.03	1.89	1.25	1.25	1.04	1.07	0.96	1.29
Zingeser and Berndt noun-verb naming test noun / verb score	1.00	0.98	1.50	1.26	1.10	1.06	3.33	1.88	0.94	1.25	1.40	0.85	1.11

Table 6.2  
*Additional Diagnostic Measures by Participant*

	Non-Fluents							Fluents					
	CT	DA	DD	ECV	JHM	TB	WS	FS	NP	NS	PS	STR	XX
COWAT letter fluency score / BNT % correct	0.18	0.20	0.03	0.08	0.07	0.19	0.04	Not Available <sup>3</sup>	0.44	0.41	0.34	0.50	0.30
COWAT category fluency score (Animals only) / BNT % correct	0.14	0.07	0.16	0.11	0.09	0.17	0.21	0.09	0.13	0.14	0.10	0.20	0.08

<sup>1</sup> WS was given the QPA but was unable to provide a sample to be analysed. He was only able to give direct responses to questions.

<sup>2</sup> PS was not given the QPA because he had a second stroke before the data could be collected.

<sup>3</sup> This score was not available because of a technical failure.



If the participants are grouped by their performances on the measures detailed in Table 6.2, a similar pattern is observed to when they are grouped by their diagnostic non-fluent/fluent distinction because the participants who had the poorest scores on the various fluency components were usually diagnosed with a non-fluent form of aphasia. One possible exception to this was TB, who was unusual in that she scored within the fluent range on the QPA words per minute score and the QPA words per minute over BNT percentage correct score. Also, she did not score comfortably within the non-fluent range on the measures not specifically related to speech rate. However, omitting TB's data from the non-fluent group appears to further support the conclusions we have made thus far: Although the non-fluent aphasics appeared to be more likely to show competition effects on the verb generation task, TB did not; also, although the fluent aphasics appeared to be marginally more likely to show competition effects on the sentence completion task, TB showed a significant latency effect. Further, on the sentence completion task distracter analysis, the non-fluent group appeared to find it more difficult to disregard the distracters than the fluent group, because they made more distracter errors, but TB made the smallest number of distracter errors of the aphasic participants.

In order to examine how the new fluency-related measures detailed in Table 6.2 were related to the results on the experimental tasks, we first performed a series of informal visual analyses: For each measure, we ordered the patients by their fluency-measure score, and then we examined the distribution of significant lexical competition effects on the experimental tasks. This approach allowed for a visual analysis of how the participants who showed significant effects were dispersed according to the various, more specific, features of their aphasias. Then we examined whether scores on these measures were correlated with difference scores between the low and high competition conditions on each task. These analyses are described next.

*Speech Rate.* For this analysis, the data were ordered by the participant's QPA words per minute scores divided by their BNT percentage correct scores (see Table 6.3). The first point to note is that this measure was quite sensitive to aphasia subtype, and most of the non-fluents had low QPA WPM/BNT scores. The two exceptions to this statement were TB who was discussed previously because of her atypical fluency scores, and ECV, who both showed unusually fluent QPA WPM/BNT scores relative to the other non-fluents. This score was chosen because it

gives a measure of speech rate that takes into account overall impairment severity, using naming as an example of a relatively rate-independent production task<sup>15</sup>.

Table 6.3

*Summary of Individual Results on Selected Tasks, Showing where Significant Effects were observed on Simple Accuracy and Latency Analyses, Ordered by QPA Words per Minute over BNT Percentage Correct Score*

	QPA words per minute by BNT percentage correct	Verb Generation		Name Agreement		Sentence Completion	
		Accuracy	Latency	Accuracy	Latency	Accuracy	Latency
WS	Not Available <sup>1</sup>	-	√	√	√	-	-
DA	0.24	-	√	-	√	-	-
JHM	0.30	√	-	-	√	-	-
CT	0.51	-	√	√	√	-	-
DD	0.69	-	-	-	-	-	-
NS	0.73	-	-	√	√	n/a	n/a
FS	0.78	-	-	-	-	√	-
ECV	0.90	-	-	√	√	√	-
TB	1.22	-	-	-	√	-	√
XX	1.35	-	√	-	-	√	-
NP	1.75	-	-	-	√	-	-
STR	2.50	-	-	-	√	-	√

<sup>1</sup> WS was given the QPA but was unable to provide a sample to be analysed. He was only able to give direct responses to questions.

Looking at the verb generation task data, four out of the five significant competition effects in accuracy and/or latency occurred in participants with low QPA words per minute/BNT scores. However, XX did not fit the general pattern shown by the other participants; he showed a significant latency effect even though he had a relatively good QPA words per minute score relative to his single word naming score. This suggests that in addition to poor fluency, some other unknown factor may influence susceptibility to competition effects.

<sup>15</sup> PS was excluded from this analysis because he had a second stroke before he could complete the QPA. Although no QPA data were available for WS either, this was because when he was given the test, he was unable to produce a sample (i.e., he was severely impaired on the task). WS's inability to complete the QPA is informative in itself, so his data have been included in the table.

In contrast to the verb generation data, examination of the sentence completion data shows that the majority of the significant competition effects occurred in participants with high QPA words per minute/BNT scores - that is, those with a relatively fast rate of speech in relation to their isolated word production ability. In this task, we suspected that a general sentence processing impairment might have hidden competition effects by extraneously influencing how participants completed the task. However, a similar analysis, where the participants are organised by their comprehension scores for sentences with reverse role distracters from the Philadelphia Comprehension Battery (see Appendix A2 for a description of this task), shows that poor comprehension of sentences with reverse-role distracters was not especially prevalent in the participants who did not show significant effects.

Another hypothesis we explored was that there may be a correlation between QPA words per minute / BNT percentage correct scores and percentage change scores between the high and low competition conditions on the four experimental tasks (see Table 6.4). Looking at the verb generation task data, there was a trend towards a significant negative correlation between participants with QPA words per minute/BNT scores, and percentage change scores in the latency data ( $p = .07$ ). In other words, there was a trend towards the non-fluent pattern of responding, where participants have a slow rate of speech on the QPA relative to their isolated word production scores, being associated with a large difference between the competition conditions. In addition, we explored the possibility that the verb generation task latency data may show a correlation between relatively low QPA words per minute scores and bigger differences between their performance in the high and low competition conditions. This hypothesis was supported by the latency data: The participants' percentage change values showed a significant negative correlation with their QPA words per minute scores ( $r(11) = -0.51, p < .05$ ). Because participants need to have good articulatory agility to get a good score on the QPA words per minute measure, it is possible that there may be a relationship between good articulation and good competition resolution as it is observed on the verb generation task.

Table 6.4

*Summary of Individual Results on Experimental Tasks, Showing Percentage Change Values between Competition Conditions on Simple Accuracy and Latency Analyses Ordered by QPA Words per Minute over BNT Percentage Correct Score, with Correlation Statistics*

	QPA words per minute by BNT percentage correct	Category Exemplar  Number of Exemplars	Verb Generation		Name Agreement		Sentence Completion	
			Accuracy	Latency	Accuracy	Latency	Accuracy	Latency
DA	0.24	4	-8	67	14	59	16	15
JHM	0.30	5	25	19	12	44	21	0
CT	0.51	30	-6	56	32	50	-11	13
DD	0.69	59	8	12	22	22	47	71
NS	0.73	13	7	33	56	47		
FS	0.78	0	-4	0	37	30	36	-32
ECV	0.90	24	29	19	31	49	38	33
TB	1.22	33	8	8	-2	34	8	44
XX	1.35	22	14	32	11	23	47	-9
NP	1.75	18	6	28	13	58	14	-3
STR	2.50	27	7	3	7	47	10	15
<i>Correlation Coefficient <math>r</math></i>		0.22	0.09	-0.47	-0.39	-0.01	-0.08	0.3
<i>df</i>		9	9	9	9	9	8	8
<i>One-tailed <math>p</math></i>		0.26	0.39	0.07	0.12	0.49	0.41	0.1

*Grammatical Class Composite Score: Open vs. Closed Class Words and Nouns vs. Verbs.* For this analysis, the data was ordered by each participant's average score from the three sub-measures included. There does not appear to be a clear relation between this component of fluency and those participants who showed significant competition effects on our tasks (see Table 6.5).

Table 6.5

*Summary of Individual Results on Selected Tasks, Showing where Significant Effects were observed on Simple Accuracy and Latency Analyses, Ordered by Grammatical Class Composite Score (the Average of the QPA Open to Closed Class Word Score, the SPT Open to Closed Class Word Score, and the Noun vs. Verb Composite Score)*

	Grammatical Class Composite Score	Verb Generation		Name Agreement		Sentence Completion	
		Accuracy	Latency	Accuracy	Latency	Accuracy	Latency
WS	2.60	-	√	√	√	-	-
DD	2.04	-	-	-	-	-	-
JHM	1.32	√	-	-	√	-	-
PS	1.16	-	-	-	√	-	-
FS	1.14	-	-	-	-	√	-
NS	1.06	-	-	√	√	n/a	n/a
CT	1.02	-	√	√	√	-	-
DA	1.00	-	√	-	√	-	-
XX	0.98	-	√	-	-	√	-
NP	0.97	-	-	-	√	-	-
TB	0.96	-	-	-	√	-	√
ECV	0.87	-	-	√	√	√	-
STR	0.87	-	-	-	√	-	√

We also explored the hypothesis that there may be a correlation between grammatical class composite scores and percentage change scores between the high and low competition conditions on the four experimental tasks (see Table 6.6). Looking at the category exemplar task data, there was a significant correlation between grammatical class composite scores, and percentage change scores. Also, looking at the name agreement task accuracy data, there was a trend towards a significant correlation between grammatical class composite scores, and percentage change scores ( $p = .06$ ). A high grammatical class composite score reflects relatively good open class word production relative to closed class word production, and relatively good noun production relative to verb production. Hence the observed correlation reflects that as participants have a grammatical class composite score more typical of non-fluent aphasia, they show a larger difference between the competition conditions.

Table 6.6

*Summary of Individual Results on Selected Tasks, Showing Percentage Change Values on Experimental Tasks Ordered by Grammatical Class Composite Score (the Average of the QPA Open to Closed Class Word Score, the SPT Open to Closed Class Word Score, and the Noun vs. Verb Composite Score), with Correlation Statistics*

	Open vs. Closed Class Words and Noun vs. Verb Composite Score (average across 3 sub- measures)	Category Exemplar	Verb Generation		Name Agreement		Sentence Completion	
		Number of Exemplars	Accuracy	Latency	Accuracy	Latency	Accuracy	Latency
WS	2.60		17	32	55	82	33	-2
DD	2.04	59	8	12	22	22	47	7
JHM	1.32	5	25	19	12	44	21	0
PS	1.16		9	27	5	71	13	6
FS	1.14	0	-4	0	37	30	36	-3
NS	1.06	13	7	33	56	47		
CT	1.02	30	-6	56	32	50	-11	13
DA	1.00	4	-8	67	14	59	16	1
XX	0.98	22	14	32	11	23	47	-4
NP	0.97	18	6	28	13	58	14	-1
TB	0.96	33	8	8	-2	34	8	4
ECV	0.87	24	29	19	31	49	38	3
STR	0.87	27	7	3	7	47	10	15
<i>Correlation Coefficient</i>		0.53	0.20	-0.04	0.45	0.27	0.39	-0.
<i>r</i>								
<i>df</i>		9	11	11	11	11	10	10
<i>One-tailed p</i>		0.047	0.26	0.45	0.06	0.19	0.11	0.2

*Nouns in Isolation vs. Nouns in Context.* For this analysis, the data were ordered by the participant's scores on the Lukkien (2006) Single Naming Test (SNT) divided by their scores for the key nouns on the Sentence Production Test (SPT) (high scores reflect a non-fluent pattern of responding with better production of nouns in isolation than nouns in context). For the verb generation and name agreement tasks there does not appear to be a clear relation between this score and the occurrence of significant competition effects on the key tasks (see Table 6.7). However, on the

sentence completion task, participants who had disproportionately poor noun production within sentences relative to their production of the same nouns in isolation appeared to be more likely to show accuracy differences. The participants who did not have disproportionately poor noun production within sentences appeared to be more likely to show latency differences.

Table 6.7

*Summary of Individual Results on Selected Tasks, Showing Where Significant Effects Were Observed on Simple Accuracy and Latency Analyses, Ordered by Nouns in Isolation vs. Nouns in Context Score*

	Nouns in Isolation / Nouns in Context	Verb Generation		Name Agreement		Sentence Completion	
		Accuracy	Latency	Accuracy	Latency	Accuracy	Latency
WS	1.89	-	√	√	√	-	-
DD	1.31	-	-	-	-	-	-
XX	1.29	-	√	-	-	√	-
CT	1.27	-	√	√	√	-	-
ECV	1.25	-	-	√	√	√	-
FS	1.25	-	-	-	-	√	-
NP	1.25	-	-	-	√	-	-
JHM	1.07	√	-	-	√	-	-
PS	1.07	-	-	-	√	-	-
NS	1.04	-	-	√	√	n/a	n/a
TB	1.03	-	-	-	√	-	√
STR	0.96	-	-	-	√	-	√
DA	0.90	-	√	-	√	-	-

We also explored the hypothesis that there may be a correlation between nouns in isolation vs. nouns in context scores and percentage change scores between the high and low competition conditions on the four experimental tasks (see Table 6.8). Looking at the name agreement accuracy data, there was a significant correlation between nouns in isolation vs. nouns in context score, and percentage change scores, indicating that better naming in isolation than naming in context was associated with larger percentage change values between the competition conditions in the accuracy data from the name agreement task. The nouns in isolation vs. nouns in context scores were chosen to reflect increased lexical load. It is interesting that these scores correlate with competition condition difference scores on the name agreement task, which involves a similar rationale – the lexical load is minimal in the

low competition condition, since only one term is used to describe the objects, and this load increases in the high competition condition, where more than one word can be used to describe the objects.



Table 6.8

*Summary of Individual Results on Selected Tasks, Showing Percentage Change Values on Simple Accuracy and Latency Analyses Ordered by Nouns in Isolation vs. Nouns in Context Score, with Correlation Statistics*

	Nouns in Isolation vs. Nouns in Context	Category Exemplar Number of Exemplars	Verb Generation		Name Agreement		Sentence Completion	
			Accuracy	Latency	Accuracy	Latency	Accuracy	Latency
WS	1.89		17	32	55	82	33	
DD	1.31	59	8	12	22	22	47	
XX	1.29	22	14	32	11	23	47	
CT	1.27	30	-6	56	32	50	-11	
ECV	1.25	24	29	19	31	49	38	
FS	1.25	0	-4	0	37	30	36	
NP	1.25	18	6	28	13	58	14	
JHM	1.07	5	25	19	12	44	21	
PS	1.07		9	27	5	71	13	
NS	1.04	13	7	33	56	47		
TB	1.03	33	8	8	-2	34	8	
STR	0.96	27	7	3	7	47	10	
DA	0.90	4	-8	67	14	59	16	
<i>Correlation Coefficient</i>								
		0.37	0.27	-0.01	0.57	0.27	0.40	
<i>r</i>								
		9	11	11	11	11	10	
<i>df</i>								
		0.13	0.19	0.49	0.02	0.19	0.10	
<i>One-tailed p</i>								

*Conceptual Fluency.* For this analysis, the data were ordered by the participant's letter fluency scores from the COWAT, divided by their BNT percentage correct scores. This measure was particularly sensitive to the non-fluent/fluent diagnostic distinction: The seven participants who were diagnosed with a non-fluent aphasia gained the seven lowest scores (including ECV and TB) (see Table 6.9). This measure taps into a different aspect of fluency than rate of speech measures, because the competition being measured does not involve syntagmatic competitors (i.e., competition from other items planned for the same utterance). Hence, it is possible that it is more relevant to the competition effects that we have examined in the current series of single word production tasks. Looking at the verb generation task data, it appears that the participants with good letter fluency/BNT percentage correct scores are unlikely to show significant competition effects. This finding suggests that the ability to use compensatory strategies during word production may limit the influence of competition resolution deficits: It is also consistent with the view that articulation

difficulties may only have minimal influence, since the trends are observed despite the minimal articulation demands present in the COWAT measures.

Table 6.9

*Summary of Individual Results on Selected Tasks, Showing where Significant Effects were observed on Simple Accuracy and Latency Analyses, Ordered by COWAT Letter Fluency over BNT Percentage Correct Score*

	COWAT Letter Fluency / BNT % Correct	Verb Generation		Name Agreement		Sentence Completion	
		Accuracy	Latency	Accuracy	Latency	Accuracy	Latency
DD	0.03	-	-	-	-	-	-
WS	0.04	-	√	√	√	-	-
JHM	0.07	√	-	-	√	-	-
ECV	0.08	-	-	√	√	√	-
CT	0.18	-	√	√	√	-	-
TB	0.19	-	-	-	√	-	√
DA	0.20	-	√	-	√	-	-
XX	0.30	-	√	-	-	√	-
PS	0.34	-	-	-	√	-	-
NS	0.41	-	-	√	√	n/a	n/a
NP	0.44	-	-	-	√	-	-
STR	0.50	-	-	-	√	-	√

We also explored whether there may be a correlation between COWAT letter fluency / BNT percentage correct scores and percentage change scores between the high and low competition conditions on the four experimental tasks. However, no significant correlations were observed (see Table 6.10).

Table 6.10

*Summary of Individual Results on Selected Tasks, Showing Percentage Change Values on Simple Accuracy and Latency Analyses Ordered by COWAT Letter Fluency Score over BNT Percentage Correct, with Correlation Statistics*

	COWAT Letter Fluency / BNT % Correct	Category Exemplar Number of Exemplars	Verb Generation		Name Agreement		Sentence Construction	
			Accuracy	Latency	Accuracy	Latency	Accuracy	Latency
DD	0.03	59	8	12	22	22	47	
WS	0.04		17	32	55	82	33	
JHM	0.07	5	25	19	12	44	21	
ECV	0.08	24	29	19	31	49	38	
CT	0.18	30	-6	56	32	50	-11	
TB	0.19	33	8	8	-2	34	8	
DA	0.20	4	-8	67	14	59	16	
XX	0.30	22	14	32	11	23	47	
PS	0.34		9	27	5	71	13	
NS	0.41	13	7	33	56	47		
NP	0.44	18	6	28	13	58	14	
STR	0.50	27	7	3	7	47	10	
<i>Correlation Coefficient</i>			-0.28	-0.35	-0.07	-0.23	0.06	-0.39
<i>r</i>								
<i>df</i>			9	10	10	10	10	9
<i>One-tailed p</i>			0.20	0.13	0.41	0.24	0.43	0.12

We also explored conceptual fluency without the strong phonological component present in the letter fluency measure discussed above. For this analysis, the data were ordered by the participant's category fluency scores from the COWAT, divided by their BNT percentage correct scores (see Table 6.11). Looking at the verb generation task data, it appears that the participants with high category fluency/BNT percentage correct scores are unlikely to show significant competition effects (i.e., those who are good at the COWAT conceptual fluency measure relative to their BNT score (the more typical fluent pattern)). This result is consistent with that observed on the analysis of the letter fluency measure discussed above.

