

THE EFFECT OF LEXICAL CONTENT ON SENTENCE PRODUCTION  
IN NONFLUENT APHASIA

by

Paula Speer

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

2014

## Abstract

Individuals with *nonfluent aphasia* are able to produce many words in isolation, but have great difficulty producing sentences. Most research to date has compared accuracy across different types of sentence structures, focussing on grammatical aspects that may be compromised in nonfluent aphasia. However, based on the premise that lexical elements activate their associated grammatical frames as well as vice versa, lexical content may also be of vital importance. For example, rapid access to lexical elements – particularly ones appearing early in the sentence - may be crucial, especially if the sentence plan is weakly activated or rapidly decaying. The current study investigated the effect of different aspects of lexical content on nonfluent aphasic sentence production. Five participants with nonfluent aphasia, four participants with fluent aphasia and eight controls completed two picture description tasks eliciting subject-verb-object sentences (e.g., *the dog is chasing the fox*). Based on existing evidence suggesting that common words are accessed more rapidly than rarer ones, Experiment 1 manipulated the frequency of sentence nouns, thereby varying their speed of lexical retrieval by varying the frequency of sentence nouns. Nonfluent participants' accuracy was consistently higher for sentences commencing with a high frequency subject noun, even when errors on those nouns were themselves excluded. This was not the case for the fluent participants. Experiment 2 manipulated the semantic relationship between subject and object nouns. Previous research suggests that phrases containing related words may be challenging for individuals with nonfluent aphasia, possibly because lexical representations are inadequately tied to appropriate structural representations. The nonfluent participants produced sentences less accurately when they contained related lexical items, even when those items were in different noun phrases. The fluent participants exhibited the opposite trend. Finally, the relationship between the patterns observed in Experiment 1 and 2 and lesion location in the aphasic participants was explored by analysing magnetic resonance

scans. We discuss the implications of our findings for theoretical accounts of sentence production more generally, and of nonfluent aphasia in particular. More precisely, we propose that individuals with nonfluent aphasia are disproportionately reliant on activated lexical representations to drive the sentence generation process, an idea we call the *Content Drives Structure* (COST) hypothesis.

## Acknowledgments

*‘Mā te huruhuru ka rere te manu’*

Many people have contributed to the completion of this dissertation and I am deeply grateful for their support. First, I would like to thank all my participants for the generous time and effort they spent to take part in this research. Thank you in particular to my participants with aphasia, who welcomed me warmly into their homes. I have learned so much from them.

My supervisor, Dr. Carolyn Wilshire, deserves a very special thank you. Working with her has been truly inspirational and I am grateful for the many hours we spent discussing this project. Without her expertise, trust, and good sense of humour, this work would not have been possible. Further, I am thankful to Josh Faulkner and Bridget Burmester for help with data collection. I would also like to thank my fellow students Adrienne Girling, Clara Cantal, Jane Goulding and Pollyane Diniz for their moral support and friendship. Thank you to Victoria University of Wellington for the financial support provided by a Victoria University PhD Scholarship (2010-2013) and two travel grants (VUWSA 2011/FSRG 2012).

Outside of the university, I was fortunate to meet many caring people in New Zealand. I would like to thank two people in particular: John MacLeod, who has been looking after me since the moment I set foot in New Zealand, and who welcomed me so warmly into his family; and Kararaina McLean, who not only brought me close to New Zealand's indigenous culture, but who also reminded me that the mind is our most powerful asset. And there are my precious friends and family back home who have always stayed in touch and supported me from a distance. I am especially grateful for the unending love and encouragement from my wonderful mother Roswitha Heller.

Finally, I thank my husband Johannes Oeffner for his abundant love and endless support – practically, financially, and emotionally. Thank you for sharing my joy during the ups and for drying my tears during the downs. I could not have done it without you.

## Contents

Abstract .....	ii
Acknowledgments.....	iv
List of Tables .....	vi
List of Figures .....	vii
Introduction .....	1
Chapter 1: Theories of Normal Sentence Production .....	3
Chapter 2: Aphasia.....	35
Chapter 3: Theories of Nonfluent Aphasic Sentence Production .....	42
Chapter 4: Introduction to the Current Study .....	59
Chapter 5: Experiment 1.....	64
Chapter 6: Experiment 2.....	121
Chapter 7: Lesion Analysis .....	144
Chapter 8: General Discussion .....	158
Appendices:	
List of Appendix Tables .....	175
Appendix A: Stimulus Material for Experiment 1 .....	177
Appendix B: Supplementary Data and Analyses for Experiment 1 .....	181
Appendix C: Stimulus Material for Experiment 2 .....	188
Appendix D: Supplementary Data and Analyses for Experiment 2.....	190
References .....	198

## List of Tables

Table 5.1 .....	72
<i>Background, medical and diagnostic information for each aphasic participant</i>	
Table 5.2 .....	73
<i>Performance of participants with aphasia on language and related cognitive tests</i>	
Table 5.3 .....	75
<i>Spontaneous speech sample of aphasic participants describing the Cookie Theft Scene in the BDAE (Goodglass et al., 2001)</i>	
Table 5.4 .....	94
<i>Description of the three different latency measures</i>	
Table 5.5 .....	98
<i>Pre-test naming errors for aphasic individuals and the three participant groups in the different frequency conditions, collapsed across all three sessions</i>	
Table 5.6 .....	103
<i>Percentage of errors for participants with aphasia and the three participant groups for the different frequency conditions</i>	
Table 6.1 .....	129
<i>Percentage of errors for participants with aphasia and the three participant groups for the different experimental conditions</i>	

## List of Figures

<i>Figure 1.1</i> Garrett's sentence production model (adapted from Garrett 1975, 1976, 1982).....	5
<i>Figure 1.2</i> Simplified depiction of the WEAVER ++ model (Levelt et al., 1999)..	9
<i>Figure 1.3</i> Simplified depiction of Dell's spreading activation model (Dell, 1986).....	14
<i>Figure 1.4</i> Stemberger's (1985) Interactive Activation Model of Language Production .....	21
<i>Figure 1.5</i> Bock's model of sentence formulation (adapted from Bock, 1982).....	24
<i>Figure 1.6</i> Simplified representation of the Dual-path model (Chang et al., 2006).....	28
<i>Figure 2.1</i> Left lateral view of the surface of the brain. ....	38
<i>Figure 2.2</i> Lesion overlay maps of individuals with different subtypes of nonfluent and fluent chronic aphasia (adapted from Dronkers et al., 2009). ....	39
<i>Figure 5.1</i> Examples of the pictured scenes from different experimental conditions. ....	67
<i>Figure 5.2</i> Lesion maps for participants with nonfluent aphasia showing axial slices of the brain.....	76
<i>Figure 5.3</i> Lesion maps for participants with fluent aphasia showing axial slices of the brain.....	77
<i>Figure 5.4</i> Lesions maps for individuals with aphasia showing the left lateral side of the brain .....	78
<i>Figure 5.5</i> Geometric mean latencies in the pre-test for aphasic individuals and the nonfluent, fluent and control group across the frequency conditions. ....	100
<i>Figure 5.6</i> Percentage of nouns incorrectly produced in subject and object position for nonfluent and fluent aphasic individuals and the three participant groups.....	104
<i>Figure 5.7</i> Percentage of overall sentence errors and sentence errors not penalising for subject errors as a function of subject frequency for nonfluent and fluent aphasic individuals and the three participant groups.....	106
<i>Figure 5.8</i> Subject and object frequency effects on sentence initiation time and verb-object noun initiation time for nonfluent and fluent aphasic individuals and the three participant groups.....	110
<i>Figure 6.1</i> Percentage of nouns incorrectly produced in subject and object position for nonfluent and fluent aphasic individuals and the three participant groups.....	130
<i>Figure 6.2</i> Percentage of overall sentence errors and combined errors on subject and object nouns as a function of relatedness for nonfluent and fluent aphasic individuals and the three participant groups.....	132