Table 6.11

*Summary of Individual Results on Selected Tasks, Showing where Significant Effects were observed on Simple Accuracy and Latency Analyses, Ordered by COWAT Category Fluency (Animals Only) over BNT Percentage Correct Score*

	COWAT Category Fluency / BNT % Correct	Verb Generation		Name Agreement		Sentence Completion	
		Accuracy	Latency	Accuracy	Latency	Accuracy	Latency
WS	0.08	-	√	√	√	-	-
ECV	0.10	-	-	√	√	√	-
FS	0.12	-	-	-	-	√	-
DA	0.13	-	√	-	√	-	-
JHM	0.14	√	-	-	√	-	-
XX	0.18	-	√	-	-	√	-
CT	0.18	-	√	√	√	-	-
NS	0.18	-	-	√	√	n/a	n/a
PS	0.20	-	-	-	√	-	-
NP	0.21	-	-	-	√	-	-
DD	0.22	-	-	-	-	-	-
TB	0.22	-	-	-	√	-	√
STR	0.30	-	-	-	√	-	√

We also explored whether there may be a correlation between COWAT category fluency / BNT percentage correct scores and percentage change scores between the high and low competition conditions on the four experimental tasks (see Table 6.12). Looking at the name agreement data, a significant negative correlation was observed between the COWAT category fluency / BNT percentage correct scores, and percentage change values on the task. This reflects that as the participants became better at the COWAT category fluency component relative to their BNT scores (as is more typical of fluents), they showed a smaller percentage change between the competition conditions. However, there was also a trend towards a positive correlation observed in the category exemplar data (i.e., the category exemplar showed a strong trend towards a correlation in the opposite direction to what was expected).

Table 6.12

*Summary of Individual Results on Selected Tasks, Showing Percentage Change Values on Simple Accuracy and Latency Analyses Ordered by COWAT Category Fluency Score over BNT Percentage Correct, with Correlation Statistics*

	COWAT Category Fluency / BNT % Correct	Category Exemplar	Verb Generation		Name Agreement		Sentence Completion	
		Number of Exemplars	Accuracy	Latency	Accuracy	Latency	Accuracy	Latency
WS	0.08	22	17	32	55	82	33	
ECV	0.10	24	29	19	31	49	38	
FS	0.12	5	-4	0	37	30	36	
DA	0.13	4	-8	67	14	59	16	
JHM	0.14	18	25	19	12	44	21	
CT	0.18	30	-6	56	32	50	-11	
NS	0.18		7	33	56	47		
XX	0.18		14	32	11	23	47	
PS	0.20	27	9	27	5	71	13	
NP	0.21	13	6	28	13	58	14	
DD	0.22	59	8	12	22	22	47	
TB	0.22		8	8	-2	34	8	
STR	0.30	33	7	3	7	47	10	
<i>Correlation Coefficient</i>		0.50	-0.24	-0.30	-0.56	-0.29	-0.36	
<i>r</i>								
<i>df</i>		9	11	11	11	11	10	
<i>One-tailed p</i>		0.06	0.21	0.16	0.02	0.17	0.13	

*Summary.* The following specific components of fluency were examined to see if they might more clearly determine performance on the experimental tasks than the diagnostic non-fluent/fluent distinction from the BDAE: rate of speech; production rates of open to closed class words and nouns to verbs; the ability to produce nouns in isolation compared to the ability to produce nouns in context; and conceptual fluency. Across the four fluency measures, the rate of speech and conceptual fluency measures appeared to predict who would show significant competition effects to some degree on select tasks, and these scores correlated with competition condition difference scores on select tasks. On the verb generation task, participants with slow speech production scores relative to their naming scores appeared to be more likely to show significant competition effects, plus slower rate of speech scores were correlated with larger differences in latency scores between the competition conditions. However, in the sentence completion task, participants with

fast rate of speech scores relative to their naming scores appeared more likely to show significant competition effects. It is not clear why these different patterns were observed on the different tasks, but they may reflect the influence of extraneous variables present in the tasks. The two fluency measures relating to grammatical class effects (rates of open to closed class words produced and rates of noun and verb production), did not appear to be related to the occurrence of significant competition condition effects: When the data was ordered by these measures, no clear groups of participants showing significant competition effects were observed. However these measures may provide some insight into which participants are more likely to show accuracy or latency effects on select tasks, because on the sentence completion task, participants with poor naming in context scores relative to their naming in isolation scores appeared to be more likely to show accuracy effects, whereas participants with good naming in context scores relative to their naming in isolation scores appeared to be more likely to show latency effects. Because non-fluent aphasics dissociate in their performance on the four fluency component measures explored, it is possible that scores on the more specific fluency measures may more accurately predict who might show competition condition effects in future research, rather than the multi-factorial BDAE non-fluent/fluent diagnostic distinction used in the current studies.

Across the four fluency measures, the rate of speech and conceptual fluency measures also appeared to predict the size of the differences observed between the competition conditions to some degree on select tasks. Further, four of the five significant correlations observed across the four tasks corresponded with the interpretation that non-fluent performance leads to larger percentage change values between the competition conditions. On the verb generation task, slower rate of speech scores were correlated with larger differences in latency scores between the competition conditions. On the category exemplar task, better scores in naming nouns than verbs and open than closed class words were correlated with larger differences between the competition conditions. On the name agreement task, better scores in naming nouns in isolation than naming nouns in context were correlated with larger differences in accuracy between the competition conditions, as were low COWAT category fluency scores / BNT percentage correct scores. However, the opposite pattern was observed on the sentence completion latency data, where a correlation was observed between the fluent pattern of good category fluency / BNT percentage correct scores, and large differences between the competition conditions.

It is not clear why these different patterns were observed on the different tasks, but again, the patterns may reflect the influence of extraneous variables present in the tasks. The fluency measures involving single word production (rates of open to closed class words produced and rates of noun and verb production, and the COWAT category fluency / BNT percentage change scores) appeared to be related to large differences between the competition conditions on the category exemplar, name agreement, and sentence completions tasks in particular. In contrast, the fluency measures relating to sentence level processes (QPA / BNT percentage correct score, QPA, and nouns in isolation vs. nouns in context scores) appeared to be related to large latency differences on the verb generation task, and accuracy differences on the name agreement task. Hence, these correlation results provide more evidence suggesting that the more specific fluency measures explored here may more accurately predict who might show competition condition effects in future research, than the multi-factorial BDAE non-fluent/fluent diagnostic distinction used in the current studies.

## General Discussion

In this project, four experiments were conducted in order to examine whether some or all non-fluent aphasics are more sensitive to lexical competition than fluent aphasics or controls. Each task contained a low competition condition, where one or few potential responses were available, and a high competition condition, where multiple potential responses were likely to become activated. The rationale behind these manipulations was that if a controlled activation process biases activation towards one possible response, it should aid lexical selection in situations where competition is unusually high, but be less important in situations where there is already one dominant response. Further, participants who experience impaired controlled activation should be particularly susceptible to performance costs in high competition conditions since the controlled activation process will be more heavily taxed in these conditions.

The four experimental tasks included a category exemplar task, a verb generation task, a name agreement task, and a sentence completion task. In the category exemplar task, participants generated items from broad groups (e.g., *Animals*: “Cat. Dog. Mouse” etc.), or narrow groups (e.g., *Animals: Pets*: “Cat. Dog” etc., then *Farm Animals*: “Cow. Sheep” etc., then *Water Animals*: “Dolphin. Badger” etc., then *Jungle Animals*: “Monkey. Snake”, etc.). We reasoned that because the broad groups were bigger, the increased number of possible exemplars would lead to activation of a greater number of potential candidates, and consequently, higher overall levels of lexical competition. In the verb generation task, participants were presented with objects and were asked to generate a related verb. The objects consistently elicited one dominant verb (e.g., *Scissors*: “Cut”) or they elicited a range of verbs (e.g., *Penny*: “Spend”/“Pay”/“Buy”/“Drop”, etc.). We reasoned that objects that elicit a range of verbs would be associated with higher levels of lexical competition. In the name agreement task, naming of objects with low name agreement (e.g., *Artist/Painter*) was compared to naming of objects with high name agreement (e.g., *Anchor*). We reasoned that because more than one possible name was available for low name agreement items, relatively strong lexical competition would be present. In the sentence completion task, extrinsic competition was introduced via presentation of auditory distracters. In the low competition condition, the auditory distracters did not make sense (e.g., *Barry wisely chose to pay*



*the RANGE*: “Bill”/“Cashier” etc), whereas in the high competition condition, the auditory distracters were plausible completions of the sentence (e.g., *Barry wisely chose to pay the FINE*: “Bill”/“Cashier” etc). We reasoned that the activation level of the distracter would be greater in the high competition condition, because it would receive activation not just from the auditory distracter stimulus, but also via its association with the sentence stem.

Below, the key findings from each task are summarised.

*Category Exemplar Task.* At the group level, all groups of participants produced fewer exemplars in the high competition (broad category) condition than low competition (narrow category) condition. However, the hypothesis that non-fluents would show larger competition condition effects than fluents and controls was not supported. Compared to the data from the other tasks, the data at the individual level was unusually consistent: A trend towards a low lexical competition advantage occurred in all but one aphasic participant, but in no case was this trend statistically reliable (see the first column in Table 6.1).

The consistency of the data trends across participants may be the result of the specific demands of this task, which could be considered the most multi-factorial of the four tasks completed. For instance, this task is likely to require some general knowledge on each category so that participants could produce names of items belonging to the group; they also need to be able to remember and monitor what they have already said to avoid repeating the same item names; and they might also benefit from being able to efficiently switch between clusters of related items. Any, or all, of these processes might be more heavily engaged in the high than the low competition condition, so the trends observed may be underpinned by different factors in different individuals. Indeed, the “competition” effect observed in the overall group analyses may actually reflect the contributions of these other processes, rather than competition *per se*.

*Verb Generation Task.* At the group level, the hypothesis that the high competition condition (many possible verb responses) would generally result in higher error rates and/or longer response latencies than low competition condition (few possible responses) was supported: A significant effect was obtained in both the accuracy and the latency analysis. However, at the group level the hypothesis that non-fluents would show larger lexical competition effects than fluent aphasics and controls was not supported. At the individual level there was limited support for this

hypothesis. Analysis of the individual participant's data revealed considerable heterogeneity within each group, but if you look at simple prevalence rates, it appears that non-fluents were more likely than fluents to show significant condition effects (see Table 6.1: four non-fluents showed a significant accuracy or latency effect, whereas only one fluent showed a significant accuracy or latency effect). Of the four experimental tasks, verb generation appears to distinguish between the non-fluent aphasics and the fluent aphasics/controls the best. The data also provided partial support for the hypothesis that non-fluent aphasics with relatively well preserved single word production but relatively impaired sentence production (non-fluents DA, ECV, and JHM), might be most likely to show exaggerated lexical competition effects. DA and JHM were amongst the three aphasic participants showing the largest differences between the conditions, although ECV was not.

Nevertheless, in the verb generation task, it is not entirely clear that the significant condition effects observed did indeed reflect differences in the levels of lexical competition across the two conditions. In the literature, there has been considerable controversy regarding the cognitive basis of effects of this kind. Some argue that the poorer performance observed on items with more potential responses reflects competition (Thompson-Schill et al., 1997; Thompson-Schill et al., 1998) (this view has been explored by comparing items with high and low response strength ratios, which reflect the proportion of the most common response to the second most common response; the relative strength of the two most common responses is thought to reflect the strength of competition present). However, others suggest that for these items, retrieval of any item is more effortful, because nothing specific comes to mind immediately (Martin & Cheng, 2006) (this view has been explored by comparing items with high and low association strength ratios, which reflect the proportion of the most common response to all other responses combined). In the current study, an exploratory analysis that took into account association strength and response strength indicated that association strength differences are potentially driving the "lexical competition" effects we observed rather than selection between competing items. However, our exploratory analysis was based on data from few stimuli and few participants. Also, it is generally difficult to un-confound response strength and association strength in this task. Indeed, it is possible that selection and controlled retrieval reduce to a common mechanism. For instance, a modulation process might function to resolve situations where the best response is ambiguous because the

potential responses have low response strength and there is a need for selection between viable alternatives: The same modulation process might be used when there is no clear response available (low association strength) and there is a need for the controlled retrieval of a suitable response.

*Name Agreement Task.* At the group level, the hypothesis that the high competition condition (low name agreement) would result in higher error rates and/or longer response latencies than the low competition condition (high name agreement) was supported by the accuracy and latency data. At the individual level, there was mixed support for this hypothesis: some of the participants showed corresponding competition effects and some did not, but no participant showed a significant effect in the unpredicted direction.

At the group level, the hypothesis that non-fluents would show larger lexical competition effects than the fluents and controls was not clearly supported. At the individual level, some differences in the prevalence of significant effects were observed in the accuracy data: Three of the seven non-fluents showed a low competition condition advantage, but only one of the six fluents did so. However, in the latency data, there was a strong trend for participants from both aphasic groups to show significant competition effects: Six of the seven non-fluents did so, as did four of the six fluents. Although there appears to be fluency based prevalence differences in the accuracy data, the high prevalence of latency differences in both groups weakens the strength of the observation that non-fluents may be particularly susceptible to competition effects because most of the fluents also show significant effects. This pattern highlights the importance of considering both accuracy and latency data, since speed accuracy tradeoffs can influence performance on either measure and the picture that emerges when both measures are considered together can differ from that which emerges when they are considered separately.

*Sentence Completion Task.* At the group level, the hypothesis that the high competition condition (plausible completion provided) would result in higher error rates than low competition condition (implausible completion provided) was supported. The hypothesis that non-fluents would show larger lexical competition effects than fluent aphasics and controls was not clearly supported in the accuracy or latency data. At the individual level, most participants showed a trend towards slower responses and/or more errors in the high competition condition, but this did not always reach significance, nor did the incidence of reliable effects appear to clearly

differ between the fluent and non-fluent aphasics. Indeed, contrary to expectations, the fluents had a marginally higher prevalence of significant competition effects in accuracy and/or latency than the non-fluents.

Nevertheless a difference between the aphasic groups was observed in the supplementary distracter error analysis. In this analysis, we counted how often the distracter names were erroneously produced as responses (e.g., *Barry wisely chose to pay the FINE*: “Fine”). We found that the non-fluent group made more distracter errors than the fluent group. Also, more distracter errors were made in the high competition condition than the low competition condition. This result suggests that there may be a qualitative difference in the way the groups respond on the task, and that the non-fluents may find it more difficult to disregard the distracter than the fluents.

*Exploratory Fluency-Component Analyses.* In addition to the task-specific analyses summarised above, we also explored whether specific components of the fluency dimension were associated with particular performance patterns across the three experimental tasks which yielded the most reliable competition effects (the verb generation, name agreement, and sentence completion tasks).

Several interesting patterns were observed relating to the verb generation task. Participants whose COWAT verbal fluency scores were poor relative to their single word naming performance on the BNT were the most likely to show significant competition effects. This measure was investigated because it reflects the participant’s cognitive fluency while taking into account the severity of their naming deficit, while keeping syntactic processing minimal. The trends shown are consistent with the observation from the verb generation task proper, that non-fluents may be more likely to show significant competition effects than fluents and controls on this task. In addition though, the results suggest that the pattern observed may not be a direct result of the participant’s word production deficit severity because naming performance was factored into the score. This result is of interest because there was some concern in the initial verb generation analyses that those participants who showed the largest response latency differences between the competition conditions also had the longest response latencies overall. A similar pattern was evident when participants were ordered according to their rate of connected speech (QPA words per minute) relative to their BNT scores - a measure which does involve syntactic processing. For the verb generation task, these two measures appeared to be more

effective predictors of significant effects than other fluency-related measures, such as relative success scores on open vs. closed class words and nouns vs. verbs, and nouns in isolation scores vs. nouns in context scores.

Several interesting patterns were also observed for the sentence completion task. In contrast to the verb generation task, those who scored highly on the speech rate measure where their QPA words per minute score was divided by their object naming score on the BNT actually appeared to be *more* likely to show significant competition effects than participants who scored low on this measure.

It is possible though, that both the verb generation and sentence completion tasks have score severity windows: Participants with a certain level of impairment may be good enough that severity related issues did not prevent observation of significant effects, but impaired enough to show behavioural consequences that are measurable using the current paradigms. This issue may be particularly relevant in the case of the sentence completion task, because the participants' naming in context/naming in isolation scores on this task (see p. 164), which were analysed because of their sensitivity to aspects of grammatical processing dysfunction, appear to be somewhat predictive of the type of competition effects that the participants show, while not necessarily predicting who will show competition effects overall. For instance, low-mid range naming in context/naming in isolation scores appear to be somewhat predictive of participants showing accuracy effects (e.g., see XX, ECV, and FS, Table 6.5, p. 164). Participants with mid-high range naming in context/naming in isolation scores do not appear to show accuracy or latency effects. In comparison, high naming in context/naming in isolation scores appear to be somewhat predictive of participants showing latency effects (e.g., see TB and STR, Table 6.5, p. 164). To illustrate this point, consider the possibility that these latency effects were significant because these participants had response latencies that were less variable than some of the others. Also, because these participants had relatively low error rates, more latencies were available for analysis. Thus, the competition effects observed may reflect relatively good performance on the task, as well as the influence of the competition manipulation. Hence it is possible that the participants were not more strongly influenced by the competition manipulation than some or all of the other participants, but that the participants who have more severe or less severe scores than those mentioned above, perform in a manner that is not conducive to observing significant competition effects in accuracy and latency.