<i>Figure 6.3</i> Noun relatedness effects on sentence initiation times and verb-object noun initiation times for nonfluent and fluent aphasic individuals and the three participant groups.....	136
<i>Figure 7.1</i> Lesion overlay maps for individuals with nonfluent aphasia.....	151
<i>Figure 7.2</i> Lesion overlay maps for individuals with fluent aphasia .....	152
<i>Figure 7.3</i> Representative axial slices showing the distribution of statistical power. ....	153
<i>Figure 7.4</i> Axial slices from VLSM maps computed for total subject noun frequency effect .....	154
<i>Figure 7.5</i> Axial slices from VLSM maps computed for total noun relatedness effect .....	155



## Introduction

According to the World Health Organization, 15 million people worldwide suffer a stroke every year (Mackay & Mensah, 2004). Approximately 40% of such cases result in loss of some language abilities, or *aphasia*, which can severely limit social interaction and lead to a significant loss of quality of life. The efficiency of behaviourally oriented treatments for aphasia, such as speech language therapy, can be improved by gaining more insight into the underlying cognitive processes that are impaired in such cases.

The condition of interest in the current research is a subtype known as *nonfluent* aphasia. In nonfluent aphasia, utterances are often reduced to one or two words and are separated by long pauses. People suffering from this condition have difficulties producing well-formed sentences, even though single word naming can be relatively well preserved (e.g., McCarthy & Kartsounis, 2000; Schwartz & Hodgson, 2002; Williams & Canter, 1982). For example, individuals with nonfluent aphasia may be able to correctly name pictures that depict a single item, but experience difficulties when asked to integrate the named items into a sentence to describe a pictured event.

The current research aims to examine the cognitive limitations that could underlie these sentence production difficulties in individuals with nonfluent aphasia. Most existing research into this difficulty has focussed on purely grammatical aspects of sentence production (e.g., Caramazza & Hillis, 1989; Goodglass, Christiansen, & Gallagher, 1994). Here, we take an alternative approach that emphasises the importance of the *interaction* of structure and content during sentence production. There is supporting evidence for this view from studies showing that structural processes in healthy speakers are influenced by the availability of lexical content (e.g., Bock, 1986a, 1987b; F. Ferreira & Engelhardt, 2006). While not many studies have systematically examined effects of lexical variables on sentence production in nonfluent aphasia, some research suggests that nonfluent aphasic speech can be

greatly influenced by the availability of lexical content elements and their semantic relationship (e.g., Freedman, Martin, & Biegler, 2004; Linebarger, McCall, Virata, & Berndt, 2007). The current research aims to investigate the influence of these different aspects of lexical content on sentence production using a picture description task, and manipulating: a) lexical frequency of sentence nouns and b) the semantic relationship between those nouns; thereby varying the ease and speed with which the constituent nouns can be retrieved. In addition, participants with aphasia were asked to undergo high resolution magnetic resonance (MRI) scans, allowing an investigation of whether damage to specific cortical structures was consistently associated with particular performance patterns in the experimental tasks.

Before the current research is described in more depth, the following chapter first provides an overview of current models of normal language production. Reviewing models of how healthy speakers produce sentences is important to set the stage for a better understanding of the different processes involved in normal sentence production, and how these processes may be compromised in nonfluent aphasia. Further, several models of normal sentence production provide valuable information of how lexical content and structural processes may interact during sentence planning, providing a background for comprehending why lexical content may have a crucial effect on nonfluent aphasic sentence production. Next, theoretical accounts of sentence production difficulties in nonfluent aphasia will be discussed, with a main focus on how these theories address and manage to explain the symptoms commonly associated with this disorder. Finally, previous research that may speak to the issue of how lexical content may impact on nonfluent aphasic sentence production will be examined, providing the theoretical basis for the current study.

## **Theories of Normal Sentence Production**

The process of sentence production is highly complex. The speaker needs to access multiple words simultaneously, specify the relation between the words and choose the correct grammatical inflections and tenses – all the time, holding the sentence elements actively in mind until it is time to produce them. Amazingly, these challenges are mastered by normal speakers without a lot of difficulty. But how do we transform an intention to communicate into a well-formed, fully articulated sentence? What processes are involved and do they interact with each other? To shed light on the underlying structures of sentence production, the most influential models of sentence production will be reviewed and discussed in the following.

### **Garrett's (1975) Two-Stage Model of Sentence Production**

One of the first persons to tackle the task of creating a model of sentence production was Merrill Garrett in 1975. He investigated a large corpus of naturally occurring speech errors and found that they occur in meaningful patterns. Moreover, these patterns were consistent when exploring the speech output of a lot of different people. For example, Garrett could identify two main types of movement errors. The first type of error involved the exchange of words, so that each occupies the other's place in the sentence (e.g., *Do you want some tea in your sugar?*). The words participating in these exchanges tended to share the same word-class (exchange of nouns for nouns, verbs for verbs, etc.) and could span many intervening words, and even whole phrases. However, there appeared to be no consistent relationship between the sound properties of the words.

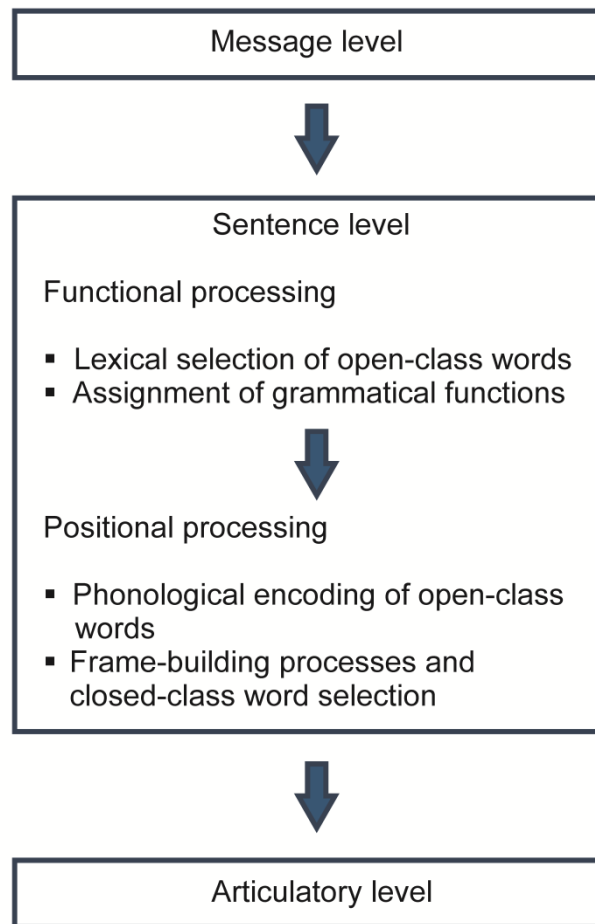
The second type of error involved the movement of word sounds (e.g., *Do you have matches to fight the lire?*) or even whole morphemes (also called *stranding errors*, e.g., *There is a stick booking out*). In contrast to word exchange errors, the sound and morpheme errors tended to occur most often on adjacent words (usually within one phrase) and did not have to