### *Overall Results Summary*

In summary, our findings showed that across all participant groups and tasks, the high competition conditions generally resulted in higher error rates and/or longer response latencies than low competition conditions. However, the non-fluent group did not show reliably larger competition effects than the fluent group in any of the tasks. This finding is consistent with two recent studies that have reported selection/control and executive control deficits in patients with anterior and posterior lesions in linguistic (Jefferies & Lambon Ralph, 2006) and spatial tasks (Peers; Ludwig, Casimir, Rorden, Cusack, Bonfiglioli, Bundesen, Driver, Antoun, Duncan, 2005), and with the observation in imaging studies that posterior temporal and/or inferior parietal regions show activation. The only group analysis where the non-fluents performed differently from the fluents was the supplementary distracter error analysis that was completed on the sentence completion task results, in which it was observed the non-fluents repeated the distracter as their response (e.g., Barry wisely chose to pay the FINE: “Fine”) significantly more frequently than the fluents. In the verb generation, name agreement, and sentence completion tasks, some individual participants showed reliable competition effects and some did not. As hypothesised, the diagnostic non-fluent/fluent distinction appeared to be somewhat predictive of these effects on the verb generation task, because more non-fluents than fluents showed significant competition effects: However, some specific components of fluency (rate of speech and cognitive fluency) appeared to be more predictive of these effects than others (grammatical processing measures such as noun vs. verb production rates). Also, a less clear, but contradictory distinction was observed in the sentence completion task data, where marginally more fluents than non-fluents showed significant competition effects (this may be due to the extraneous sentence level processing required for this task, however, further research is required before this can be established).

Overall, there was also little clear support for the hypothesis that those non-fluent aphasics with relatively well preserved single word production but relatively impaired sentence production might be most likely to show exaggerated lexical competition effects. The participants who fitted this profile were over-represented in the group of participants showing the largest and most consistent competition effects, but there was a lack of cross-task consistency: These individuals showed large effects on some tasks but not others. This suggests that some other, as yet unknown, feature

may have influenced the participants' performance, but again, the exploratory analysis of individual fluency components suggests that characteristics such as rate of speech production and conceptual fluency may be more predictive of competition condition effects than more grammatical aspects of fluency, such as rates of open and closed class word production.

### *Theoretical Implications*

In each of the experimental tasks competition effects were observed. These findings are problematic for the Dell (1986) and MacKay (1987) models of speech production, because these models do not clearly predict competition effects. In these models, spreading activation activates nodes containing relevant information, and the most highly activated nodes are selected for production. If two nodes are activated to the same level for the same purpose (i.e., if there is a high level of lexical competition), it is not clear which node will be selected. It is possible that over time one node may retain/accrue more activation than the other and will eventually be selected. The limitation of this explanation is that it does not discriminate competition effects from other effects that result in increased reaction times, such as word frequency effects (see Arpita, van Lieshout, & Square, 2007). As was illustrated in the verb generation task, competition effects and frequency effects can dissociate: This suggests that different processes may be relied on to resolve these disparate effects.

At the individual level, it was observed that although participants showed significant competition effects on specific tasks, none consistently showed significant effects on all four tasks. One potential reason for these inconsistencies is that the different tasks may place more demand on a single level of processing (i.e., the semantic, lexical, or phonological level). For example, in the verb generation task, competition arises between semantically distinct verbs (*Cat* = "Purr" / "Scratch" etc.), hence competition resolution must occur at the semantic level. Semantic level competition will also dominate in the category exemplar and sentence completion tasks: In the category exemplar task potential exemplars will compete, and in the sentence completion task competition will occur between the auditory distracters and potential responses. In contrast, in the name agreement task, lexical competition could result but semantic competition would be minimal, because during the

production of low name agreement (LNA) items, the semantic description is identical but the words used to describe the items are competitors (e.g., “Couch” / “Sofa”).

Logically, a participant with a deficit to the lexical level but not the semantic level may be more likely to show a competition effect on the name agreement task than the verb generation, sentence completion, or category exemplar tasks. Unfortunately, it is difficult to clearly assess this possibility because aphasics often have deficits to more than one level of processing, plus cascading activation means that deficits at one level can also influence processing at other levels. Still, four of the current aphasic participant pool performed within the normal range on the tests of single word comprehension included in our diagnostic test battery (DA, DD, JHM, and NP), indicating that their semantic level processing is relatively intact. Two of these participants did show competition effects on the verb generation task (DA, JHM), where you would expect their deficits to have minimal effect. Three of these participants showed competition effects on the name agreement task (DA, JHM, and NP), where you would expect their deficit to have a more robust effect. However, other participants also showed similar effects on the name agreement task (CT, ECV, TB, WS, NS, PS, and STR), perhaps due to flow-down effects from impaired semantic processing, or because their poor performance on semantic tests reflects something other than impaired semantics, such as picture matching difficulties.

To conclude, the process of resolving deficits at specific levels of processing to performance on these tasks is problematic. However, one potential avenue that could be explored in future studies is to examine whether performance on different skills that reflect intact/deficient performance on particular processes that are included in language models (such as types of naming errors, picture-word matching scores, sentence comprehension scores, function word omission scores, and span tasks) are correlated with the size or occurrence of competition effects on tasks that target competition at theoretically driven stages of speech production.

### *Suggestions for Future Research*

Despite the observation of mixed, at best, support for the selection and control hypothesis in the current tasks, several patterns were observed which might direct future research down more informative routes. This final section discusses several ideas for future research that have been identified as a result of the current



experiments, and how these ideas may be useful in terms of addressing some unanswered questions regarding the selection and control hypothesis.

The participants' performances on the verb generation task dissociated along the non-fluent/fluent dimension more frequently than their performances on the other tasks. However, it is not clear why this is the case. Items which have more than one plausible response may indeed induce greater competition for selection than items with one dominant response. However, an alternative possibility is that for at least some of these items, there is simply no obvious response that comes to mind. In other words, the stimulus is not strongly associated with any particular response (see Martin & Cheng, 2006). In this case participants might be required use a conscious search strategy (for instance, if they are not able think of what an item does, they may then change their approach and try to think about what we do with the item). This may also result in a larger and/or more consistent change in behaviour, making it easier to measure.

One possible approach that could be used in the future to tease apart why competition condition effects more successfully distinguished between our participant groups on the verb generation task than the other tasks is to use a naming to definition paradigm: Participants could be asked to provide examples of names to highly constraining definitions (e.g., A general name given to male siblings: "Brothers"), or weakly constraining, underdetermined, definitions (e.g., A type of animal: "Cat"/ "Dog"/ "Elephant" etc.). One advantage of this paradigm would be that it might provide an additional avenue for exploring the selection vs. controlled retrieval debate discussed in the previous paragraph. It gives potential for the development of a better selection of test items that fit the three categories explored by Martin and Cheng (2006) in their verb generation task (i.e., items that are underdetermined (low response strength) but highly associated, items that are underdetermined but weakly associated, and items that are highly determined (high response strength)). Another advantage of this paradigm would be that the stimuli are less likely to automatically generate a response by virtue of a strong semantic association alone. This would minimise some concerns that arise in verb generation, such as how to account for competition from non-verb associates (e.g., where for the item *Nurse*, the non-verb "Doctor" competes with related verbs such as "Heal", "Care", or "Treat"). Finally, the paradigm also provides an avenue through which results that converge with those from verb generation tasks might be observed.

Another paradigm that may provide a productive avenue for future research on lexical competition effects is picture-word interference (see Wilshire et al., 2006). In this paradigm participants are asked to name pictures that are presented simultaneously with auditory distracters (e.g., they could be shown a picture of a *Dog*, and hear the word *Duck*). In this situation, competition from an external source is biasing activation towards particular nodes, as it is in the sentence completion task. Three reasons why this paradigm may be useful are described next.

One reason why the picture-word interference paradigm may be useful is because it relies on extrinsic competition. This may be an important task feature because the other analysis in the current studies where a non-fluent/fluent distinction was observed was the group level sentence completion distracter analysis: The non-fluents made more distracter errors than the fluents in both competition conditions (both groups made more distracter errors in the high than the low competition condition). What was it about the sentence completion task that led to this pattern? One unique task feature was that competition was created extrinsically by exposing participants to distracters. The external presentation of the distracters presumably boosted activation of the corresponding lexical nodes. Potentially, the non-fluents found this extrinsic competition particularly hard to overcome. In activation terms, perhaps the other groups were able to use strategies to inhibit the distracters pre-presentation. If they were able to consciously ignore the distracters the impact of distracter presentation might be reduced. However, this possibility is still speculative because this was the only task with a design that allowed for this kind of analysis. Still, the data do identify a qualitative difference between the groups that, if further investigated, may be informative in terms of the patterns of breakdown present.

A second reason why the picture-word interference paradigm may be useful is because the use of distracters in the task gives the experimenter more control over where competition is coming from, and how much competition can be expected from given distracters. It must be acknowledged that at a general level the results of the current experimental tasks do not provide unequivocal support for the selection and control hypothesis. There is a very real possibility that the notion of lexical competition may not be the best way to characterise the kinds of word production problems seen in non-fluent aphasia. It is possible that a non-competitive controlled retrieval process may more accurately describe the variations observed amongst fluent and non-fluent aphasics. At least on some tasks (e.g., verb generation), competition

condition effects may reflect differences in the extent to which other processes, such as effortful searches, are required. One way to explore whether controlled retrieval or competitive processes potentially explain the cognitive processes involved most accurately is to focus on paradigms that explicitly focus on either association strength or competition. However, it is not clear how you would operationalise a task that specifically manipulates controlled retrieval/association strength: This is particularly difficult because the concept is less specific than that of competition, and hence it is more difficult to isolate and manipulate. Nevertheless, some aspects of the data do hint that the selection and control hypothesis may be worth pursuing using different types of paradigms and manipulations, and there is potential to further isolate the influence of competition in picture-word interference because, as was mentioned previously, the examiner can control where competition is coming from, and there is also more scope to analyse the effects of this competition than there is on many other tasks that manipulate competition (e.g., through distracter error analyses).

A third reason why the picture-word interference paradigm may be useful is because it may be possible to use it to address the concern, to some level at least, that tasks we developed to explore competition were not sufficiently “pure” to highlight genuine differences in lexical competition effects between the key aphasic groups. It is possible that the similar behavioural manifestations present for both groups might result from different functional deficits. Further, current models of language and current models of executive function do not clearly predict how the different components of each task might influence performance, or how impairments to various cognitive functions may interfere with performance. This makes it difficult to isolate and control for the different factors that might contribute to the observed results. Systematically isolating and manipulating different task components may provide more insights on this matter though, and lead to more specified theories. For instance, as was discussed above, the influence of extrinsic competition from distracter stimuli could be investigated further in a picture-word interference task (see Wilshire et al., 2006). Also, a picture-word interference paradigm would provide another opportunity to analyse the influence of auditory distracters in a task that, relative to the current sentence completion task, further minimises the amount of sentence level processing required, and hence minimises any extraneous influences from sentence level processes.

Another possible reason why we did not observe the predicted significant differences and interactions at the group level on our tasks, which could be addressed in future research, is that the participants within our different comparison groups varied so widely. This wide intra-group variability is not only likely to have reduced our chances of obtaining significant group differences, but it may also have had some more direct statistical consequences: for example, the assumption of homogeneity of variance that underlie most parametric statistical methods is likely to have been violated. However, despite the failure to observe predicted group differences in the current research, heterogeneity of participants within groups does not always prevent the observation of significant group differences in aphasia research. To the best of our knowledge, the cyclic blocked naming task by Schnur et al. (2006) is the only group study that has been completed on context-specific naming deficits in aphasia where significant interactions between aphasic groups have been observed. The Schnur et al. study used a different criteria to select participants than the current study and had some groups that were more homogeneous than ours (for example, their non-fluent group was limited to Broca's aphasics only): Their participant selection criteria may have contributed to the significant results observed by decreasing the response variability present. The size of the groups in Schnur et al's (2006) study also differed from ours as they had as many Broca's as we had non-fluents, but they had 11 fluents whereas we had 6. This increased number of participants may have contributed to the significant effects observed, since it would have increased the total number of data points, and statistical power available (it is also worth noting that their paradigm also gave a large number of data points from each participant since there were many trials included, which would similarly influence the statistical analyses used).

In summary, despite the limited support for the selection and control hypotheses from the current tasks' results, the idea of lexical competition in non-fluent aphasia might be worth pursuing further using more tightly constrained task designs. This could be achieved in several ways. For example, one potential way forward is to use more specific participant selection criteria. We observed that rate of speech and conceptual fluency measures appeared to be more predictive of competition effects than fluency measures that focus on grammatical processes, such as relative rates of noun and verb production or open to closed class word production. Potentially, if participants groups were defined by scores on these particular components of fluency, rather than diagnostic categories such as the BDAE's non-

fluent/fluent diagnostic distinction, cleaner distinctions between group results may be observed and provide more informative results. A second way to move forward could be to use paradigms that target specific types of competition. The current data suggest that paradigms using extrinsic competition or underdetermined responses may be particularly promising because these types of paradigms appeared to discriminate most effectively between our participant groups. Possible avenues for future research that could be used to do this are to use picture-word interference or naming to definition paradigms.

Our findings also emphasise the many complexities of lexical selection, which were reflected in the way that individual participants performed differently on the four experimental tasks completed, even though the tasks were all designed to manipulate levels of lexical competition. This highlights the need for models of tasks that provide finer-grained specifications of the particular competition and modulation mechanisms that various tasks may require. Once we understand what precise mechanisms are required in various tasks it may be possible to design tasks that more directly investigate the influence of hypothesised task components, such as competition from different sources, cognitive control, and lexical selection.

## Appendix A1: Category Exemplar Pilot Task Instructions

[Pilot: High Competition Condition]

### Category Exemplar Task

#### Instructions:

Please do not turn the page until you are instructed to do so.

Overleaf is written the name of a category. Your task is to write down as many examples of items that belong to that category as possible, during a limited time period. Following is an example of what you need to do.

Category: 'Types of Vehicle'

Possible responses: Car  
Truck  
Train  
Bus  
Plane  
Helicopter  
Bicycle  
Cable car  
Tractor

etc... Continue writing down examples until you run out of time, or you cannot think of anymore.

There are several categories to complete, each on a separate page. Please do not move on to the next category until you are instructed to do so. When instructed to turn the page, please do so immediately.

Are you ready?

[Pilot: Low Competition Condition]

### Category Exemplar Task

#### Instructions:

Please do not turn the page until you are instructed to do so.

Overleaf is written the name of a category. Your task is to write down as many examples of items that belong to that category as possible, during a limited time period. To guide your recall, your instructor will verbally tell you sub-categories that you should use to guide your recall of objects. You are required to write down as many items in each sub-category as possible during the period directly after your instructor has told you the sub-category. You need to write down items that belong to the given sub-category only. Once your instructor has told you a new sub-category, you must move on and start writing down items that belong to the new sub-category only. Following is an example of what you need to do.

Category: 'Types of Vehicle'

Sub-category One (from your instructor): 'Vehicles that Fly'

Possible responses: Plane  
Helicopter  
747  
Cessna  
Glider

etc... Continue writing down examples until you run out of time, or you cannot think of anymore.

Sub-category Two (from your instructor): 'Farm Vehicles'

Possible responses: Tractor  
Quad bike  
Ute

etc... Continue writing down examples until you run out of time, or you cannot think of anymore.

There are several categories to complete, each on a separate page. Please do not move on to the next category until you are instructed to do so. When instructed to turn the page, please do so immediately.

Are you ready?



## Appendix A2: Descriptions of the Tests Used for Diagnostic Purposes

### *General Tests*

#### *The Boston Diagnostic Aphasia Examination, Third Edition (BDAE)*

(Goodglass, Kaplan, & Barresi, 2001) is a test battery that can be used to identify which, if any, of the classic aphasia syndromes best describes a language deficit (e.g., Broca's Aphasia, Wernicke's Aphasia, etc.). The diagnoses are based on score ratings for the following eight skills: 1) articulatory agility; 2) melodic line (prosody); 3) the length of phrases produced in running speech; 4) the variety of grammatical forms used in running speech; 5) the prevalence of paraphasias in running speech; 6) word finding ability relative to speech fluency; 7) the ability to repeat sentences, and; 8) auditory comprehension (including single word discrimination, the ability to follow commands, and the understanding of complex ideational material).

#### *The Psycholinguistic Assessments of Language Processing in Aphasia*

(PALPA) battery consists of 60 tests that are designed to assess different aspects of language processing (Kay, Lesser & Coltheart, 1992). From this battery, three reading and two repetition tasks were administered. These tests are discussed in more detail in the reading and repetition sections below.

### *Comprehension Tests*

#### *The Peabody Picture Vocabulary Test (PPVT)* assesses auditory

comprehension of single words using non-verbal responses (Dunn & Dunn, 1997). In this test participants are shown a card with four pictures on it (e.g. a broom, a bell, an elephant, and a bus). Then the examiner asks the participant to point out a specific picture (e.g., "Show me the bus"). The items get increasingly difficult throughout the test.

*The Pyramids and Palm Trees Test* assesses semantic access without requiring a linguistic response (Howard & Patterson, 1992). The three-picture version of the test was used. In this version, participants are shown a picture (e.g., spectacles) and are required point to a semantically associated target picture (e.g., an eye) rather than a distracter picture (e.g., an ear).