share the same word-class. Also they were highly dependent on the sound properties of the utterance: For example, a large number of phoneme exchanges involved phonemes that differed only in one distinctive feature, such as voicing (e.g., /b/ vs. /p/), place of articulation (e.g., /k/ vs. /g/) or manner of articulation (e.g., /t/ vs. /s/). These findings, taken together, led Garrett to the conclusion that there must be two major stages of word production during sentence formulation: one in which words are represented in terms of their grammatical class, but not their sound form, and another where the words were fully specified as to their sound form and appropriate morphological inflections had been added. Garrett combined the idea of a two-stage process of word production with a two-stage process of sentence planning.

According to Garrett (1975, 1976, 1982), the speech production process involves three major components (see Figure 1.1): It begins at a preverbal message level (consisting of a non-linguistic representation of a concept that the speaker intends to express), followed by a sentence level consisting of two separate stages, and it ends at an articulatory level (where motor commands guide the vocal tracts to produce the sound structure of the utterance). As the sentence level is of central relevance for the current work, it will be explained in more detail in the following. Sentence planning begins at the *functional processing* stage. At this stage, lexical selection of open-class words<sup>1</sup> takes place (for example, *dog*, *chase*, *cat*), which are then assigned grammatical functioning (for example, *subject: dog; object: cat; verb (transitive): chase*). Whole word exchange errors occur here, when words are incorrectly assigned to their grammatical functions, or semantic substitution errors, when the wrong word gets selected. This stage is not sensitive to the phonological properties of the words or their serial order in the eventual utterance, which explains why word exchange errors are not dependent on sounds and can move quite far.

---

<sup>1</sup> Open-class words (synonymous with content words) carry meaning and include nouns, verbs, adjectives, adverbs etc.



*Figure 1.1.* Garrett's sentence production model (adapted from Garrett 1975, 1976, 1982, 1984; modified for simplification). Arrows represent the flow of activation.

The second stage of sentence planning is the *positional processing* stage. Here, the phonological forms of the selected open-class words are retrieved, and these fully phonologically specified words are organised into a syntactic planning frame. Based on their functional roles specified at the previous stage, the frame consists of “slots” that define the correct sequential order of the content words (Garrett, 1980). Phoneme errors occur at this stage. Since word order is represented at the positional stage, and also phonological content, this explains why phoneme movement errors are more likely to occur in proximal words and in a phonologically similar environment, while they show no sensitivity for the same word class (Garrett, 1975, 1976).

Bound morphemes<sup>2</sup>, such as -ed or -ing, are also selected at the positional processing stage. The evidence supporting this proposal comes from stranding errors in which two open-class lexical elements exchange, leaving behind a bound morpheme (e.g., *I am not in the read for mooding*). Garrett argued that these stranding errors only ever involved bound morphemes, usually “syntactically active” ones – that is, those indicating tense or number (Garrett, 1975). These findings suggest that the selection of closed-class words<sup>3</sup> may occur at a later stage of processing to that of open-class words (Garrett, 1982). Further, this evidence suggests that a complete syntactic frame for the sentence is not created until the positional processing stage. This frame specifies the closed-class elements, as well as the appropriate inflectional affixes and the prosodic structure of the sentence, and also contains the “slots” for each lexical content element based on the information specified at the functional level.

Importantly, Garrett’s model postulates a strict top-down flow of activation from higher to lower stages of processing (Garrett, 1976). Consequently, the processes occurring at the functional and positional processing stages do not interact. The evidence Garrett cited in support of this claim is that word exchange errors show sensitivity to the participating words’ grammatical form (which is specified at the functional level) but not to properties specified at the lower, position level, such as phonological form. In addition, Garrett cited a further piece of evidence supporting this non-interactive view: He was able to show that substitutions were either related to the intended word by form (*confirmation* instead of *conservation*, both words consisting of similar phonemes and having the same number of syllables) or meaning (*husband* instead of *wife*), but never both (examples taken from Garrett, 1976, p. 246).

---

<sup>2</sup> Bound morphemes cannot stand alone and can only appear with another word (e.g., inflections and affixes). In contrast, free morphemes can stand alone and include content and function words.

<sup>3</sup> Closed-class words (synonymous with function words) serve syntactic purposes and include determiners, pronouns, and prepositions.

Garrett's model of sentence production laid the foundation for later models, many of which share the same basic processing stages, but not necessarily the postulation of a strict top-down flow of information. We will come back to this point later.

### **The Importance of Verbs: Verb-Based and other Lexically-Driven Theories**

Another finding from speech error analysis has concerned the role of the verb in the production process. Hotopf (1980, as cited in Bock & Levelt, 1994) noticed that the verb of a sentence was hardly ever involved in substitution errors, a finding which may reflect its central role in the sentence production process, making it resistant to error. For example, the meaning of the verb seems to restrict the range of appropriate candidates. Jarvella and Sinnott (1972) examined the effect of the verb on noun selection. They compared the choice of nouns in subject and object positions for verbs with different selectional rules as to animacy. Participants were asked to fill in a subject and an object in a cloze procedure, while the verb was given (e.g., *The\_\_ thanked the\_\_*). Verbs that usually go along with animate nouns elicited the production of more animate nouns, while verbs that allow for inanimate nouns caused a decrease in animacy in the produced nouns. Also, nouns produced in the subject position seemed to be more animate than the nouns produced for the object position.

More recent research has specifically investigated the effect of verb argument structure on sentence processing. Verb argument structure refers to the number and types of entities that can participate in the event described by the verb, which can range from one (e.g., *The boy runs*) to three (e.g., *The boy gave the present to the girl*). This information seems to be crucial for the speaker to construct an appropriate syntactic frame for a sentence, and many studies have shown that information about the verb's argument structure becomes readily available during sentence processing (e.g., Ahrens, 2003; Shapiro & Levine, 1987, 1990; Shapiro, Zurif, & Grimshaw, 1987; Trueswell & Kim, 1998). For example, Shapiro and Levine (1987) asked participants to listen to a sentence for meaning and to perform a

visual lexical decision at the same time. In the critical condition, the lexical decision task was presented immediately after the verb. In this condition, the authors observed longer reaction times when the key verb had a complex argument structure than when it had a simpler one. The authors argued that all potential argument structure representations become activated during verb processing, and that this leads to a higher processing load. Taken together, these findings suggest that the verb of a sentence can influence both the potential lexical content and the structure used to convey it.