*The Philadelphia Comprehension Battery (PCB)* consists of four subtests: 1) Lexical Comprehension, 2) Sentence Comprehension, 3) Grammaticality Judgments, and 4) Synonymy Triplets (Saffran, Schwartz, Linebarger, Martin & Bochetto, 1988). The *Lexical Comprehension subtest* assesses single noun comprehension. In each

trial, participants are shown four pictures and are asked to indicate the picture that the examiner names. The distracter pictures include items that are semantically, phonologically, and perceptually related to the target picture. For example, the semantically related distracters for the target ‘Octopus’, are ‘Crab’, ‘Walrus’, and ‘Shark’; the phonologically related distracter for the target ‘Steak’ is ‘Snake’ (the other distracters are ‘Chair’ and ‘Knife’); and the perceptually related distracter for the target ‘Shovel’ is ‘Fork’ (the other distracters are ‘Tie’ and ‘Flamingo’). Poor performance on items with semantic distracters can suggest that lexical/semantic access is impaired. Poor performance on items with phonological distracters can suggest that phonological processes are impaired. Poor performance on trials containing perceptually related distracters can indicate that visual disturbances are present that may influence performance overall. The *Sentence Comprehension subtest* assesses auditory comprehension of sentences. For each trial, participants are shown two pictures and are read a sentence that describes one of the pictures. They are asked to point to the picture that was described in the sentence. Half of the test items include reverse role distracters (e.g., for the sentence “The girl who washed the boy was talkative”, there is a target picture of a girl washing a boy and a distracter picture of a boy washing a girl): These items are designed to assess syntactic comprehension. The remaining items include distracters with extraneous lexical items (e.g., for the sentence “The woman that photographed the man was friendly”, there is a target picture of a woman taking a photograph of a man and a distracter picture of a woman handing a man a plate of food): These items are designed to assess lexical comprehension. The *Grammaticality Judgments Subtest* assesses sensitivity to information that is structurally encoded in sentences, and the *Synonymy Triplets subtest* is a difficult test of semantic access: However, these sub-tests were not included in the diagnostic testing for this study.

#### *Production Tests*

*The Boston Naming Test (BNT)* assesses confrontation naming (Goodglass et al., 2001). Participants are shown 60 pictures of increasing difficulty and are asked to name each picture, with the aid of semantic and phonological cues if required. This test forms part of the BDAE.

*The Noun/Verb Naming Test* assesses single noun and single verb production (Zingeser & Berndt, 1988). In this test, participants are shown a set of 30 pictures depicting some sort of action, and are asked to tell the examiner what action is shown

in each picture (e.g., “Shoot”, “Sharpen”). Participants are also shown two sets of 30 pictures depicting objects, and are asked to name each picture (e.g., ‘Sun’, ‘Stove’). The first set of object pictures is matched to the verb set on cumulative frequency (the combined frequency of all morphological variants of the word combined, e.g., shoot, shoots, shooting, shot). The second set of objects is matched to the verb set on base frequency (the frequency of the base word only, e.g., shoot). This test can indicate dissociations between noun and verb naming ability.

*The Controlled Oral Word Association Task (COWAT)* is a test of verbal fluency (Spreeen & Strauss, 1991). In this task, participants are given a limited time period to say as many words as possible that meet specified conditions (e.g. words that start with a given letter, or words that belong to a semantic category). We used five different groups of items: words starting with F, A, and S, Animals, and Fruits. Participants were given one minute for each group of items.

*The Sentence Production Test* (Lukken, 2006) is a quantitative assessment of sentence production. In this test, participants are shown line drawings of 30 scenes, and are asked to describe each scene in one sentence (e.g., “The dog is swimming”, “The fairy is giving a crown to the girl”). Transitive, single object, double object, passive, and embedded target sentences are included in the test. Neurologically intact individuals show a minimum of 80 % agreement on their scene descriptions. The quantitative measures included in this test include: a complete sentence score, which reflects the number of sentences produced that are identical to the target; a correct syntax score, which reflects the number of sentences that are produced with syntax that is identical to the target, not penalising for content word errors; and scores for the correct production of specific items within the sentences, such as a nouns, root verbs, and a function words. Further, the *Single Noun Test* (Lukken, 2006) presents pictures of 35 objects that are also depicted in the Sentence Production Test scenes (e.g., “Dog”, “Fairy”). The objects are tested in a single-word naming paradigm, allowing for comparison between single-word naming and naming within sentences.

*The Qualitative Production Analysis (QPA)* (Saffran, Berndt, & Schwartz, 1989) is a quantitative assessment of spontaneous speech. In the test, participants are asked to tell a well-known story, preferably the Cinderella fairytale. A range of measures are taken from their narrative sample, including the number of words produced per minute, and the proportion of open class words produced (i.e., nouns, verbs, adjectives and adverbs). This test was not included in our original test battery,

so QPA scores were not considered when our participants were assigned to their diagnostic groups. However, the selected measures from this test that are described above were used in the post-hoc fluency analyses that were completed.

### *Repetition Tests*

*The PALPA Non Word Repetition Test (sub-test 8)* is a test of acoustic-phonological conversion (Kay et al., 1992). In this task, the examiner reads aloud 30 pronounceable non-words of variable length (e.g., splank, vater, ality) and the participant attempts to repeat the non-words back to the examiner.

*The PALPA Auditory Word Repetition Test (sub-test 9)* assesses imageability and frequency effects in word repetition (Kay et al., 1992). In this task, the examiner reads aloud 80 words from four categories: 1) Low Imageability, Low Frequency; 2) Low Imageability, High Frequency; 3) High Imageability, Low Frequency; 4) High Imageability, High Frequency. Participants are required to repeat the words back to the examiner.

### *Reading Tests*

*The PALPA Imageability and Frequency Reading Test (subtest 31)* assesses imageability and frequency effects in single word reading (Kay et al., 1992). In this task, the participant is asked to read 80 words from four categories: 1) Low Imageability, Low Frequency; 2) Low Imageability, High Frequency; 3) High Imageability, Low Frequency; 4) High Imageability, High Frequency.

*The PALPA Spelling-Sound Regularity Test (subtest 35)* assesses whether reading is affected by spelling-sound regularity (Kay et al., 1992). In this task, participants are asked to read thirty regular words (e.g., nerve), and 30 exception words (e.g., island).

*The PALPA Nonword Reading Test (subtest 36)* assesses oral reading of unfamiliar letter strings (Kay et al., 1992). In this task, participants are shown 24 non-words (e.g., ked, birl, smode), and are asked to tell the examiner how they think each non-word should be pronounced.

### Appendix A3. Case Descriptions of Aphasic Participants

#### *Non-Fluent Aphasics*

##### *CT.*

In 1995, aged 35, CT suffered a cerebrovascular accident (CVA) involving the left middle cerebral artery. This resulted in a large lesion incorporating her left frontal and parietal lobes. In 2001 CT suffered another CVA involving her left frontal lobe and right cerebellum. CT produces grammatically simple spontaneous speech that has a slow and hesitant quality. Her speech is disrupted by frequent word finding pauses, and occasional phonemic and semantic paraphasias (see Table 2.7 for a sample of her speech).

CT was diagnosed with Broca's aphasia in a speech and language pathology report dated 2004. On administration of the BDAE during 2006, CT did not fit into any of the BDAE aphasia profiles. She presented with better melodic line, phrase length and articulatory agility scores than a pure Broca's aphasic. However, she also produced more paraphasias than are usually observed in Broca's aphasics (1-2 per minute of conversation), and her comprehension score fell below the Broca's aphasia range. At best, CT could be described as a mixed non-fluent aphasic.

*Comprehension.* CT has a comprehension deficit that mainly affects her sentence level comprehension. Her semantic processing of pictures and single words is largely preserved: On the Three-Pictures version of the Pyramids and Palm Trees Test, CT scored within the normal range; on the BDAE Word Comprehension subtest, CT scored 36/37; on the PPVT, CT scored in the 'moderately low' range; and on the PCB Lexical Comprehension subtest, CT scored within the normal range. On the PCB Sentence Comprehension subtest CT scored just below the normal range on sentences with lexical distracters. However, when the sentences contained reverse role distracters, CT scored well below the normal range, at chance level. Her sentence level comprehension impairment was also reflected in poor BDAE scores for following commands (9/15), and for comprehending complex ideational material (2/12).

*Production.* CT has a clear speech production deficit that affects single word production and sentence production, especially the latter. Her single word production deficit was reflected on the BNT, where she scored well below the normal range. She

also showed evidence of impairment on the Noun-Verb Naming Test, where her verb and high-frequency noun production were below the normal range, although her low-frequency noun production was within the normal range. She did not show a clear dissociation between her noun and verb production, however her verb score was slightly lower than her high and low frequency noun scores. CT did well on the COWAT, particularly the category fluency component, where her score was within the normal range when she was generating items for the *Animals* category: However, her score was well below the normal range for the letter fluency component. On the Lukkien Single Noun Test, CT scored 28/35 correct. However, when she was asked to name the same items within sentences in the Lukkien Sentence Production Test, she showed a trend towards less successful performance (22/35) (Yates'  $\chi^2(1) = 2.52$ ,  $p = .19$ ). Of the 35 nouns tested in the Lukkien Sentence Production test, CT omitted 6, and made 7 substitutions. The substitution errors included three phonemic paraphasias, two semantic paraphasias, one formal paraphasia, and one unrelated paraphasia. Her complete sentence score (the total number of target sentences correctly produced) was 0/30, although her correct syntax score was 4/30.

*Repetition.* CT has a mild repetition deficit that especially affects low-imageability words and non-words. On the PALPA Auditory Word Repetition Test (Subtest 9), her scores were below the normal range. She did very poorly on the PALPA Non-Word Repetition Task (Subtest 8) scoring 10/30 items correct: Of her 20 errors, 13 were real word substitutions (e.g., 'splint' for 'splant', and 'radio' for 'adio').

*Reading.* CT has a severe phonological dyslexia, presenting with a marked high imageability word advantage. On the PALPA Imageability and Frequency Reading Test (Subtest 31), CT scored below the normal range overall: CT's score was significantly better for the high imagery words (35/40) than the low imagery words (18/40) (Yates'  $\chi^2(1) = 14.312$ ,  $p < .01$ ); She scored 29/40 high frequency words correct, and 24/40 low frequency words correct. On the PALPA Spelling-Sound Regularity Test, CT's scores were below the normal range, with 16/30 regular words correct, and 18/30 exception words correct. Although CT performed below the normal range on all of these tests, she showed a much more severe deficit on the Non-word Reading Task, scoring 1/24.

DA.

DA had a CVA during 2000, aged approximately 59. Further medical/anatomical information is not available. DA presents with non-fluent speech, which is characterised by articulation difficulties, flat intonation, use of fillers, long word finding pauses, and grammatical errors that he occasionally self-corrects (see Table 2.7 for a sample of his speech)

On administration of the BDAE during November 2006, DA was diagnosed with Broca's aphasia. This diagnosis is consistent with an early diagnosis given in a speech and language pathology report from 2001.

*Comprehension.* DA has intact single word comprehension, but a mild sentence comprehension deficit. On the Three-Pictures version of the Pyramids and Palm Trees test, DA performed within the normal range, indicating that he has good non-verbal semantic access. His intact single word comprehension was reflected in a 'moderately high score' on the PPVT, and ceiling performance on the PCB Lexical Comprehension test. However, on the PCB Sentence Comprehension test, DA showed a clear impairment. On sentences with lexical distracters he scored below the normal range (25/30), and on sentences with reverse role distracters his score was considerably below normal, but still above chance (23/30).

*Production.* DA has a mild single word production deficit, but severely non-fluent sentence production. On the BNT, DA scored well within the normal range. On the Noun-Verb Naming test, DA's performance on the verbs and the high frequency nouns was within the normal range, but he scored slightly below the normal range on the low frequency nouns. On the COWAT, DA scored within the normal range on the category (animals only) and the letter fluency components. On the Lukkien Single Noun Test, DA scored 27/35: When he was asked to name the same items within sentences in the Lukkien Sentence Production Test, he had a similar success rate (30/35). Of the 35 nouns tested in the Lukkien Sentence Production test, DA omitted 1, and made 4 substitutions. The substitution errors included: Two semantic paraphasias; and two 'other' errors (unspecified). His complete sentence score (the total number of target sentences correctly produced) was 7/30, and his correct syntax score was 13/30.

*Repetition.* DA has a repetition deficit that affects his sentence and single word repetition. On the PALPA Auditory Word Repetition Test (Subtest 9), DA scored below the normal range: He showed a trend towards better performance on the

high imageability items (32/40) than the low imageability items (25/40) (Yates'  $\chi^2(1) = 2.99, p < .14$ ), but he did not show a frequency effect. DA showed a severe deficit on the PALPA Non-Word Repetition Test (Subtest 8), scoring 4/30.

*Reading.* DA has a mild phonological dyslexia. On the PALPA Imageability and Frequency Reading Test (Subtest 31), DA scored just below the normal range: He did slightly better on the high imageability (38/40) and high frequency words (38/40) than the low imageability (36/40) and low frequency words (36/40). On the PALPA Spelling-Sound Regularity Test (Subtest 35), DA performed at ceiling on the regular words (e.g., shoe, peril), but just below the normal range on the exception words (e.g., bouquet, mortgage). On the PALPA Non-Word Reading Test (Subtest 36), DA scored well within the normal range for the 3, 4, and 5-letter non-words, although he was under the normal range for the 6-letter non-words.

*DD.*

In July 2000, aged 53, DD had a left temporal lobe abscess surgically removed. Subsequently, he suffered an acute infarct in the left frontal and temporal areas, involving the infero-lateral frontal lobe and the peri-sylvian fissure. DD presents with severely non-fluent spontaneous speech, which primarily consists of content words and fillers. His speech appears effortful, and other people are often relied on to drive conversation (see Table 2.7 for a sample of his speech).

DD was diagnosed with Broca's aphasia on a speech and language pathology report dated 2004. Administration of the BDAE during 2006/2007 confirmed that this was the best diagnosis, although DD's comprehension score fell below the Broca's range (17/100), as did his sentence repetition score (0/10).

*Comprehension.* DD has a comprehension deficit that mainly affects his sentence level comprehension. On the Three-Pictures version of the Pyramids and Palm Trees Test, DD scored within the normal range: This suggests that his non-verbal semantic processing is preserved. His single word comprehension is also well preserved: On the PPVT, he scored in the 'moderately low' range, and on the PCB Lexical Comprehension subtest, he scored at ceiling. However, on the PCB Sentence Comprehension subtest, DD's scores were below the normal range for sentences that contain lexical distracters (26/30), and well below the normal range for sentences that



contain reverse role distracters, where he had close to chance level success (14/30): This difference was highly significant (Yates'  $\chi^2(1) = 10.8, p < .01$ )

*Production.* DD has a severe non-fluent speech production deficit that affects his single-word and sentence production. On the BNT, DD scored well below the normal range. He also performed well below the normal range on the Noun-Verb Naming test, where his results showed a trend towards particularly bad verb production (Yates'  $\chi^2(1) = 3.214, p = .12$ ). On the category fluency component of the COWAT (animals only), DD's score was also below the normal range, and he was not able to produce a single exemplar on the letter fluency component. On the Lukkien Single Noun Test, DD scored 21/35. However, when he was asked to name the same items within sentences in the Lukkien Sentence Production Test, he showed a trend towards less successful performance (16/35) (Yates'  $\chi^2(1) = 1.433, p < .34$ ). Of the 35 nouns tested in the Lukkien Sentence Production test, DD omitted 9, and made 10 substitutions. The substitution errors included: seven semantic paraphasias, 1 unrelated paraphasia, and 2 other (unspecified) errors. DD's complete sentence score (the total number of target sentences correctly produced) was 0/30, as was his syntax score.

*Repetition.* DD has considerable repetition impairment. On the PALPA Auditory Word Repetition Test (Subtest 9), his scores were below the normal range: He scored significantly better on the high imageability words (34/40) than the low (21/40) (Yates'  $\chi^2(1) = 9.833, p < .01$ ); but did not show a clear difference between the high frequency (26/40) and low frequency words (29/40). He also did poorly on the PALPA Non-Word Repetition Test (Subtest 8), only scoring 8/30 correct. DD was unable to repeat any of the sentences included in the sentence repetition component of the BDAE.

*Reading.* DD has severe deep dyslexia. On the PALPA Imageability and Frequency Reading Test (Subtest 31), he scored well below the normal range on all subcategories, and was especially poor on the low imageability items. On the PALPA Spelling-Sound Regularity Test, DD's scores were also well below the normal range: He read 8/30 regular words correctly, and 10/30 exception words. DD's word reading errors included semantic substitutions (e.g., *Ceiling*: "Floor", *Bury*: "Die", *Choir*: "Sing", and *Island*: "Sea"), visual/formal errors (e.g., *Friction*: "Fridge", *Quay*: "Quake"), and derivational errors (e.g., *Sew*: "Sewing"). DD was completely unable

to read any of the non-words included in the PALPA Non-Word Reading Test (Subtest 36).

### *ECV.*

ECV presents with generally non-fluent speech, which is occasionally disrupted by a surprisingly long sentence. Her spontaneous speech is marked by frequent word finding pauses and fillers, and occasional paraphasias. However, she produces a range of grammatical constructions, and has good word repetition (see Table 2.7 for a sample of her speech).

On a speech and language pathology report from 2001, ECV was diagnosed with transcortical motor aphasia. In 2006, when examined using the BDAE, ECV's profile best matched that of an anomic aphasic: Her spontaneous speech was characterised by good articulation and few errors, and her sentence repetition was fair (7/10). However, her auditory comprehension score fell below the range expected of both transcortical motor and anomic aphasics (37/100). Further testing indicated that she has sentence level comprehension difficulties, which were more typical of the non-fluent than fluent aphasias. Because of her sentence level comprehension difficulties, her predominantly non-fluent speech, and her previous diagnosis with a non-fluent aphasia, ECV has been classified as non-fluent despite her occasional well-formed sentence production.

*Comprehension.* ECV has a borderline single-word comprehension deficit, and moderately impaired sentence comprehension. On the Three-Pictures version of the Pyramids and Palm Trees Test, ECV scored 46/52: This is a score that the test authors suggest reflects a clinically significant semantic access deficit. In contrast, she scored in the moderately low range on the PPVT, and at the bottom of the normal range on the PCB Lexical Comprehension subtest. It is unusual to see impaired performance on the easy Pyramids and Palm Trees test, with relatively intact single word comprehension: This pattern may reflect a loss of the executive skills required to match pictures, rather than a semantic processing impairment *per se*. ECV was impaired on the PCB Sentence Comprehension subtest: She scored below the normal range on sentences with lexical distracters (24/30), and sentences with reverse-role distracters (20/30). Taken together, these scores suggest that ECV may have some mild difficulties accessing semantic information or some disrupted executive

processes, and that these impairments may be compounded by difficulties comprehending grammatically encoded information.

*Production.* ECV has a moderate single word production deficit, and a moderate sentence production deficit. On the BNT, ECV scored on the lower borderline of the normal range. However, on the COWAT, ECV's scores were slightly below the normal range on the category (animals only) and letter fluency components. ECV also performed below the normal range on all three sub-tests of the Noun-Verb Naming Test, scoring 19/30 on the verb component, 23/30 on the high frequency nouns, and 25/30 on the low frequency nouns score. ECV also achieved a moderately low score on the Lukkien Single Noun Test (25/35). However, when she was asked to name the same items within sentences in the Lukkien Sentence Production Test, she showed a trend towards less success (20/35) (Yates'  $\chi^2(1) = 1.556, p < .32$ ). On the Lukkien Sentence Production test, her complete sentence score (the total number of target sentences correctly produced) was 4/30, and her correct syntax score was 5/30. Of the 35 nouns tested in the Sentence Production test, ECV omitted 2, and made 13 substitutions. The substitution errors included: five semantic paraphasias; three neologisms; three unrelated words; and two mixed errors.

*Repetition.* Despite ECV's fair performance on the repetition component of the BDAE (7/10), she presents with a moderate word repetition deficit. On the PALPA Auditory Word Repetition Test (Subtest 9), she scored within the normal range on the high imageability items (36/39), but below the normal range on the low imageability items (32/39), producing phonemic and formal errors. However, the difference between her high and low imageability repetition scores was not significant (Yates'  $\chi^2(1) = 1.835, p < .31$ ). ECV scored 17/30 on the PALPA Non-Word Repetition Task (Subtest 8), again making phonemic and formal errors.

*Reading.* ECV has phonological dyslexia. She presents with reading that is sensitive to imageability. On the PALPA Imageability and Frequency Reading Test (Subtest 31), she scored below the normal range: Her reading of high imagery words (36/39) was significantly better than her reading of low imagery words (13/40) (Yates'  $\chi^2(1) = 29.988, p < .01$ ). On the PALPA Spelling-Sound Regularity Test (Subtest 35), she got 11/30 regular words correct, and 18/30 exception words correct: Both of these scores were well below the normal range. Although ECV performed below the normal range on all of these reading tests, she showed a much more severe

deficit on the PALPA Non-word Reading Test (Subtest 36), scoring 1/24. Her reading errors were generally formal, or she would say that she did not know the word.

### *JHM.*

JHM suffered a left middle cerebral artery CVA during 2000, aged 40. She presents with severely non-fluent spontaneous speech that is characterised by long word finding pauses, effortful articulation, and frequent phonemic paraphasias (see Table 2.7 for a sample of her speech).

JHM's language was assessed during March 2004, and she was diagnosed with Broca's aphasia. She has a flat melodic line, short phrase length, and poor sentence repetition. Her auditory comprehension was towards the bottom of the Broca's range, but her rate of semantic paraphasia production was low.

*Comprehension.* JHM has good comprehension of single words, but a sentence level deficit that particularly affects her understanding of grammatically complex sentences. JHM's score on the Three-Pictures version of the Pyramids and Palm Trees Test was within the normal range, indicating that she has good non-verbal semantic access. Her good single-word comprehension is reflected in a low-average score on the PPVT, and ceiling performance on the PCB Lexical Comprehension subtest. On the PCB Sentence Comprehension subtest, JHM scored on the borderline of the normal range on the sentences containing lexical distracters (29/30). However, on the sentences containing reverse role distracters, she showed a clear impairment and near chance performance (16/30).

*Production.* JHM has a mild single word production deficit, but a severe sentence production deficit. On the BNT, JHM scored just below the normal range. On the category fluency component of the COWAT, her score was within the normal range (animals only). However, her score was well below the normal range for the letter fluency component. On the Noun-Verb Naming test she scored just below the normal range on the high frequency nouns and the verbs, but she performed at ceiling on the low frequency nouns: She did not show a clear dissociation between noun and verb production. On the Lukkien Single Noun Test, JHM scored 30/35. However, when she was asked to name the same items within sentences in the Lukkien Sentence Production Test, she made a similar number of errors (28/35). Of the 35 nouns tested in the Lukkien Sentence Production test, JHM omitted 2, and made 5 substitutions.

The substitution errors included three phonemic paraphasias and two semantic paraphasias. On the Lukkien Sentence Production Test, her complete sentence score (the total number of target sentences correctly produced) was 2/30, and her correct syntax score was 9/30.

*Repetition.* JHM has a severe sentence repetition deficit, but her single word repetition is considerably better. Given the PALPA Auditory Word Repetition Test (Subtest 9), she scored within the normal range overall. Her high imageability word score was at ceiling (40/40) but her repetition of the low imageability words was below the normal range (35/40); the difference between her high and low imageability word reading scores approached significance (Yates'  $\chi^2(1) = 5.333, p < .06$ ). JHM also performed well on the PALPA Non-Word Repetition Test (Subtest 8).

*Reading.* JHM has a mild real word reading deficit, with a severe non-word reading deficit. Generally, her reading usually fits the pattern of a phonological dyslexic: However she very occasionally made semantic errors, suggesting that her reading may be on the borderline of deep dyslexia. On the PALPA Imageability and Frequency Reading Test (Subtest 31), JHM scored slightly below the normal range on high imageability words (38/40), but her score was considerably below the normal range on low imageability words (30/40). The difference between her high and low imageability word reading was significant (Yates'  $\chi^2(1) = 6.275, p < .02$ ). She showed a trend towards better high frequency words reading (37/40) than the low frequency word reading (31/40) (Yates'  $\chi^2(1) = 3.529, p < .12$ ). On the PALPA Spelling-Sound Regularity Test (Subtest 35), she showed no effect of regularity scoring 28/30 regular words correct, and 29/30 exception words correct. On the PALPA Non-Word Reading Test (Subtest 36) she was unable to sound out any of the letter strings she attempted (testing was discontinued at item 11 at her request).

*TB.*

During September 2002, TB suffered a left middle cerebral artery infarction / CVA infarct with oedema. This resulted in a mass effect on the anterior horn of her left lateral ventricle. Her spontaneous speech production is mildly impaired, but includes frequent word finding pauses. However, she is able to hold fairly normal conversations, and she successfully works in a cafeteria part time (see Table 2.7 for a sample of her speech).

Examination of TB's speech using the BDAE during 2007 showed minimal evidence of aphasia. However, a speech and language pathology report from 2002 indicates that she presented with Broca's aphasia. A subsequent report from 2004 indicated some language recovery, as she then fitted the classification of an anomic aphasic. Her performance on the more specific tests summarised below, suggests that she still has considerable language impairments, especially with sentence level processing. Taken together the tests suggest that TB is best described as a recovered Broca's aphasic.

*Comprehension.* TB has good single word comprehension, but an impairment of sentence level comprehension that particularly affects grammatically encoded information. Her score on the Three-Pictures version of the Pyramids and Palm Trees Test was within the normal range, indicating that her non-verbal semantic access is intact. Her good single word comprehension was reflected in a moderately low score on the PPVT, and scores just below the normal range on the PCB Lexical Comprehension subtest. On the PCB Sentence Comprehension subtest, TB performed at ceiling on the sentences containing lexical distracters. However, on the sentences containing reverse role distracters, she scored well below the normal range, although still well above chance (23/30).

*Production.* TB has moderate single-word and sentence production deficits. Her single-word production deficit was reflected in a score that was below the normal range on the BNT. Also, on the Noun-Verb Naming test, TB scored below the normal range on the high frequency nouns, although she scored on the borderline of the normal range for the verbs, and at ceiling on the low frequency nouns. She did not show a clear dissociation between noun and verb production. TB also performed within the normal range on the category fluency component of the COWAT (animals only). However, her score was well below the normal range for the letter fluency component. On the Lukkien Sentence Production Test, TB's complete sentence score (the total number of target sentences correctly produced) was 8/30, although her correct syntax score was 20/30. Of the 35 nouns included in both the Lukkien Sentence Production Test and the Lukkien Single Noun Test, TB produced 31 nouns correctly in the context of single nouns, and 30 nouns correctly within sentences. Of the five noun errors she made on the Lukkien Sentence Production test, three were semantic paraphasias, and two were unrelated words.

*Repetition.* TB's ability to repeat words and sentences remains intact. Given the PALPA Auditory Word Repetition Test (Subtest 9), TB's scores were within the normal range, regardless of the imageability or frequency of the words that were to be repeated. TB also performed well on the PALPA Non-Word Repetition Test (Subtest 8), and the sentence repetition component of the BDAE (10/10).

*Reading.* TB has a mild phonological dyslexia. Her real-word reading is mildly impaired, but she cannot read non-words. She scored just below the normal range on the PALPA Imageability and Frequency Reading Test (Subtest 31) and the PALPA Spelling-Sound Regularity Test. In contrast, on the PALPA Non-Word Reading Test (Subtest 36), she scored 0/24. On the PALPA Non-Word Reading Task, she tended to pass on items, or to make a formal paraphasia and say that she knew that her answer was wrong because it was a real word. For example, in response to the non-word 'Dringe', TB replied "I want to say it's is um drainage, but that's not a word".

*WS.*

WS had a stroke in 1994, aged 63. A CT scan performed during March 1996 showed a large low attenuation area in his left fronto-temporo parietal region, consistent with a left middle cerebral artery infarct and associated frontal lobe atrophy. WS's spontaneous speech is severely non-fluent, and marked by a common perseveration of his favourite horse's name, "Him I Gotta Go" (WS was a successful jockey and horse trainer before he retired) (see Table 2.7 for a sample of his speech).

When examined using the BDAE, WS's language did not correspond with one aphasia profile. The closest match was to Broca's aphasia, where he scored within the specified range on all measures apart from the sentence repetition component where he scored at floor, and the auditory comprehension measures, where he scored below the specified range (42.5/100).

*Comprehension.* WS has good single word comprehension, but a moderate sentence comprehension deficit that particularly affects his understanding of grammatically encoded information. WS scored within the normal range on the Three-Pictures version of the Pyramids and Palm Trees test, indicating that his non-verbal semantic access is good. His good single word comprehension is reflected in a score within the normal range on the PPVT (a moderately low score), and a score on the lower borderline of the normal range on the PCB Lexical Comprehension subtest.

On the PBC Sentence Comprehension test, WS performed just below the normal range on sentences with lexical distracters (26/30), but well below the normal range on sentences with reverse role distracters, although still above chance (19/30): This difference approached significance (Yates'  $\chi^2(1) = 4.356, p < .07$ ).

*Production.* WS has a severe language production impairment that affects his single-word and sentence production. WS showed a severe impairment on the BNT, scoring well below the normal range. WS also showed a clear deficit on the COWAT: He was unable to produce any items on the letter fluency component, and on the category fluency component (animals only) he only produced three items (horse, mares, hog). On the Noun-Verb Naming Test, WS again scored well below the normal range. He was impaired on all three categories, but his impairment was significantly worse on the verbs (6/30 correct) than the high frequency nouns (20/30) and the low frequency nouns (20/30) (Yates'  $\chi^2(1) = 17.431, p < .01$ ). On the Lukkien Sentence Production Test, his complete sentence score (the total number of target sentences correctly produced) was 0/30, as was his correct syntax score. Of the 35 nouns included in both the Lukkien Sentence Production Test and Lukkien Single Noun Test, WS produced 17/35 nouns correctly in the context of single nouns, and 9/35 nouns correctly within sentences: This difference approached significance (Yates'  $\chi^2(1) = 3.916, p = .08$ ). Of the 26 noun errors WS made on the Sentence Production Test, 6 were omissions, 14 were 'other' (unspecified), 2 were phonemic paraphasias, and 2 were semantic paraphasias.

*Repetition.* WS has a moderate repetition impairment. On the PALPA Auditory Word Repetition Test (Subtest 9), WS scored well below the normal range. His performance appeared to be influenced by both imageability (high 29/40, low 18/40), and frequency (high 29/40, low 18/40). WS also performed very poorly on the PALPA Non-Word Repetition Test (Subtest 8) scoring 4/30: His errors consisted of phonologically related neologisms (e.g., gaffic "/dæfɪk/"), and phonologically related real word substitutions (e.g., splack "black").

*Reading.* WS has severe phonological dyslexia, and his reading is affected by word frequency and imageability. On the PALPA Imageability and Frequency Reading Test (Subtest 31), WS performed well below the normal range and, as with his repetition scores, his reading scores appeared to be influenced by both imageability (high 29/40, low 13/40) and frequency (high 26/40, low 16/40). On the



PALPA Spelling-Sound Regularity Reading Test (Subtest 35), WS again performed well below the normal range, although no differentiation between regular (16/30) and exception (17/30) word reading appeared. WS also showed a considerable impairment on the PALPA Non-word Reading Test (Subtest 36), where he made only two correct responses, both of which were of single syllable, three-letter non-words. WS's reading errors were mainly phonologically related real word substitutions.

### *Fluent Aphasics*

#### *FS.*

FS presents with fluent speech that is interrupted by word finding pauses. While her melodic line and articulatory agility are good, her speech sometimes presents with a slightly slurred quality (see Table 2.7 for a sample of her speech).

When examined using the BDAE, FS presented with good melodic line, phrase length, and use of grammatical forms. She had poor repetition, word finding, and auditory comprehension and was best diagnosed as a Wernicke's aphasic, although she made fewer paraphasias than is stipulated in the BDAE Wernicke's aphasia profile.

*Comprehension.* FS has deficit of single-word and sentence comprehension. Her score on the Three-Pictures version of the Pyramids and Palm Trees Test was below the normal range, indicating that her non-verbal semantic access was also impaired. Her single-word comprehension deficit was reflected in her score on the PPVT, which was in the extremely low range. Also, her scores on the PCB Lexical Comprehension subtest were at borderline for within and between category distracters. Evidence of a sentence comprehension deficit was also shown on the PCB Sentence Comprehension subtest, as she performed well below the normal range on sentences with lexical distracters and sentences with reverse role distracters.

*Production.* FS has a moderate speech production deficit. Her object naming is impaired, as reflected by a BNT score that fell below the normal range. In addition, her score on the category fluency component of the COWAT was borderline (animals only). The letter fluency component of the COWAT was not scored due to a recording failure. On the Noun-Verb Naming test, FS scored below the normal range on the high (24/30) and low frequency (21/30) nouns, but she had a significantly lower score on the verbs (12/30) (Yates'  $\chi^2(1) = 10.55, p < .01$ ). On the Lukkien

Sentence Production Test, FS's complete sentence score (the total number of target sentences correctly produced) was 0/30, and her correct syntax score was 4/30. Of the nouns included in both the Lukkien Sentence Production Test and the Lukkien Single Noun Test, FS produced 20/35 correctly in the context of single nouns, and 17/35 correctly within sentences. Of the 18 noun errors FS made on the Lukkien Sentence Production Test, she omitted 6, and made 4 semantic paraphasias, 3 phonemic paraphasias, 2 formal paraphasias, 2 neologisms, 2 other (unspecified) errors, and 1 unrelated error. Her speech error rate was considerably higher on the Lukkien Sentence Production Test than the connected speech measures included in the BDAE. This difference may reflect the highly constraining nature of the Lukkien Sentence Production Test, where one target sentence is expected, which can be contrasted with the freer measure in the BDAE which allows for more compensatory speech. It may also reflect FS's prior exposure to the measures included in the BDAE. As mentioned previously, FS did not make enough paraphasias on the BDAE to fit the Wernicke's profile. However the number of paraphasias made on the Lukkien Sentence Production Test implies that she does have a considerable impairment in this area.

*Repetition.* FS has a severe repetition impairment. Given the PALPA Auditory Word Repetition Test (Subtest 9), FS scored well below the normal range on high imageability words (28/40), however her low imageability word repetition was significantly worse (14/40) (Yates'  $\chi^2(1) = 9.825, p < .01$ ). FS also showed a clear deficit on the PALPA Non-Word Repetition Test (Subtest 8).

*Reading.* FS has a phonological dyslexia. On the PALPA Imageability and Frequency Reading Test (Subtest 31) and the PALPA Spelling-Sound Regularity Test (Subtest 35), FS scored below the normal range on all subcategories: On these tasks, FS's reading errors tended to be phonemic or formal paraphasias. On the PALPA Non-Word Reading Test (Subtest 36) FS read 10/24 items correctly: For three-letter non-words (e.g., 'ked', 'bem', 'nar'), FS's reading was within the normal range, but for the four, five, and six letter non-words, her score was below the normal range.

*NP.*

NP suffered a sudden onset CVA in August 1999, aged 60. At that time, NP presented with a right homonymous hemianopia, right hemiplegia, right neglect, and expressive aphasia. A CT scan revealed reduced attenuation in the left occipital lobe,

the left basal ganglion, and the left temporal lobe. NP presents with fluent speech that incorporates a variety of grammatical constructions, but is disrupted by frequent word finding pauses (see Table 2.7 for a sample of his speech). When NP's language abilities were examined using the BDAE, he showed word finding difficulties with some paraphasias, and was diagnosed with Anomic aphasia.

*Comprehension.* NP has good comprehension. On the Three-Pictures version of the Pyramids and Palm Trees test he scored within the normal range, indicating that he has good non-verbal semantic access. His good single-word and sentence level comprehension were reflected in score within, or on the borderline of, the normal range on the PPVT (with a moderately low score), the PCB Lexical Comprehension subset, and the PCB Sentence Comprehension subtest, where he scored within the normal range on sentences containing lexical distracters, and sentences containing reverse-role distracters.

*Production.* NP has a moderate single-word production deficit, but frequently produces syntactically correct sentences. His single-word production deficit was reflected in a BNT score that was below the normal range. His performance on all three sections of the Noun-Verb Naming test were also below the normal range, but his consistency in performance between the nouns and verbs did not exhibit a disproportionate deficit on the noun or verb syntactic form class, although he reports having the most difficulty naming objects. However, on the COWAT, he scored at the low end of the normal range on both the letter fluency and category fluency (animals only) components. On the Lukkien Single Noun Test, NP scored 30/35. However, when he was asked to name the same items within sentences in the Lukkien Sentence Production Test, he showed a trend towards poorer performance, scoring 24/35 (Yates'  $\chi^2(1) = 2.917, p = .15$ ). Of the 35 nouns tested in the Lukkien Sentence Production test, NP made 11 substitutions. The substitution errors included: six semantic paraphasias, one phonemic paraphasia, one unrelated paraphasia, and three other errors. On the Lukkien Sentence Production Test, NP's complete sentence score (the total number of target sentences correctly produced) was 7/30, and his correct syntax score was 18/30.