A theory of sentence production that particularly ascribes a central role to the verb was proposed by Willem Levelt (1989, 1999). Similar to Garrett's model (1975), the WEAVER++ model of Levelt and colleagues (1999) assumes the existence of three distinct components: the concept, lemma and form component (or *stratum*). Each component contains different types of nodes that are interconnected with each other. Information is retrieved via a process of *spreading activation* amongst the interconnected nodes. For example, if the concept node of *dog* is activated by a picture, it will automatically send activation to its corresponding lemma node on the lemma stratum, which will then further activate its constituent morpheme and phoneme nodes in the form component. The total activation level of each node depends on the activation coming from its interconnected nodes. If the sum of this activation exceeds that of all other nodes on that particular component, the node is selected for production. However, other units than the target unit may be highly activated at the same time. This may be caused because they share conceptual features with the target, or by random activation noise in the model. The co-activation of other units in the network leads to *competition* for production in the network, with the possible result that an unintended, but highly activated word will be produced instead of the target word.

Figure 1.2 displays the structure of the WEAVER++ model with an example. The lexical concept nodes on the conceptual stratum share links to other conceptually related



nodes (e.g., *swim* as a subordinate relative to *move*, or *hunt* as an equivalent action to *chase*).

Lexical concept nodes also store information about the conceptual arguments they can take (for example, the concept node of *hunt* specifies that it can take two semantic arguments).

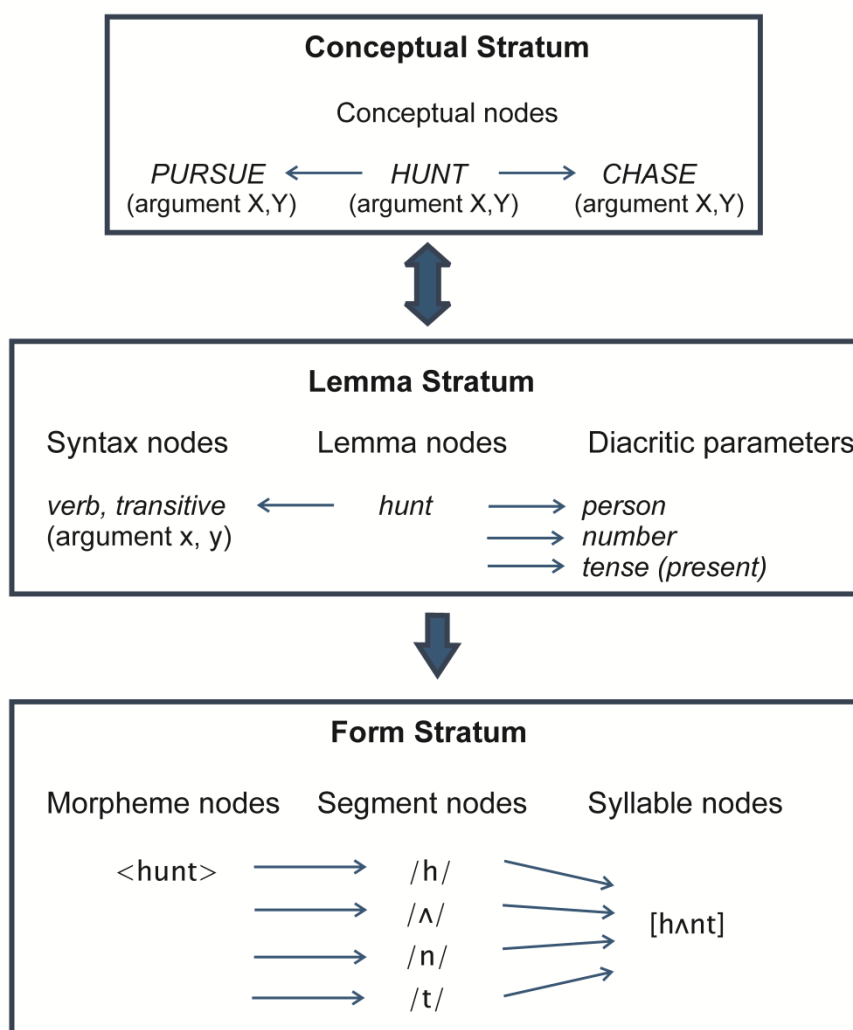


Figure 1.2. Simplified depiction of the WEAVER ++ model, using the verb “hunt” as an example (Levelt et al., 1999). Arrows represent the flow of activation.

Once a lexical concept node on the concept stratum is selected, it feeds activation down to the next level, the lemma stratum. The lemma nodes represent lexical elements, whose meanings are defined by their connection to their corresponding concept nodes on the concept stratum. Most importantly for the process of sentence production, lemma nodes are

linked to syntax nodes that specify their syntactic properties. As mentioned before, the verb lemma plays a crucial role in the production model since it defines the argument structure of a sentence. For example, the verb lemma *hunt* is linked to a transitive verb node that specifies that it can carry two argument positions. The verb lemma is further linked to nodes that specify the appropriate inflectional features (so called diacritic parameters), depending on tense, number and person of the subject as indicated by the conceptual message. In the model, both function and content words are represented by lemmas. However, while content word lemmas receive their activation from the concept nodes that specify their meaning, many function word lemmas do not have a strong conceptual representation. Selection of function word lemmas is therefore driven by their syntactic properties, although the model is not explicit about the exact selection mechanism.

The lowest level of the WEAVER ++ model, the form stratum, represents the specific sound structure of a word. It contains the morpheme nodes, which receive activation from their corresponding lemma nodes. Morpheme nodes are linked to corresponding phonemic segment nodes that contain information about the spelling of the word. The segment nodes are ordered with regards to their position in the word, which is important because they are not exclusively linked to specific morpheme nodes but also engage in the spelling of other words. The segment nodes activate appropriate syllable nodes, which ensure the correct syllabification according to the context in which the word is produced.

While the WEAVER ++ model remains rather silent about specific syntactic structure building processes, an earlier sentence production model of Levelt (1989) gives some indication of how grammatical encoding may take place. The model postulates a completely lexically driven grammar, with all the necessary syntactic information stored in the lemmas. The syntactic frame is constructed around the verb lemma, consisting of slots for each of the arguments that are stored with it. The arguments are then mapped onto the respective

grammatical functions (e.g., subject, object). For each selected lemma, the syntactic category that it belongs to activates a “categorical procedure” instruction as to where the lemma will go in the syntactic planning frame. For example, only a noun can go at the head of a noun phrase. The selected lemma will be assigned to the left most appropriate slot, while the right part of the sentence needs to be maintained active until it is time to produce it. If a sentence contains two nouns, the noun lemma that is higher activated will be assigned to the first noun slot. Depending on whether the subject or the object noun lemma is higher activated, an active or a passive structure will be generated. The specification of lemma nodes and the assumption of a lexically driven syntax differ from Garrett’s sentence production model (1975), which claims separate syntactic and lexical processes.