*Repetition.* NP has a moderate word repetition deficit, and a severe non-word repetition deficit. On the PALPA Auditory Word Repetition Test (Subtest 9), NP scored slightly below the normal range. His performance was fairly consistent across the different item categories, scoring 36/40 for the high imageability items, 32/40 for

the low imageability items, and 34/40 for both the high and low frequency items. However, he performed poorly on the PALPA Non-word Repetition Task (Subtest 8): His errors were almost exclusively neologisms that were phonologically related to the target non-word, with two phonologically related real word substitutions.

*Reading.* NP has a mild reading deficit. On the PALPA Imageability and Frequency Reading Test (Subtest 31), NP performed just beneath the normal range. His performance on the high imagery and high frequency (19/20), and high imagery and low frequency (20/20) words was within the normal range, but he scored below the normal range for the low imagery high frequency (18/20), and low imagery low frequency words (16/20). On the PALPA Spelling-Sound Regularity Reading Test (Subtest 35), NP performed just below the normal range, but he got the same score on the regular and exception words (27/30). Finally, on the PALPA Non-Word Reading Test (Subtest 36), NP scored at the low end of the normal range: Three of his four errors occurred on the longest, 6-letter non-words.

*NS.*

NS suffered a sudden onset CVA on February 28, 2000, aged 83. A CT scan, undertaken three days post stroke, revealed a small region of hypodensity in his left posterior parietal lobe. He presents with fluent speech that is characterised by frequent paraphasias, and he can have obvious difficulty following conversations (see Table 2.7 for a sample of his speech). When NS's language was examined using the BDAE, he presented with a borderline Wernicke's / conduction aphasia.

*Comprehension.* NS has a comprehension deficit, but good non-verbal semantic access. NS scored highly on the Three-Pictures version of the Pyramids and Palm Trees test suggesting that his non-verbal semantic access is good. NS's reading comprehension is better than his auditory comprehension, and he frequently asks you to write a word down if he does not comprehend it when spoken. On the PPVT, NS achieved a 'moderately low score', which is within the normal range. However, on other tasks that require auditory input, his performance was impaired. For instance, on the PCB Lexical Comprehension subtest NS scored well below the normal range, as he did on the PCB Sentence Comprehension Subtest.

*Production.* NS's single word production is moderately impaired. On the BNT, NS scored below the normal range. On the Noun – Verb Naming test, he scored below the normal range for all three groups of items, and a clear dissociation

between the nouns and the verbs did not appear. On the COWAT, he scored within the normal range on both the letter fluency and category fluency (animals only) components. On the Lukkien Single Noun Test, NS scored 27/30. When he was asked to name the same items within sentences in the Lukkien Sentence Production Test, he made one more error (26/35). Of the 35 nouns tested in the Lukkien Sentence Production test, NS omitted 1, and made 8 substitutions. The substitution errors included: two semantic paraphasias, one phonemic paraphasia, one unrelated paraphasia, and three other (unspecified) errors. On the Lukkien Sentence Production test, NS's complete sentence score (the total number of target sentences correctly produced) was 2/30, although his correct syntax score was 8/30.

*Repetition.* NS has a severe word repetition deficit, and he is completely unable to repeat non-words. On the PALPA Auditory Word Repetition Test (Subtest 9) NS scored 36/80, which is well below the normal range. NS's performance was influenced by word imageability, as he had a significantly poorer performance on the low imageability words (12/40) than high imageability words (24/40) (Yates'  $\chi^2(1) = 7.273, p < .05$ ). Word frequency did not appear to influence his repetition. On the PALPA Non-Word Repetition Test (Subtest 8), he scored 0/30. On both repetition tests, NA's errors were usually phonologically related non-words.

*Reading.* NS's reading is well preserved. On the PALPA Spelling-Sound Regularity Reading Test (Subtest 35), NS scored at ceiling on the 30 regular words. However, his performance on the exception words was slightly below the normal range (26/30). NS performed within the normal range on the PALPA Non-Word Reading Test (Subtest 36) by reading 21/24 items correctly.

*PS.*

During July 2002 at age 50, PS collapsed while playing netball and had a seizure. A CT scan revealed a left hemisphere subarachnoid haemorrhage, and an angiogram showed a left middle cerebral artery aneurysm. During surgery the aneurysm was dissected, and an additional large bleeding temporal vein was cauterised. Post operatively, PS has continued to have seizures, and is taking medication to control these. PS presents with fluent spontaneous speech, but experiences word finding difficulties (see Table 2.7 for a sample of his speech).

During 2006, PS's language abilities fitted the BDAE classification of an anomia aphasic: He showed good melodic line, phrase length, articulatory agility,

and auditory comprehension, but mild word finding and repetition deficits, some paraphasias in running speech, and used a slightly limited variety of grammatical constructions.

*Comprehension.* PS has good comprehension. He performed within the normal range on the Three-Pictures version of the Pyramids and Palm Trees test, indicating that his non-verbal semantic access is intact. At the single-word comprehension level, he also performed well: On the PPVT, he scored within the normal range with a low-average score; and on the PCB Lexical Comprehension test, he scored at the lower borderline of the normal range for both the within and between category items. On the PCB Sentence Comprehension test, PS scored at ceiling on the items with lexical distracters, and although he made three errors on the reverse-role sentences, his performance was still within the normal range.

*Production.* PS has a mild word production deficit. On the BNT, he performed within the normal range for his age group. On the COWAT, PS was again within the normal range, on both the letter fluency and category (animals only) fluency components. PS's ability to produce nouns and verbs was compared using the Noun-Verb test: PS scored within the normal range on the high frequency and the low frequency nouns, however, he scored well below the normal range on the verbs: His verb score was significantly lower than his overall noun score (Yates'  $\chi^2(1) = 10.827, p < .01$ ). On the Lukkien Single Noun Test, PS scored 31/35. When he was asked to name the same items within sentences in the Lukkien Sentence Production Test, he made two more errors (19/31). Of the 35 nouns tested in the Lukkien Sentence Production test, PS made 6 substitutions, but no omissions. On the Lukkien Sentence Production test, PS's complete sentence score (the total number of target sentences correctly produced) was 12/30, although his correct syntax score was 21/30.

*Repetition.* PS's word repetition is intact, although he shows a mild non-word repetition impairment. On the PALPA Auditory Word Repetition Test (Subtest 9), PS scored within the normal range. On the PALPA Non-Word Repetition Test (Subtest 8) he scored 24/30, but produced several phonologically related neologisms (e.g., “/st-ʒbəl/” for stirple) and phonologically related real word substitutions (e.g., “struggle” for truggle).

*Reading.* PS's reading ability is good, and does not appear to be influenced by frequency, imageability, or spelling-sound regularity. He performed within the

normal range on the PALPA Imageability and Frequency Reading Test (Subtest 31). On the PALPA Spelling-Sound Regularity Reading Test (Subtest 35), he performed within the normal range on the regular words, and at the lower borderline of the normal range on the exception words. On the PALPA Non-Word Reading Test (Subtest 36), PS scored just below the normal range.

### *STR.*

During September 2001, at 70 years of age, STR was admitted to hospital, presenting with left-side weakness, slurred speech, left facial weakness, and homonymous hemianopia. A CT scan confirmed a right occipital region ischemic stroke, and a possible left parietal area stroke. Her bilateral CVA extended 5 days after the initial incident, and she showed the additional symptoms of left hemiparesis, left hemineglect, and dysphasia. STR's spontaneous speech is fluent, but she makes paraphasias and word finding pauses (see Table 2.7 for a sample of her speech). STR was diagnosed with anomic aphasia using the BDAE: She has good auditory comprehension, melodic line, articulatory agility, and she uses a range of grammatical forms, but her speech has a mildly empty quality, and she makes some errors, particularly phonemic paraphasias.

*Comprehension.* STR's single word comprehension and semantic access are good, however she exhibits a mild comprehension deficit on grammatically complex sentences. Her score on the Three-Pictures version of the Pyramids and Palm Trees Test was at ceiling, suggesting that her non-verbal semantic access is good. She also performed well at the single word comprehension level: She scored within the normal range on the PPVT with a low-average score, and at the lower borderline of the normal range on the PCB Lexical Comprehension subtest. On the PCB Sentence Comprehension test, she scored at the lower borderline of the normal range on the sentences that contained lexical distracters: However, she scored slightly below the normal range when the sentences contained reverse-role distracters.

*Production.* STR has a moderate single-word production deficit. On the BNT, STR scored well below the normal range. On the Noun-Verb Naming test, her performance on all three sections was below the normal range. However, on the COWAT category (animals only) and letter fluency components, STR scored within the normal range. Of the 35 nouns included in the Luskien Sentence Production test and the Luskien Single Naming test, STR produced 25 correctly in the single-noun

context and 26 correctly within sentences. Of the noun errors that STR made during the Lukkien Sentence Production Test, seven were semantic paraphasias, two were phonemic paraphasias, and one was an unrelated word. On the Lukkien Sentence Production test, STR's complete sentence score (the total number of target sentences correctly produced) was 7/30, although her correct syntax score was 22/30.

*Repetition.* STR has a moderate single-word repetition deficit that is affected by word imageability. Given the PALPA Auditory Word Repetition Test (Subtest 9), STR showed a mild-moderate deficit on the high imageability words (34/40) and the low imageability words (29/40). STR performed fairly well on the PALPA Non-Word Repetition Test (Subtest 8) (26/30).

*Reading.* Although STR has a mild reading impairment, she has considerable success reading single words and non-words. On the PALPA Imageability and Frequency Reading Test (Subtest 31), STR's reading was affected by imageability: She showed mild-moderate impairments on the high imageability words (37/40) and the low imageability words (32/40). On the PALPA Spelling-Sound Regularity Reading Test (Subtest 35), STR scored just below the normal range on the regular and exception words. STR also performed fairly well on the PALPA Non-Word Reading Test (Subtest 36), scoring 21/24.

XX.

XX had a left middle cerebral artery CVA during August 2002, aged 59. He presents with fluent speech with an empty quality, and many word finding pauses (see Table 2.7 for a sample of his speech). Examination of XX's language using the BDAE indicated that XX has Wernicke's Aphasia.

*Comprehension.* XX has moderate single-word and sentence comprehension deficits. However, on the Three-Pictures version of the Pyramids and Palm Trees Test, he scored within the normal range, indicating that his non-verbal semantic access is good. XX's moderate single-word comprehension deficit was reflected in scores that were below the normal range on the PPVT (he scored in the extremely low range), and the PCB Lexical Comprehension subtest. On the PCB Sentence Comprehension subtest he scored below the normal range on the sentences that contained lexical distracters (27/30), but he made significantly more incorrect responses on the sentences that contained reverse role distracters (19/30) (Yates'  $\chi^2$  (1) = 5.963,  $p < .05$ ).



*Production.* XX has moderate to severe single-word and sentence production deficits. At the single word level XX scored well below the normal range on the BNT. On the Noun-Verb Naming test, he also scored below the normal range on the high frequency nouns, the low frequency nouns, and the verbs, but he did not show a clear dissociation between his noun and verb production. XX did better on the COWAT: On the letter fluency component, his score was within the normal range, however his score was below the normal range for the category fluency component (animals only). Of the 35 nouns included in the Lukkien Sentence Production test and the Lukkien Single Naming test, XX produced 22 correctly in the single-noun context and 17 correctly within sentences. On the Lukkien Sentence Production test, XX's complete sentence score (the total number of target sentences correctly produced) was 1/30, although his correct syntax score was 6/30. Of the 18 noun errors XX made on the Lukkien Sentence Production test, seven were omissions, five were semantic paraphasias, four were other (unspecified) errors, one was a neologism, and one was a mixed error.

*Repetition.* XX's ability to repeat aurally presented stimuli is impaired. Given the PALPA Auditory Word Repetition Test (Subtest 9), XX consistently scored below the normal range. His repetition appears to be influenced by imageability: He was correct on 28/40 high imageability words, but had significantly worse performance on the low imageability words 17/40 (Yates'  $\chi^2(1) = 6.146, p < .05$ ). XX also showed a deficit on the PALPA Non-Word Repetition Test (Subtest 8): Of his 23 errors, 4 were phonologically related real word substitutions, but the majority were phonologically related neologisms.

*Reading.* XX has impaired reading. On the PALPA Imageability and Frequency Reading Test (Subtest 31), he scored below the normal range on the high and low imageability words, but his reading was significantly better on the high imageability words (33/40) than the low imageability words (16/40) (Yates'  $\chi^2(1) = 15.221, p < .01$ ). He also scored below the normal range on the PALPA Spelling-Sound Regularity Reading Test (Subtest 31): His reading showed a trend towards worse reading of exception words (14/30) than regular words (22/30) (Yates'  $\chi^2(1) = 4.444, p < .07$ ). XX also produced a low score on the PALPA Non-Word Reading Test (Subtest 36), where he produced 13 phonologically related real word substitutions, 4 phonologically related neologisms, and 2 phonologically unrelated

real word substitutions (e.g., *Thease*: “Dust”).

## Appendix A4: Complete Control Participant Results from the Category Exemplar Task.

### *Design*

Due to limited availability some controls completed one experimental session, during which they completed one version of the task. Hence, these participants were presented with a category once only, and condition was controlled for across participants<sup>16</sup>. The data from the first session completed by all of the controls (including the three who completed two sessions) is considered separately here. The control participants were pseudo-randomly assigned a task version, with as even a distribution of participants completing each version as possible.

### *Data Analysis*

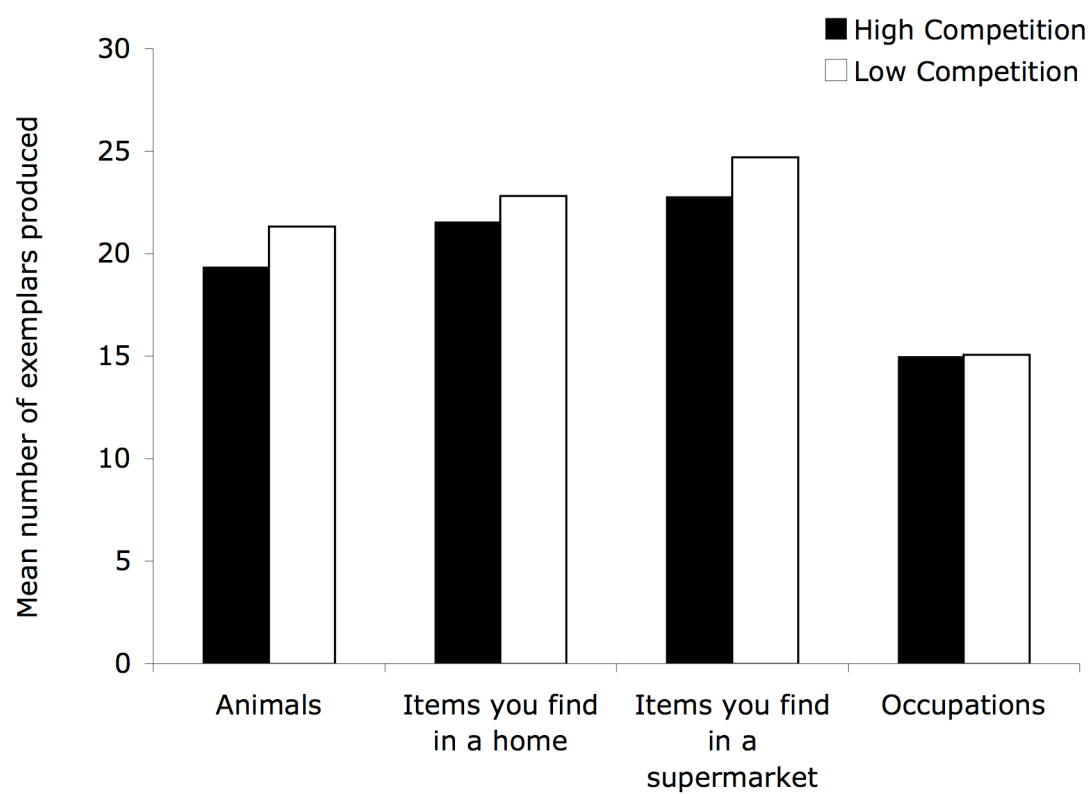
The number of correct responses made in the high and low competition conditions were analysed at the group level. A general linear mixed model was performed, including the following independent variables: condition, category, and a condition by category interaction.

### *Results*

The mean number of correct exemplars produced by category and competition condition is illustrated in Table A1. There was a general trend for the controls to produce more exemplars in the low competition condition than the high competition condition. However, a main effect of competition condition was not observed ( $F(1, 11) = 0.00, p > .05$ ), and nor was an interaction of competition condition and category ( $F(1, 10) = 0.19, p > .05$ ). The main effect of category was highly significant ( $F(1, 11) = 11.28, p < .01$ ).

---

<sup>16</sup> The three controls who were available for two sessions (AP, EM, and GA) completed an identical task design to the aphasic participants, and their data was analysed in the same way as the aphasic participants: These results are reported in the main text.



*Figure A1.* Category exemplar task: Complete control data from session one showing the mean number of correct exemplars produced by category and competition condition.

## Appendix B1: Verb Generation Pilot Task Instructions

**Verb Generation Task**Instructions:

Please do not turn the page until you are instructed to do so.