### **Sequential or Interactive Stages? Dell’s (1986) Model of Sentence Production**

Both Levelt’s model (1989) and the WEAVER ++ model (Levelt et al., 1999) largely conform with Garrett’s postulation of distinct production stages and feed-forward activation. That is, once a representation is selected, its activation only proceeds down to lower levels of production, so that processing at one level does not commence until processing is complete at the next highest level. This means that processing at early levels cannot be influenced by processes that occur further downstream: For example, the process of lemma retrieval should not be influenced by the phonological properties of the words being retrieved.

Nevertheless, findings from both natural and experimental speech error data challenge a strict top-down flow of activation and strongly support the existence of mutual interactions between the representations at different levels. For example, semantic substitution errors have been found to share a similar sound structure with the target word more frequently than would be expected by chance (the *mixed error effect*, e.g., producing *rat* not *dog* instead of *cat*; example taken from Martin, 1996; see also Del Viso, Igoa, & Garcia-Albea, 1991; Dell & Reich, 1981; Fromkin, 1971). This finding indicates that sound representations on the

lower production level (Garrett's positional processing level and Levelt's form stratum) seem to influence the semantic processes on the higher production levels (Garrett's functional processing level and Levelt's lemma stratum). In a model that permits interactive flow of activation, words that are similar in sound receive activation from the phonological nodes that they share, and are so more likely to get confused than those with little phonological overlap. Also, the interaction of meaning and sound exists in the opposite direction, too: There seems to be a lexical bias concerning meaningful combinations of sound units. Phonological errors are more likely to create existing words or morphemes rather than nonwords (e.g., producing *bed* and not *ked* for *red*; Baars, Motley, & MacKay, 1975; Dell, 1986; Dell & Reich, 1981; Stemmer, 1984, but see Garrett, 1976). Again, an interactive model can account for these findings: Activation from the phonological level may flow upward, which can only result in activating real words or morphemes, because nodes for nonwords do not exist.

However, it has to be noted that during computer simulation using the *WEAVER ++* model (Levelt et al., 1999), a disproportionally high lexical bias effect in speech errors was found, even though the model does not allow for a feedback loop. Levelt and colleagues (1999) suggest that this occurs due to a monitoring system that is incorporated in the *WEAVER ++* model, the role of which is to compare intended and actual output and to screen both for inconsistencies. If a speech error accidentally results in a real word, the output editor will be less likely to detect and correct it since the phonetic plan is not faulty. A similar explanation is used to account for the mixed error effect: A word that is phonologically AND semantically related to the target words is less likely to be noticed by the monitor due to its high perceptual similarity to the target word (Levelt et al., 1999). However, some have argued that such an output editor would add too much complexity to a production model (e.g., Dell, 1986; Stemmer, 1985), when the same phenomena can effortlessly be explained by an interactive account.

Based on the observations reported above, Dell (1986) proposed one such interactive model of sentence production. He identified three major levels: a *syntactic*, a *morphological*, and a *phonological* level. The evidence Dell used to support the idea of a separate morphological level was the observation that inflectional morphemes can take part in errors, and when they do they often violate syntactic rules (e.g., *look atting* instead of *looking at*; example taken from Stemberger, 1982), suggesting they occur after syntactic frame construction. At the same time, they are qualitatively different from phonological errors because the whole inflection moves in its entirety.

As shown in Figure 1.3, the different levels in Dell's model are connected to a lexical network of nodes for words, morphemes and phonemes, which themselves are interconnected via excitatory links and retrieved through the process of spreading activation. At each level, a specific insertion rule ensures that only acceptable combinations of nodes are selected and produced according to their category (for example, the correct sequence of onset, nucleus and coda on the phonological level). To ensure that activation of a selected target node effectively activates its corresponding nodes at the higher and lower levels, it receives a further boost of what Dell calls *signalling activation* once it is selected (Dell, 1986). This mechanism applies to each level in the model, and significantly enhances the activation of the selected node over all others as well as the activation of its corresponding nodes on different levels.

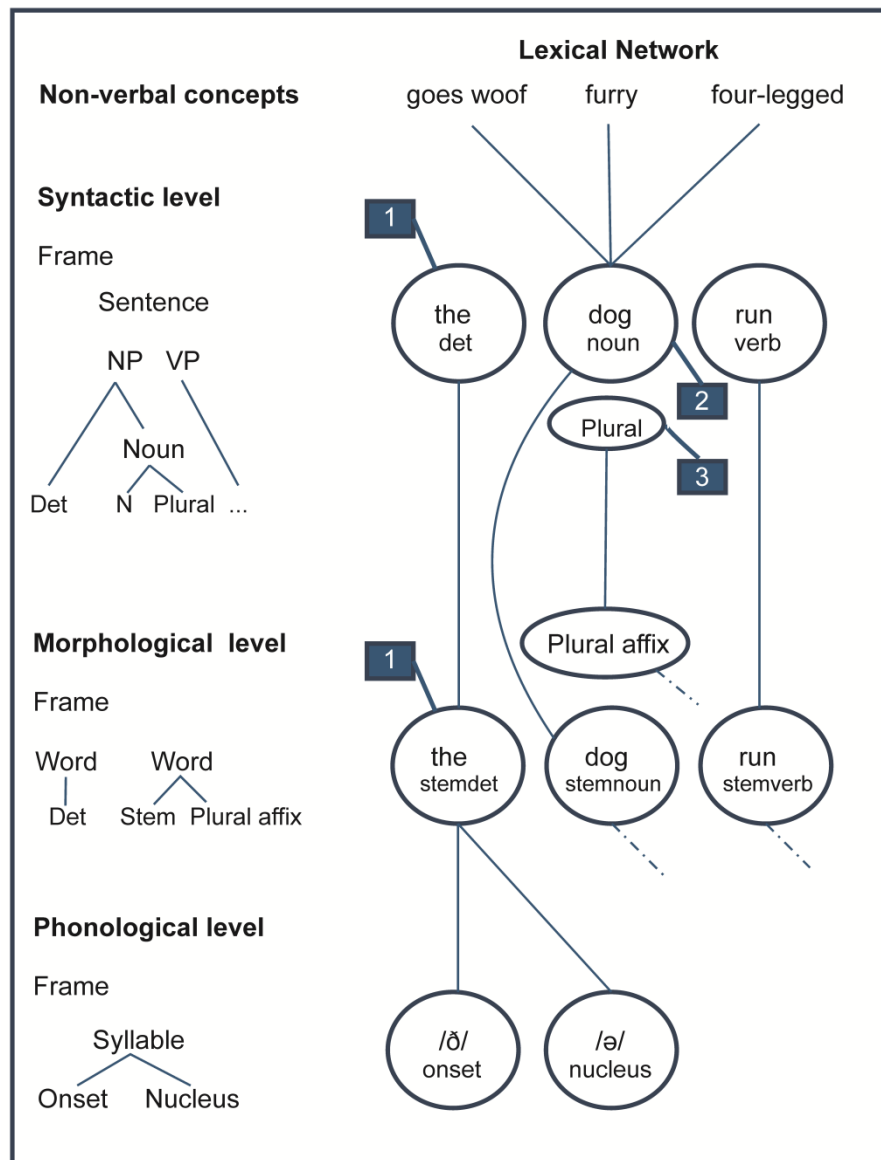


Figure 1.3. Simplified depiction of Dell's spreading activation model, using the sentence "the dogs run" as an example (Dell, 1986).

Let us consider how a sentence is produced in Dell's model (see Figure 1.3 for an example). The first step of the sentence production process in the model is similar to the models described so far (Garrett, 1975; Levelt, 1989; Levelt et al., 1999): It begins with a nonverbal-conceptual message (e.g., a semantic representation of two domestic animals which are moving fast). This message in turn activates word nodes that define the corresponding lexical elements (e.g., *dog* and *run*). Word nodes are linked to a syntactic

frame that consists of slots specifying the sequence in which the words need to be ordered (e.g., determiner – noun + plural – verb). The word node with the highest activation is selected for production by the insertion rule and will be tagged to its specific position in the frame according to its syntactic category (e.g., the noun will be assigned the second position in a determiner-noun-verb frame). Furthermore, a node specifying the quantity of the noun may be selected and subsequently tagged (e.g., a plural node for *dogs*). The tagging process orders all word nodes that form part of the sentence and ensures that they take the correct position in the frame. A node can receive multiple tags if it appears several times in the sentence. At the time of insertion into the syntactic frame, the word node receives a further boost of signalling activation, which significantly enhances its activation above all others. The node's increased activation then feeds down to the next level, activating the corresponding morphological representations.

As at the syntactic level, an insertion rule ensures that the morphological node with the highest activation is inserted into a morphological frame at the appropriate position (e.g., *dog* will be inserted into the stem-slot). Next, the most highly activated affix will be tagged and inserted (in this example, the plural affix for the noun). Finally, at the phonological level, phonemes that correspond to the selected morpheme nodes are activated and integrated into a phonological frame via the insertion rule, which specifies the different slots that code the sound structure of a word (onset, nucleus, and coda). Once the selection process is completed at one level, the activation level of a node is reset to zero. Selection of a “new” higher level unit will follow (e.g., *run*) and initiate the same process until sentence production is completed (e.g., *the dogs run*). Importantly, content and function words in the model are processed in an identical way on the phonological level, as opposed to Garrett's model (1975), which proposes that content words are selected during functional processing, whilst function words are specified during positional processing.

Speech errors in Dell's model occur due to general activation noise, which is attributed to the spreading nature of the model: "Random fluctuation" and feedback from other interacting levels may lead to under- or over-activation of nodes' resting levels. If a target word's resting level is below its usual threshold, more activation is needed for its production, and the selection might be more prone to errors because of the activation surplus of other competing nodes. In agreement with that, a resting level above the threshold for non-target units would make their erroneous selection very likely. Consequently, the syntactic category constraint observed in word exchanges (e.g., *Would you like some tea in your sugar?*) can be attributed to a higher activation level of the second noun (*tea*) at the time for tagging the first noun (*sugar*) for its position in the syntactic frame. Since nodes are tagged by the specific insertion rules which ensure that only acceptable combinations of nodes at every level get selected and produced, only elements of the same syntactic (e.g., noun), morphologic (e.g., suffix), or phonological (e.g., vowel) category can swap their position in the frames, which fits nicely with the previously discussed patterns of speech errors.

However, although Dell's (1986) model can elegantly account for many speech error phenomena, there are also challenges to his model: For example, it fails to account for the fact that sound errors mostly occur in content words rather than in function words, because it does not distinguish between the processing of these two types of items (as opposite to Garrett, 1975, for example). Dell suggests that the high frequency of most function words may account for their resistance to error, but this suggestion remains speculative. Furthermore, the model remains silent about the specific processes that guide insertion rules on the different levels, and about how the highest (nonverbal-conceptual) and lowest (articulatory) processes of the production system are organised.



## Evidence for Interactivity of Structure and Content

While the models described so far make different assumptions regarding the degree of interaction between the particular *stages* of the production system, they all have in common the proposal that they do not assume a full interaction between the *lexical* and *structure* components. However, findings from both speech error analyses and from experimental studies provide strong evidence that syntactic processes can indeed affect lexical processes and vice versa throughout the sentence production process.

For example, many studies have shown that specific aspects of lexical content can influence the grammatical structure chosen for a sentence (Bock, 1986a, 1986b; F. Ferreira & Engelhardt, 2006; Konopka & Bock, 2009; Levelt & Kelter, 1982; Pickering & Branigan, 1998; Stemberger, 1985). Let us consider evidence from language experiments first. In lexical priming studies, a lexical item required for a particular sentence is primed ahead of time (e.g., *dog* in *the cat chased the dog*). Many studies showed that speakers are more likely to select a sentence plan that places this element early in the sentence, even when it means opting for a non-canonical structure (e.g., *the dog was chased by the cat*; Bock, 1986a, 1987; F. Ferreira & Engelhardt, 2006; Konopka & Bock, 2008; Levelt & Kelter, 1982; Pickering & Branigan, 1998; Slevc, 2011). This finding can be accounted for in Levelt's model (1989, 1999), because the primed lemma will have a higher activation and therefore be assigned to the first noun position in the sentence. However, there is also evidence that the sound of a lexical item may have an influence on the sentence structure: Bock (1986b, 1987) was able to show that the priming of a phonological structure (e.g., *dot* for *dog*) led participants to produce the target word later in the sentence when describing a picture (e.g., *the cat and the dog are sleeping*; but see Cleland & Pickering, 2003, for a study that failed to replicate this finding).

V. S. Ferreira and Firato (2002) provided evidence that the availability effect on syntactic structures could extend to the use of optional words, like complementisers (*that*; see also V. S. Ferreira & Dell, 2000). They asked participants to recall sentences that began with a main subject including three noun phrases, and a single embedded-subject noun phrase, which allowed for an optional production of *that* immediately before the embedded subject (e.g., *The man, the woman, and the child heard (that) the dog was barking*). The conceptual similarity of the embedded-subject noun phrase to the main-subject noun phrases was manipulated. Based on the premise that conceptual similarity between the two subject noun phrases would induce proactive interference and slow down retrieval processes, V. S. Ferreira and Firato predicted that in order to gain more processing time, speakers would produce more *thats* in sentences that contained conceptually similar main-subject and embedded-subject noun phrases (e.g., *The author, the poet, and the biographer recognized (that) the writer was boring*), than if the phrases were dissimilar (e.g., *The author, the poet, and the biographer recognized (that) the golfer was boring*). This is exactly what was found.