Overleaf you will find a list of objects. Your task is to write down what the object does, or what you can do with the object. In other words, you are required to write down the first verb that comes to your mind that is associated with the object. Some examples follow:

<u>Object:</u>	<u>Possible Response:</u>
----------------	---------------------------

Cat	Purr
-----	------

Scissors	Cut
----------	-----

## Appendix B2: List of Items Included in the Verb Generation Task Proper by Competition Condition

### High Competition Condition Items:

Alligator	Curtains	Key	Shark
Axe	Daisy	Leg	Soldier
Baby	Duck	Moon	Sparrow
Barbecue	Ear	Needle	Suitcase
Binocular	Envelope	Package	Tail
s	Feet	Pan	Teeth
Boat	Fence	Penny	Tiger
Candle	Frog	Pills	Tongue
Caravan	Heart	Radio	Watch
Carnation	Hinge	Razor	Worm
Cigarette	Horse	Road	
Crab	Ice	Seesaw	

### Low Competition Condition Items:

Airplane	Dice	Picture	Straw
Arrow	Elbow	Pipe	Sugar
Ball	Fire	Pool	Sun
Basin	Ghost	Priest	Telephone
Basket	Hawk	Rope	Towel
Bed	Kettle	Scales	Trapeze
Bell	Ladder	Scissors	Van
Broom	Lion	Shovel	Wool
Can	Lips	Snow	Yacht
Chair	Mosquito	Stethoscope	
Church	Nun	Stomach	
Crane	Piano	Stove	

## Appendix B3: List of Items Included in the Verb Generation Task by Frequency

## Condition

## High Frequency Condition Items:

Baby	Fence	Piano	Stomach
Ball	Fire	Picture	Straw
Bed	Ghost	Pills	Sugar
Bell	Heart	Pipe	Sun
Boat	Horse	Pool	Tail
Chair	Ice	Priest	Telephone
Church	Key	Radio	Tongue
Cigarette	Leg	Road	Van
Daisy	Lion	Rope	Watch
Ear	Moon	Scales	
Elbow	Pan	Snow	
Envelope	Penny	Soldier	

## Low Frequency Condition Items:

Airplane	Crab	Nun	Trapeze
Alligator	Crane	Package	Wool
Arrow	Curtains	Razor	Worm
Axe	Dice	Scissors	Yacht
Barbecue	Duck	Seesaw	
Basin	Feet	Shark	
Basket	Frog	Shovel	
Binocular	Hawk	Sparrow	
s	Hinge	Stethoscope	
Broom	Kettle	Stove	
Can	Ladder	Suitcase	
Candle	Lips	Teeth	
Caravan	Mosquito	Tiger	
Carnation	Needle	Towel	

#### Appendix B4: Additional Results for the Verb Generation Task Frequency Condition Analyses

*Accuracy.* Non-fluent JHM was significantly more accurate in the low competition condition than the high competition condition ( $\chi^2(1) = 7.91, p < .01$ ), and non-fluent ECV showed a trend towards a main effect of competition condition ( $p = .06$ ). The following participants showed significant main effects of session: FS ( $p < .01$ ), XX ( $p < .01$ ), and PS ( $p = .04$ ) were less accurate in the first session than the second; NP was significantly more accurate in the first session than the second ( $p < .01$ ). No other significant differences or trends towards significant differences ( $p > .05$  and  $p < .1$ ) were observed.

*Latencies.* The following participants had significantly faster RTs during the second testing session than the first: Non-fluents CT ( $F(1, 34) = 15.30, p < .01$ ), ECV ( $F(1, 25) = 19.38, p < .01$ ), and JHM ( $F(1, 60) = 7.72, p < .01$ ). The following participants had significantly slower RTs during the second testing session than the first: Non-fluents DD ( $F(1, 23) = 9.46, p < .01$ ), and TB ( $F(1, 54) = 6.41, p < .05$ ). NS showed a trend towards slower responses in the second session ( $p = .07$ ). No other significant differences or trends towards significant differences were observed ( $p > .05$  and  $p < .1$ ).



## Appendix B5: List of Items Included in the Verb Generation Task by Association Strength Condition

### Low Competition Condition:

Bed	Mosquito	Scales	Sun
Church	Picture	Snow	Trapeze
Crane	Pool	Stomach	Van
Hawk	Rope	Sugar	

### High Competition/High Association Condition:

Carnation	Hinge	Shark	Tail
Daisy	Horse	Soldier	Watch
Envelope	Leg	Sparrow	Worm
Frog	Penny	Suitcase	

### High Competition/Low Association Condition

Axe	Teeth
Baby	Tongue
Barbecue	
Binocular	
s	
Boat	
Candle	
Feet	
Heart	
Ice	
Key	
Moon	
Needle	
Pan	

Appendix C1: List of Items Included in the Name Agreement Pilot Task, with  
Response Strength Ratios

Acorn / 1.31	Box / 38
Alligator / 1.17	Bracelet / 16
Anchor / 100 % Agreement	Bread / 18
Ant / 14.5	Brow / 11
Apple / 100 % Agreement	Bucket / 100 % Agreement
Arm / 1.22	Bug / 1.69
Arrow / 38	Bus / 6.2
Artist / 1.35	Butterfly / 39
Astronaut / 17	Button / 39
Axe / 18.5	Cactus / 39
Baby / 6.8	Cage / 18.5
Backbone / 16	Cake / 100 % Agreement
Bag / 2.45	Calculator / 100 % Agreement
Balloon / 100 % Agreement	Calendar / 36
Banana / 100 % Agreement	Camel / 39
Barber / 11	Camera / 100 % Agreement
Barn / 9.33	Candle / 100 % Agreement
Barrel / 100 % Agreement	Car / 4.67
Basket / 100 % Agreement	Carousel / 1.43
Beaver / 2.14	Carrot / 100 % Agreement
Bed / 18.5	Cat / 39
Bee / 8.25	Cemetery / 1.43
Beer Mug / 17	Chair / 39
Bell / 39	Chest / 6.4
Bike / 2.45	Chicken / 1.05
Binoculars / 100 % Agreement	Chimney / 100 % Agreement
Blackboard / 9	Chopsticks / 38
Bones / 17.5	Cigar / 1.18
Book / 17.5	Clock / 39
Bow / 1.22	Cloud / 6.2
Bowl / 18	Clown / 100 % Agreement

Coat / 2.7	Guillotine / 14.5
Cockroach / 1.23	Guitar / 39
Collar / 34	Gun / 11
Comb / 100 % Agreement	Hall / 1.58
Cookies / 1.8	Handbag / 1.07
Corkscrew / 1.08	Harmonica / 14.5
Corn / 17.5	Harp / 6.8
Couch / 4.29	Helicopter / 18.5
Cradle / 1.11	Helmet / 19
Crown / 100 % Agreement	Hook / 8.75
Curler / 2.43	Horn / 3.22
Curtain / 12.33	Hourglass / 1.88
Deer / 3.25	House / 38
Dirt / 1.8	Hut / 2.14
Door / 100 % Agreement	Igloo / 18.5
Ear / 100 % Agreement	Iron / 100 % Agreement
Egg / 100 % Agreement	Jacket / 6.4
Eggplant / 3.13	Jigsaw / 2.44
Elephant / 100 % Agreement	Key / 100 % Agreement
Eye / 17.5	King / 15
Faucet / 38	Kite / 39
Finger / 12	Knife / 100 % Agreement
Foot / 39	Ladle / 5.8
Fork / 19	Leaf / 100 % Agreement
Foyer / 1.3	Lemon / 100 % Agreement
Frog / 18.5	Letter / 6.25
Garbage / 1.4	Lighthouse / 100 % Agreement
Giraffe / 38	Lion / 100 % Agreement
Grapes / 12	Lipstick / 100 % Agreement
Grass / 39	Lobster / 2.27
Grater / 17.5	Lollies / 6.2
Gravestone / 2	Maid / 11
Greenhouse / 1.77	Mallet / 1.06
Groceries / 2.71	Map / 8.75

Match / 3.88	Planet / 1.86
Medal / 39	Plate / 100 % Agreement
Minister / 2.33	Plug / 37
Money / 5.67	Pocket / 39
Monkey / 1.17	Porcupine / 18.5
Moon / 100 % Agreement	Potato / 39
Mosquito / 4	Prawn / 1.6
Motor / 4	Present / 1.86
Mule / 1.12	Prison / 1.11
Napkin / 1.14	Prisoner / 5
Neck / 100 % Agreement	Pumpkin / 100 % Agreement
Necklace / 38	Queen / 36
Net / 39	Rabbit / 18.5
Nose / 100 % Agreement	Ring / 8.5
Nun / 100 % Agreement	Road / 100 % Agreement
Nurse / 100 % Agreement	Rocket / 16
Octopus / 100 % Agreement	Rooster / 8
Onion / 100 % Agreement	Rug / 17
Orange / 6.8	Ruler / 100 % Agreement
Paddock / 3.8	Saddle / 36
Padlock / 2.08	Sandwich / 100 % Agreement
Pail / 12.33	Saw / 38
Paintbrush / 1.44	Scarecrow / 19
Palace / 4.83	Scissors / 18.5
Pants / 3	Screwdriver / 10.67
Peanut / 5.17	Seahorse / 100 % Agreement
Pear / 100 % Agreement	Shelf / 1.23
Pen / 100 % Agreement	Shell / 39
Pencil / 100 % Agreement	Ship / 2.63
Piano / 100 % Agreement	Shoe / 37
Pig / 19	Shop / 3.86
Pipe / 18	Shutter / 1.13
Pitcher / 4.4	Skateboard / 100 % Agreement
Pizza / 100 % Agreement	Skis / 18.5

Slide / 100 % Agreement	Tomato / 100 % Agreement
Slippers / 100 % Agreement	Tongue / 39
Smoke / 1.75	Toothbrush / 100 % Agreement
Snail / 36	Top / 1.4
Snowman / 100 % Agreement	Tornado / 3.75
Sock / 39	Train / 19
Soldier / 1.56	Trophy / 12
Spider / 8.75	Truck / 18.5
Spoon / 8.75	Turtle / 39
Steak / 1.2	Tuxedo / 11
Steps / 3.22	TV / 1.47
Stethoscope / 11	Typewriter / 38
Stick / 6.25	Tyre / 3.33
Stomach / 1.75	Umbrella / 100 % Agreement
Stones / 2.08	Unicorn / 100 % Agreement
Stool / 19	Valley / 11.33
Stove / 1.79	Wagon / 1.6
Strawberry / 100 % Agreement	Waiter / 100 % Agreement
Stream / 3.17	Watch / 100 % Agreement
Suitcase / 11.33	Whistle / 100 % Agreement
Suspenders / 1.56	Windmill / 19
Swing / 100 % Agreement	Window / 18.5
Tear / 2.33	Wineglass / 1.38
Telephone / 2.17	Witch / 17.5
Thumb / 39	Wolf / 4.14
Tie / 100 % Agreement	Wood / 1.82
Tights / 5	Wrist / 100 % Agreement
Toe / 4.67	Zebra / 100 % Agreement
Toilet / 38	Zucchini / 2.5

Appendix C2: Items Included in the Name Agreement Task by Competition  
Condition

Low Competition Item Names:

Anchor	Lion	Watch
Apple	Lipstick	Whistle
Balloon	Moon	Wrist
Banana	Neck	Zebra
Barrel	Nose	
Basket	Nun	
Binocular	Nurse	
s	Octopus	
Bucket	Onion	
Cake	Pear	
Calculator	Pen	
Camera	Pencil	
Candle	Piano	
Carrot	Pizza	
Chimney	Plate	
Clown	Pumpkin	
Comb	Road	
Crown	Ruler	
Door	Sandwich	
Ear	Seahorse	
Egg	Skateboard	
Elephant	Slide	
Iron	Snowman	
Key	Strawberry	
Knife	Swing	
Leaf	Tomato	
Lemon	Umbrella	
Lighthous	Unicorn	
e	Waiter	

## High Competition Item Names:

Artist	Mallet	Zucchini
Baby	Minister	
Bag	Money	
Bike	Mosquito	
Bug	Paddock	
Carousel	Padlock	
Cemetery	Paintbrush	
Coat	Pants	
Cockroac	Prawn	
h	Present	
Cookies	Prison	
Couch	Prisoner	
Cradle	Rooster	
Curler	Ship	
Deer	Shop	
Eggplant	Soldier	
Foyer	Steak	
Graveston	Steps	
e	Stomach	
Greenhou	Stones	
se	Stove	
Groceries	Stream	
Hall	Suspenders	
Handbag	Tear	
Hourglass	Telephone	
Hut	Tights	
Jacket	Top	
Jigsaw	Tornado	
Ladle	Tyre	
Lobster	Wagon	
Lollies	Wineglass	

Appendix D1: Sentence Stems and Distracters Included in the Sentence Completion Task by Competition Condition and Constraint Condition

Sentence Number	Sentence Stem	Low Competition Condition Distracter	High Competition Condition Distracter	Constraint Condition
1	The wooded lake made a pretty _____	Reach	Scene	Weak
2	In the first space enter your _____	Animal	Initials	Strong
3	Don found that he had no spare _____	South	Change	Weak
4	Barry wisely chose to pay the _____	Range	Fine	Weak
5	Dan caught the ball with his _____	Brood	Glove	Weak
6	Captain Sheir wanted to stay with the sinking _____	Pest	Raft	Strong
7	Without food a man would die in several _____	Halves	Weeks	Weak
8	The earth is shaped like a _____	Wish	Ball	Weak
9	The death of his dog was a great _____	Trail	Shock	Weak
10	The lawyer feared that his client was _____	Filling	Lying	Weak
11	You can't take the test without a _____	Bell	Pen	Weak
12	The pain she felt was all in her _____	Area	Mind	Weak
13	The birds in the yard ate every last _____	Match	Seed	Weak
14	When you go to bed turn off the _____	Progress	Radio	Strong
15	Every month Rick had to clean his _____	Word	Car	Weak
16	Most cats see very well at _____	Gaze	Climbing	Strong
17	The person who caught the thief deserves our _____	Kiss	Praise	Weak
18	My aunt likes to read the daily _____	Attorney	Newspaper	Weak
19	The long test left the class _____	Liquid	Tired	Weak
20	Bob would often sleep during his lunch _____	Key	Break	Weak
21	Few nations are now ruled by a _____	Recorder	Dictator	Weak



Sentence Number	Sentence Stem	Low Competition Condition Distracter	High Competition Condition Distracter	Constraint Condition
22	Some of the ashes dropped on the _____	Patch	Rug	Weak
23	Jean was glad the affair was _____	Battled	Finished	Strong
24	The surgeon tried vainly to save his _____	Concept	Patient	Weak
25	The old house will be torn _____	Beneath	Apart	Strong
26	He mailed the letter without a _____	Erosion	Signature	Strong
27	Jill looked back through the open _____	Chief	Window	Weak
28	At last the time for action had _____	Subjected	Arrived	Weak
29	Her job was easy most of the _____	Same	Day	Strong
30	The game was called when it started to _____	Mine	Snow	Strong
31	His leaving home amazed all his _____	Interest	Family	Weak
32	The surface of the water was nice and _____	Bound	Smooth	Weak
33	Lois is taller than most _____	Increase	Women	Weak

## Appendix D2: Additional Results for the Sentence Completion Task

*Accuracy by competition condition.* The following aphasics did not show a significant main effect of competition condition in their accuracy data (Non-fluents; CT ( $\chi^2(1) = .09$ ,  $p = .77$ ), DA ( $\chi^2(1) = 1.14$ ,  $p = .29$ ), JHM ( $\chi^2(1) = 2.27$ ,  $p = .13$ ), TB ( $\chi^2(1) = 0.37$ ,  $p = .54$ ), WS ( $\chi^2(1) = 1.29$ ,  $p = .26$ ): Fluents NP ( $\chi^2(1) = 1.60$ ,  $p = .21$ ), STR ( $\chi^2(1) = 3.00$ ,  $p = .08$ ), PS ( $\chi^2(1) = 2.67$ ,  $p = .10$ )).

*Accuracy by session.* None of the aphasics showed a significant main effect of session in their accuracy data (Non-fluents; CT ( $\chi^2(1) = 2.78$ ,  $p = .10$ ), DD ( $\chi^2(1) = 0.05$ ,  $p = .82$ ), DA ( $\chi^2(1) = 0.26$ ,  $p = .61$ ), ECV ( $\chi^2(1) = 0.26$ ,  $p = .61$ ), JHM ( $\chi^2(1) = 0.79$ ,  $p = .37$ ), TB ( $\chi^2(1) = 1.60$ ,  $p = .21$ ), WS ( $\chi^2(1) = 1.29$ ,  $p = .26$ ): Fluents; FS ( $\chi^2(1) = 0.42$ ,  $p = .52$ ), NP ( $\chi^2(1) = 0.37$ ,  $p = .54$ ), PS ( $\chi^2(1) = 0.78$ ,  $p = .38$ ), STR ( $\chi^2(1) = 0.43$ ,  $p = .51$ ), XX ( $\chi^2(1) = 0.13$ ,  $p = .72$ )).

*Accuracy competition condition by session interaction.* None of the aphasics showed a significant interaction of competition condition and session (Non-fluents; CT ( $\chi^2(1) = .17$ ,  $p = .68$ ), DD ( $\chi^2(1) = 0.88$ ,  $p = .35$ ), DA ( $\chi^2(1) = 0.10$ ,  $p = .75$ ), ECV ( $\chi^2(1) = 0.54$ ,  $p = .46$ ), JHM ( $\chi^2(1) = 0.86$ ,  $p = .36$ ), TB ( $\chi^2(1) = 1.95$ ,  $p = .16$ ), WS ( $\chi^2(1) = 0.29$ ,  $p = .59$ ): Fluents; FS ( $\chi^2(1) = 0.03$ ,  $p = .85$ ), NP ( $\chi^2(1) = 0.15$ ,  $p = .70$ ), PS ( $\chi^2(1) = 0.25$ ,  $p = .61$ ), STR (could not get model convergence), XX ( $\chi^2(1) = 0.28$ ,  $p = .60$ )).

## References

- Arpita, B., van Lieshout, P., & Square, P. A. (2007). Word frequency and bigram frequency effects on linguistic processing and speech motor performance in individuals with aphasia and normal speakers. *Journal of Neurolinguistics*, 20(1), 65-88.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modelling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390-412.
- Banich, M. T. (2004). *Cognitive Neuroscience and Neuropsychology* (2<sup>nd</sup> ed.). Boston, USA: Houghton Mifflin Company.
- Barcelo, F., Suwazono, S., & Knight, R. T. (2000). Prefrontal modulation of visual processing in humans. *Nature Neuroscience*, 3(4).
- Barker-Collo, S. L. (2001). The 60-item Boston Naming Test: Cultural bias and possible adaptations for New Zealand. *Aphasiology*, 15(1), 85-92.
- Berndt, R. S., Wayland, S., Rochon, E., Saffran, E., & Schwartz, M. (2000). *Quantitative production analysis: A training manual for the analysis of aphasic sentence production*. Hove, UK: Psychology Press.
- Biegler, K., Crowther, J. E., & Martin, R. C. (2006). Activation vs. inhibition accounts of semantic blocking effects in production and comprehension. *Brain and Language*, 99, 96 - 97.
- Bloom, P. A., & Fischler, I. (1980). Completion norms for 329 sentence contexts. *Memory and Cognition*, 8 (6), 631-642.
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict Monitoring and Cognitive Control. *Psychological Review*, 108, 3, 624-652.
- Breedin, S. D. & Saffran, E. M. (1999). Sentence Processing in the Face of Semantic Loss: A Case Study. *Journal of Experimental Psychology*, 128(4), 547-562.
- Brysbaert, M. (2007). *"The language-as-fixed-effect fallacy": Some simple SPSS solutions to a complex problem (Version 2.0)*. Royal Holloway, University of London.
- Butterworth, B. (1981). Speech errors: old data in search of new theories. *Linguistics*, 19, 627-662.