Findings from syntactic priming studies offer further support for the interaction of structure and content. In syntactic priming studies, a prime sentence is presented to – or generated by – the participant, and then subsequently a target sentence is elicited, often using a picture event description task. Many studies have shown that participants are more likely to generate target sentences using the structure of the primed sentence (e.g., Bock, 1986b; Levelt & Kelter, 1982). In other words, the primed syntactic structure generalises across different lexical elements, influencing the way that lexical content is sequenced. The syntactic priming effects have been shown to even persist over the production of a series of target sentences, indicating that structural information can have a long-lasting effect on the order of lexical content (e.g., Bock & Griffin, 2000; Hartsuiker & Kolk, 1998). However, Pickering and Branigan (1998) found that syntactic priming was not completely independent

of the lexical content of the prime and target sentences: They found that participants showed an increased syntactic priming effect when the verb in both sentences was identical. This phenomenon, also called a *lexical repetition effect* or *lexical boost*, demonstrates that lexical information can influence the selection of syntactic structure (see also Konopka & Bock, 2008). Lexical variables also seem to influence syntactic encoding in other ways. For example, James, Thompson, and Baldwin (1973) asked participants to recall verbally presented sentences which contained lexical elements of varying concreteness. They found that sentences containing a concrete subject (e.g., *burglar*) were more accurately recalled than sentences containing an abstract subject (e.g., *rumour*). Furthermore, sentences with an abstract subject were more likely to be recalled in a passive sentence, thereby allowing the concrete, and therefore more accessible, item to fill the position of the subject.

Finally, evidence for an interaction between content and structure also comes from speech error analysis. So-called *bumper car errors* (Stemberger, 1985) are characterised by the occurrence of both an anticipation error and a substitution error in the same sentence (e.g., *There is a web in the net* instead of *There is a spider in the web*) and show a clear interaction of structural sequencing (the target word is inserted into the wrong slot) and lexical access (the substitution of the target words with a meaning-related word).

### **Stemberger's (1985) Fully Interactive Model of Sentence Production**

Based on findings from speech error analysis, Joseph Stemberger introduced an interactive activation model of language production (Stemberger, 1985). Similar to Dell's theory (1986), the model incorporates spreading activation as the retrieval mechanism and proposes the existence of an interconnected network of nodes that allows for interaction between all levels of the production process. However, unlike Dell, Stemberger proposes that connections between the nodes in the network can be excitatory (mainly between levels) OR inhibitory (mainly within one level). If a node is highly activated, it will send inhibitory

activation through to other competing nodes on the same level (e.g., nodes that might also be activated due to shared semantic features with the target node). Selection for production is thereby ensured in two ways: The target word will have the highest activation level and therefore also the strongest inhibitory effect. This phenomenon of lateral inhibition is what Stemmer calls “the rich get richer, the poor get poorer” (Stemmer, 1985, p. 146) and serves the same purpose as Dell’s signalling activation mechanism.

As shown in Figure 1.4, Stemmer’s model (1985) consists of four main levels: a semantic level, a lexical level, a syntactic level, and a phonological level (including several sub-components). Like Dell (1986) and Garrett (1975), Stemmer proposes a frame-and-slot system at the syntactic level for the ordering of multiple words in an utterance. However, in Stemmer’s model, the frames and slots also operate according to the principles of spreading activation. So-called *phrase-structure units* become activated depending upon the thematic roles and grammatical relations that are specified at the semantic level (e.g., agent, recipient, etc.). Once the appropriate phrase-structure units are selected, they in turn activate their consistent phrases (e.g., subject noun phrase, verb phrase). Also, the phrase-structure units themselves are interconnected in such a way that they always become activated in a syntactically appropriate manner. For example, in English, first the initial noun phrase of a sentence will be highly activated and selected for production, and then activation will spread to the verb phrase, and potentially to another noun phrase and so on. When the most appropriate syntactic structure for a message is chosen, it will stay active in working memory until the full message can be produced.

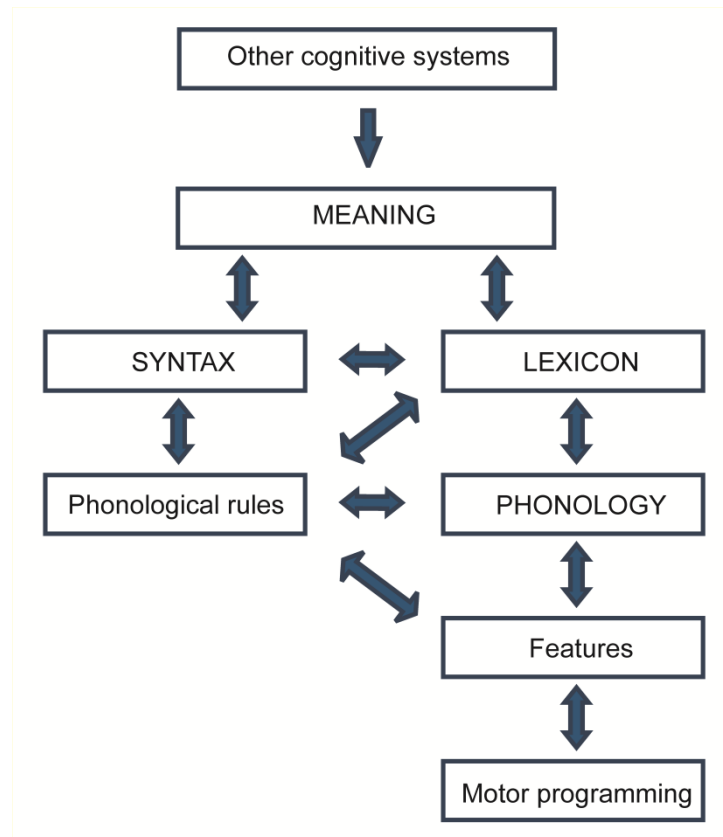


Figure 1.4. Stemmerger's (1985) Interactive Activation Model of Language Production.

Arrows indicate flow of activation.

One of the most important aspects in Stemmerger's model is the proposal that the lexical and the syntactic level are highly interactive, and that processing of words and structure occurs simultaneously. This means that phrase-structure units can influence the selection of lexical elements and vice versa. Selected phrases require lexical units of a certain category to fill the slots in the phrase frame and they send activation to all suitable candidates for the phrase position. For example, a selected noun phrase (e.g., *the dog*) will first send activation to the determiner, and then to lexical elements marked as nouns, while inhibiting all unsuitable candidates (e.g., adjectives). Since the target noun receives further activation from the semantic feature nodes, its activation will exceed that of other activated elements in the noun category. This proposal fits nicely with the observed category constraint in word

exchange errors. As mentioned before, Stemberger argues that not only syntactic structure can influence lexical retrieval, but also lexical retrieval can influence syntactic structure. In other words, each lexical element is capable of activating its corresponding slot in the frame. This means that currently active lexical content has the capacity to influence the overall syntactic plan.