- Butterworth, B., Howard, D., & McLoughlin, P. (1984). The semantic deficit in aphasia: the relationship between semantic errors in auditory comprehension and picture naming. *Neuropsychologia*, 22 (4), 409-426.
- Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, 14, 177-208.
- Chan, A. H., Liu, H. L., Yip, V., Fox, P. T., Gao, J. H., & Tan, L. H. (2004). Neural systems for word meaning modulated by semantic ambiguity. *Neuroimage*, 22, 1128-1133.
- Copland, D. A., de Zubicaray, G.I., McMahon, K., Wilson, S. J., Eastburn, M., & Chenery, H.J. (2003). Brain activity during automatic semantic priming revealed by event-related functional magnetic resonance imaging. *Neuroimage*, 20, 302-310.
- Cohen, J.D. & Huston, T. A. (1994). Progress in the use of interactive models for understanding attention and performance. In C. Umiltà & M. Moscovitch (Eds.), *Attention and Performance XV* (pp. 453-456). Cambridge, MA: MIT Press.
- Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: A connectionist approach to behaviour and Biology in Schizophrenia. *Psychological Review*, 99 (1), 45-77.
- Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology*, 33A, 497-505.
- Crowther, J. E., & Martin, R. C. (2007). Further evidence for a post-selection inhibitory mechanism in lexical retrieval. *Brain and Language*, 103(1-2), 130-131.
- Crutch, S. J., & Warrington, E. K. (2005). Gradients of semantic relatedness and their contrasting explanations in refractory access and storage semantic impairments. *Cognitive Neuropsychology*, 22 (7), 851-876.
- Crutch, S. J., & Warrington, E. K. (2004). The semantic organisation of proper nouns: The case of people and brand names. *Neuropsychologia*, 42(5), 584-596
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93 (3), 283-321.
- Dell, G. A., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, 104 (4), 801-838.

- Desimone, R. & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18, 193-222.
- Donovan, K., Siegert, R., McDowell, J., & Abernethy, D. (1999). Clustering and switching in verbal fluency in Parkinson's disease. *New Zealand Journal of Psychology*, 28 (1), 61-66.
- Grindrod, C. M., Bilenko, N. Y., Myers, E. B., & Blumstein, S. E. (2008). The role of the left inferior frontal gyrus in implicit semantic competition and selection: An event-related fMRI study. *Brain Research*, 1229, 167-178.
- Dunn, Lloyd. M., & Dunn, Leota. M. (1997). *Peabody Picture Vocabulary Test: Third Edition*. Minnesota: American Guidance Service.
- Ellis, A. W., & Morrison, C. M. (1998). Real age-of-acquisition effects in lexical retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24 (2), 515-523.
- Freedman, M. L., Martin, R. C., & Biegler, K. (2004). Semantic relatedness effects in conjoined noun phrase production: Implications for the role of short-term memory. *Cognitive Neuropsychology*, 21, 245-265.
- Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), *The psychology of learning and motivation*. New York: Academic Press.
- Garrett, M. F. (1976). Syntactic processes in sentence production. In R. J. Wales & E. Walker (Eds.). *New Approaches to Language Mechanisms* (pp. 231-256). Amsterdam: North-Holland.
- Garrett, M. F. (1982). Production of speech: Observations from normal and pathological language use. In A. Ellis (Ed.), *Normality and pathology in Cognitive Functions* (pp. 19-76). London: Academic Press.
- Glaser, W. R., & Glaser, M. O. (1989). Context effects in Stroop-like word and picture processing. *Journal of Experimental Psychology: General*, 118, 13-42.
- Goodglass, H., Christiansen, J. A., & Gallagher, R. (1994). Syntactic constructions used by agrammatic speakers: Comparison with conduction aphasics and normals. *Neuropsychology*, 8(4), 598-613.
- Goodglass, H., Kaplan, E. & Barresi, B. (2001). *Boston Diagnostic Aphasia Examination: Third Edition*. Maryland, USA: Lippencott Williams & Wilkins.
- Grant, D. A., Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, 38(4), 404-411.

- Hamilton, A. C., & Martin, R. C. (2005). Dissociations among tasks involving inhibition: A single case study. *Cognitive, Affective, and Behavioural Neuroscience*, 5(1), 1-13.
- Hanes, K. R., Andrewes, D. G., & Pantelis, C. (1995). Dysfluency in Huntington's disease, Parkinson's disease, and schizophrenia. *Applied Neuropsychology*, 2, 29-34.
- Harley, T. (2001). *The Psychology of Language: From Data to Theory*. Hove: Psychology Press Ltd.
- Hillis, A. E., Rapp, B. C., Romani, C. & Caramazza, A. (1990). Selective impairment of semantics in lexical processing. *Cognitive Neuropsychology*, 7(3), 191-243.
- Hoenig, K., & Scheef, L. (2005). Mediotemporal contributions to semantic processing: fMRI evidence from ambiguity processing during semantic context verification. *Hippocampus*, 15, 597-609.
- Howard, D., & Gatehouse, C. (2006). Distinguishing semantic and lexical word retrieval deficits in people with aphasia. *Aphasiology*, 20, 921-950.
- Howard, D. & Orchard-Lisle, V. (1984). On the origin of semantic errors in naming: Evidence from the case of a global aphasic. *Cognitive Neuropsychology*, 1(2), 163-190.
- Howard, D. & Patterson, K. (1992). *The Pyramids and Palm Trees Test: A Test of Semantic Access from Words and Pictures*. Bury St Edmonds, England: Thames Valley Test Company.
- Hulleman, J. & Humphreys, G. W. (2007). Maximizing the power of comparing single cases against a control sample: An argument, a program for making comparisons, and a worked example from the Pyramids and Palm Trees Test. *Cognitive Neuropsychology*, 24(3), 279-291.
- Ihara, A., Hayakawa, T., Wei, Q., Munetsuna, S., & Fujimaki, N. (2007). Lexical access and selection of contextually appropriate meaning for ambiguous words. *Neuroimage*, 38, 576-588.
- Jefferies, E., Baker, S. S., Doran, M., & Lambon Ralph, M. A. (2007). Refractory effects in stroke aphasia: A consequence of poor semantic control. *Neuropsychologia*, 45, 1065-1079.
- Jefferies, E., & Lambon Ralph, M. A. (2006). Semantic impairment in stroke aphasia versus semantic dementia: A case-series comparison. *Brain: A Journal of Neurology*, 129(8), 2132-2147.

- Jescheniak, J. D. & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(4), 824-843.
- Kan, I. P., Kable, J. W., Van Scoyoc, A., Chatterjee, A., & Thompson-Schill, S.L. (2006). Fractionating the left frontal response to tools: Dissociable effects of motor experience and lexical competition. *Journal of Cognitive Neuroscience*, 18(2), 267-277.
- Kan, I. P., & Thompson-Schill, S. L. (2004). Effect of name agreement on prefrontal activity during overt and covert picture naming. *Cognitive, Affective, & Behavioural Neuroscience*, 4 (1), 45-57.
- Kay, J., Lesser, R., & Coltheart, M. (1992). *Psycholinguistic assessments of language processing in aphasia (PALPA): An Introduction*. Hove, UK: LEA
- Kempen, G. & Huijbers, P. (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition*, 14(2), 185-209.
- Kertesz, A. (1982). *Western Aphasia Battery*. New York: Grune & Stratton.
- Kimberg, D. Y. & Farah, M. J. (1993). A unified account of cognitive impairments following frontal lobe damage: The role of working memory in complex, organized behaviour. *Journal of Experimental Psychology: General*, 122(4), 411-428.
- Kimberg, D. Y. & Farah, M. J. (2000). Is there an inhibitory module in the prefrontal cortex? Working memory and the mechanisms underlying cognitive control. In S. Monsell & J. Driver (eds.). *Control of Cognitive Processes: Attention and Performance XVIII*. Cambridge: MIT Press.
- Kolk, H. (1995). A time-based approach to agrammatic production. *Brain and Language*, 50 (3), 282-303.
- Kolk, H. H. J. (1987). A theory of grammatical impairment in aphasia. In G. Kempen (Ed.), *Natural language generation: New results in artificial intelligence, psychology, and linguistics* (pp. 379-391). Dordrecht: Martinus Nijhoff.
- La Heij, W. (1988). Components of Stroop-like interference in picture naming. *Memory and Cognition*, 16, 400-410.

- Laiacina, M., Luzzatti, C., Zonca, G., Guarnaschilli, C. & Capitani, E. (2001). Lexical and semantic factors influencing picture naming in aphasia. *Brain and Cognition*, 46 (1-2), 184-187.
- Lahar, C. I., Tun, P. A., & Wingfield, A. (2004). Sentence-Final Word Completion Norms for Young, Middle-Aged, and Older Adults. *Journals of Gerontology: Series B: Psychological Sciences and Social Sciences*, 59B(1).
- Lambon Ralph, M. A., Sage, K. & Roberts, J. (2000). Classical anomia: A neuropsychological perspective on speech production. *Neuropsychologia*, 38(2), 186-202.
- Lee, S. S., & Dapretto, M. (2006). Metaphorical vs. literal word meanings: fMRI evidence against a selective role of the right hemisphere. *Neuroimage*, 29 (2), 536-544.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioural and Brain Sciences*, 22, 1-75.
- Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review*, 98 (1), 122-142.
- Lukkien, C. C. (2006). *A sentence production test for aphasia*. Unpublished doctoral dissertation, Victoria University of Wellington, New Zealand.
- MacKay, D. G. (1987). The organization of perception and action: A theory for language and other cognitive skills. Springer-Verlag: New York.
- Martin, R. C., & Byrne, M. D. (2006). Why opening a door is as easy as eating an apple: A reply to Thompson-Schill and Botvinick (2006). *Psychonomic Bulletin & Review*, 13 (3), 409 - 411.
- Martin, R. C., & Cheng, Y. (2006). Selection demands versus association strength in the verb generation task. *Psychonomic Bulletin & Review*, 13 (3), 396 - 401.
- Mason, R.A., & Just, M. A., (2007). Lexical ambiguity in sentence comprehension, *Brain Research*, 1146, 115-127.
- Max Planck Institute for Psycholinguistics (2001). *Celex Database*. Retrieved August 18, 2008, available from <http://www.mpi.nl/world/celex>.
- McCarthy, R. A., & Kartsounis, L. D. (2000). Wobbly words: refractory anomia with preserved semantics. *Neurocase*, 6, 487 – 497.
- Miller, E. K. & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual review of Neuroscience*, 24, 167-201.



- Milner, B. (1963). Effects of different brain lesions on card sorting: the role of the frontal lobes. *Archives of Neurology*, 9, 100-110.
- Milner, B. (1964). Some effects of frontal lobectomy in man. In J. M. Warren & K. Akert (Eds.), *Frontal granular cortex and behavior* (pp. 313-331). New York: McGraw-Hill.
- Mohr, J. P., Pessin, M. S., Finkelstein, S., Duncan, G. W., & Davis, K. R. (1978). Broca's aphasia: Pathologic and clinical. *Neurology*, 28, 311-324.
- Morrison, C. M., Chappel, T. D., & Ellis, A. W. (1997). Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables. *The Quarterly Journal of Experimental Psychology*, 50A, 528-559.
- Peers, P. V., Ludwig, C. J. H., Rorden, C., Cusack, R., Bonfiglioli, C., Bundesen, C., Driver, J., Antoun, N., Duncan, J. (2005). Attentional functions of parietal and frontal cortex. *Cerebral Cortex*, 15(10), 1469-1484.
- Perret, E. (1974). The left frontal lobe of man and the suppression of habitual responses in verbal categorical behaviour. *Neuropsychologia*, 12(3), 323-330.
- Randolf, C., Braun, A. R., Goldberg, T. E., & Chase, T. N. (1993). Semantic fluency in Alzheimer's, Parkinson's, and Huntington's disease: Dissociation of storage and retrieval failures. *Neuropsychology*, 7(1), 82-88.
- Rapp, B. & Goldrick, M. (2000). Discreteness and Interactivity in Spoken Word Production. *Psychological Review*, 107 (3), 460-499.
- Robinson, G., Shallice, & Cipolotti, L. (2005). A failure of high level verbal response selection in progressive dynamic aphasia. *Cognitive Neuropsychology*, 22(6), 661-694.
- Rodd, J. M., Davis, M. H., & Johnsrude, I. S. (2005). The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cerebral Cortex*, 15, 1261-1269.
- Roelofs, A. (1992). A spreading activation theory of lemma retrieval in speaking. *Cognition*, 42, 107-142.
- Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, 64, 249-284.
- Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (1989). The quantitative-analysis of agrammatic production – procedure and data. *Brain and Language*, 37 (3), 440-479.

- Saffran, E. M., Schwartz, M. F., Linebarger, M., Martin, N. & Bochetto, P. (1988). *Philadelphia Comprehension Battery*. Unpublished test.
- Schnur, T. T., Schwartz, M. F., Brecher, A., & Hodgson, C. (2006). Semantic interference during blocked-cyclic naming: Evidence from aphasia. *Journal of Memory and Language*, 54, 199-227.
- Schwartz, M. F., & Hodgson, C. (2002). A new multiword naming deficit: Evidence and interpretation. *Cognitive Neuropsychology*, 19 (3), 263-288.
- Schwartz, M. F., Linebarger, M. C., & Saffran, E. M. (1985). The status of the syntactic deficit theory of agrammatism. In M-L. Kean (Ed.). *Agrammatism* (pp. 83-104). New York: Academic Press.
- Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96 (4), 523-568.
- Shallice, T., & Burgess, P. (1996). The domain of supervisory processes and temporal organization of behaviour. *Phil. Trans. R. Soc. Lond. B*, 351, 1405-1412.
- Shatzman, K. B., & Schiller, N. O. (2004). The word frequency effect in picture naming: Contrasting two hypotheses using homonym pictures. *Brain and Language*, 90, 163-169.
- Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 174-215.
- Spreen, O., & Strauss, E. (1991). *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary*. NY: OUP.
- Starreveld, P. A., & La Heij, W. (1995). Semantic interference, orthographic facilitation and their interaction in naming tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 686-698.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643-662.
- Stuss, D. T., & Benson, D. F. (1984). Neuropsychological studies of the frontal lobes. *Psychological Bulletin*, 95 (1), 3-28.
- Székely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., Jacobsen, T., & Bates, E. (2003). Timed picture naming: Extended norms and

- validation against previous studies. *Behavior Research Methods, Instruments & Computers*. Vol 35(4), 621-633.
- Thompson-Schill, S. L. (2005). Dissecting the language organ: A new look at the role of Broca's area in language processing. In A. Cutler (Ed.), *Twenty-first Century Psycholinguistics: Four Cornerstones*. New Jersey: Lawrence Erlbaum Associates, Publishers.
- Thompson-Schill, S. L., Bedny, M., & Goldberg, R. F. (2005). The frontal lobes and the regulation of mental activity. *Current Opinion in Neurobiology*, 15, 219 - 224.
- Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. A., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proc. Natl. Acad. Sci.*, 94, 14792-14797.
- Thompson-Schill, S. L., D'Esposito, M., Kan, I. P. (1999). Effects of repetition and competition on activity in left prefrontal cortex during word generation. *Neuron*, 23(3), 513-522.
- Thompson-Schill, S. L., Swick, D., Farah, M.J., D'Esposito, M., Kan, I. P., & Knight, R. T. (1998). Verb generation in patients with focal frontal lesions: A neuropsychological test of neuroimaging findings. *Proceedings of the National Academy of Sciences*, 95, 15855-15860.
- Troyer, A. K., Moscovitch, M., Winocur, G., Alexander, M. P., & Stuss, D. Clustering and switching on verbal fluency: the effects of focal frontal- and temporal-lobe lesions. *Neuropsychologia*, 36(6), 499-504.
- Vitkovitch, M., & Tyrrell, L. (1995). Sources of disagreement in object naming. *The Quarterly Journal of Experimental Psychology*, 48A (4), 822-848.
- Wagner, A. D., Paré-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning: Left prefrontal cortex guides controlled semantic retrieval. *Neuron*, 31, 329 – 338.
- Weidner, R., Pollmann, S., Müller, H. J., & von Cramon, D. Y. (2002). Top-down controlled visual dimension weighting: An event-related fMRI study. *Cerebral Cortex*, 12, 318-328.
- Williams, S. E., & Canter, G. J. (1982). The influence of situational context on naming performance in aphasic syndromes. *Brain and Language*, 17(1), 92-106.

- Wilshire, C. E., & McCarthy R. A. (2002). Evidence for a context-sensitive word retrieval disorder in a case of non-fluent aphasia. *Cognitive Neuropsychology*, 19(2), 165-186.
- Wilshire, C. E, Scott, R., & Stuart. E. J. (2006). Effects of competitor words on production in a case of non-fluent aphasia. *Brain and Language*, 99, 92-93.
- Zempleni, M. Z., Renken, R., Hoeks, J. C. J., Hoogduin, J. M., & Stowe, L. A. (2007). Semantic ambiguity processing in sentence context: evidence from event related fMRI. *Neuroimage*, 34, 1270-1279.
- Zingeser, L. B., & Berndt, R. S. (1988). Grammatical class and context effects in a case of Pure Anomia: Implications for models of language production. *Cognitive Neuropsychology*, 3, 474-513.
- Zingeser, L. B., & Berndt, R. S. (1990). Retrieval of Nouns and Verbs in Agrammatism and Anomia. *Brain and Language*, 39, 14-32.