One final aspect of Stemberger's model worth noting concerns how it deals with the distinction between open- and closed-class elements. Similar to Dell (1986), the model represents open- and closed-class items in an identical fashion. However, open-class elements receive activation primarily from semantic feature nodes, whereas closed-class elements receive proportionally more activation from phrase-structure units. Like Dell, Stemberger explains the less frequent occurrence of closed-class items in speech errors by their higher frequency and therefore higher resting levels of activation.

### **Capacity Limitations in Sentence Production**

The models of sentence production described so far consider how different types of lexical elements and syntactic structures become selected during sentence planning, and how or whether these processes interact. However, they do not address in any detail how the various overlapping processes are coordinated and managed so that they do not exceed the limits of a speaker's mental processing capacity. There is support for the idea that speech production draws on general processing resources (Ford & Holmes, 1978; Goldman-Eisler & Cohen, 1970; Power, 1986). For example, Power (1986) asked participants to talk about certain topics and perform a tracking task in isolation. On some trials, participants were required to simultaneously talk and track. Results showed that tracking errors increased when participants were talking and tracking at the same time, compared to when they only performed the tracking task in isolation. This finding indicates that speech production requires resources from a capacity-limited system, which may then not be available for

accurately performing a secondary task. Furthermore, other research shows that speech is more fluent when speakers are given time to plan the content of an utterance in advance than when they have to produce speech impromptu (Deese, 1978, 1980). The length of pauses during speech is also longer before syntactically complex utterances than before simpler ones (Ford, 1982). Finally, speakers' tendency to place available lexical elements and concepts in an early position in a sentence and thus avoiding to hold information active for longer than necessary points to a capacity-reducing strategy during language production (see above, e.g., Bock, 1986a, 1987; F. Ferreira & Engelhardt, 2006; Slevc, 2011).

Bock (1982) attended to this matter and emphasised the need for a “capacity-limited processing system” of language production. She postulates five arenas in her framework of sentence formulation, some of which have direct links to working memory (synonymous with central processing capacity in the following). Figure 1.5 gives an outline of the model's architecture. Bock's model aligns with interactive models such as those by Dell (1986) and Stemmerger (1985), and allows for backward flow of activation between linguistic representations. Similar to other models, communication starts in the *referential arena*. Here, the nonverbal idea that is intended to be communicated is transformed into a propositional format, in which the various conceptual components are individually represented, such as the agent of the action, the object involved in the action, etc. This process is proposed to draw heavily on working memory capacity. Once the propositional format of the message is created, it simultaneously proceeds into both the *lexical* and *syntactic arena*.

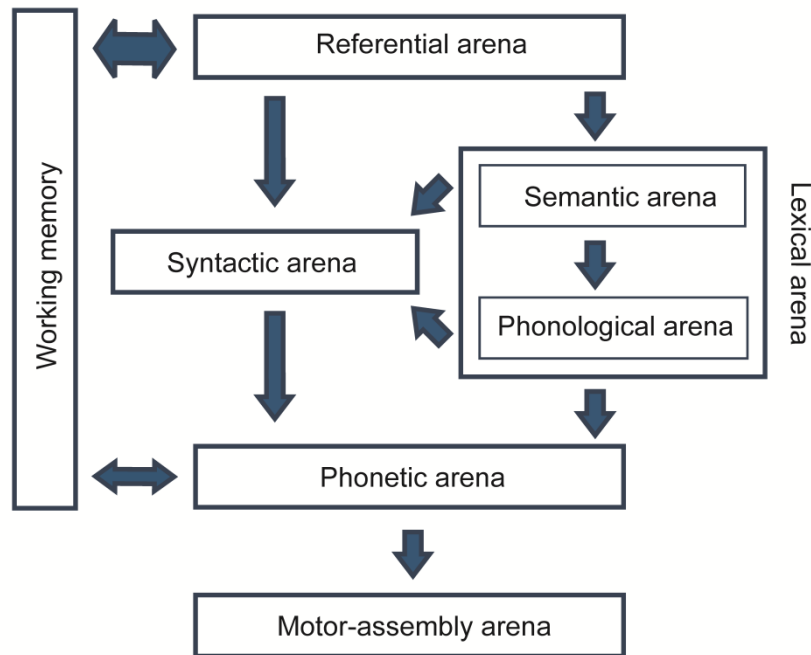


Figure 1.5. Bock's model of sentence formulation (adapted from Bock, 1982). Thickness of arrows pointing to the working memory is indicative of the degree that the arenas draw on processing capacity.

The lexical arena deals with the specifics of individual content word retrieval. First, the concepts that comprise the nonverbal proposition are matched with *lexical concepts*, which carry both meaning and syntactic properties of a word. This process is guided by the grammatical class of the lexical concept (e.g., the agent of an action is usually assigned a noun). Bock postulates that the simpler the mapping between semantic information and its corresponding lexical concept, the faster the formulation process becomes. For example, matching the nonverbal concept of an apple to its lexical representation should be fast due to its singularity, while matching a nonverbal concept of an unfamiliar object may be slower due to many possible lexical representations (see Chafe, 1977, for a study supporting this suggestion). Second, the lexical concepts are translated into a phonological representation.



Bock (1982) argues that this aspect of lexical processing has become automatized in (adult) speakers through repeated practice, and thus does not require any processing capacity.

Output from the referential and lexical arenas feeds into the syntactic arena. Based on the intended message and the lexical elements required to convey it, a syntactic frame consisting of slots for the individual elements is created. The frame generation occurs in parallel with lexical processing, and is also considered automatic, not drawing on processing capacity of the working memory.

In the *phonetic arena*, the outcome of lexical and syntactic processes is coordinated. The frame slots that were generated in the syntactic arena are filled with the appropriate phonological representations, which are further phonetically encoded. Here, the order of the words is established. Like processes in the referential arena, phonetic processing is proposed to require working memory capacity, because the syntactic frame needs to be held active in working memory while phoneme retrieval takes place. Finally, the phonetic representation feeds into the *motor-assembly arena* that generates motor plans to drive the movement of muscles in the vocal tract.

The unique contribution of Bock's sentence production model is that it specifically addresses the processes that draw on capacity during speech production, and to what extent this occurs. Within this framework, the reason that syntactic structure selection can be influenced by lexical availability is that it reduces overall processing demands - more readily accessible words (be they recently primed, or simply more common or concrete) are placed in an early position in the sentence to avoid maintaining the word's activation until it is time to produce it, which would draw unnecessarily on processing resources.

When compared to the other models described so far, Bock shares Dell's (1986) and Stemmerger's (1985) proposal that processing at different stages in language production can occur at the same time. However, Bock's model does not place a strong emphasis on

specifying the structure of representation units in the different arenas of the model, and does not give any details about the principles of activation spreading in and between the different arenas. Furthermore, and unlike Stemmerger, Bock only proposes a unidirectional flow from the lexical arena to the syntactic component, and does not specifically address a potential impact of syntax on lexical processes. Furthermore, Bock does not address the nature of closed-class and open-class elements, and in what way these elements are represented in the language system.

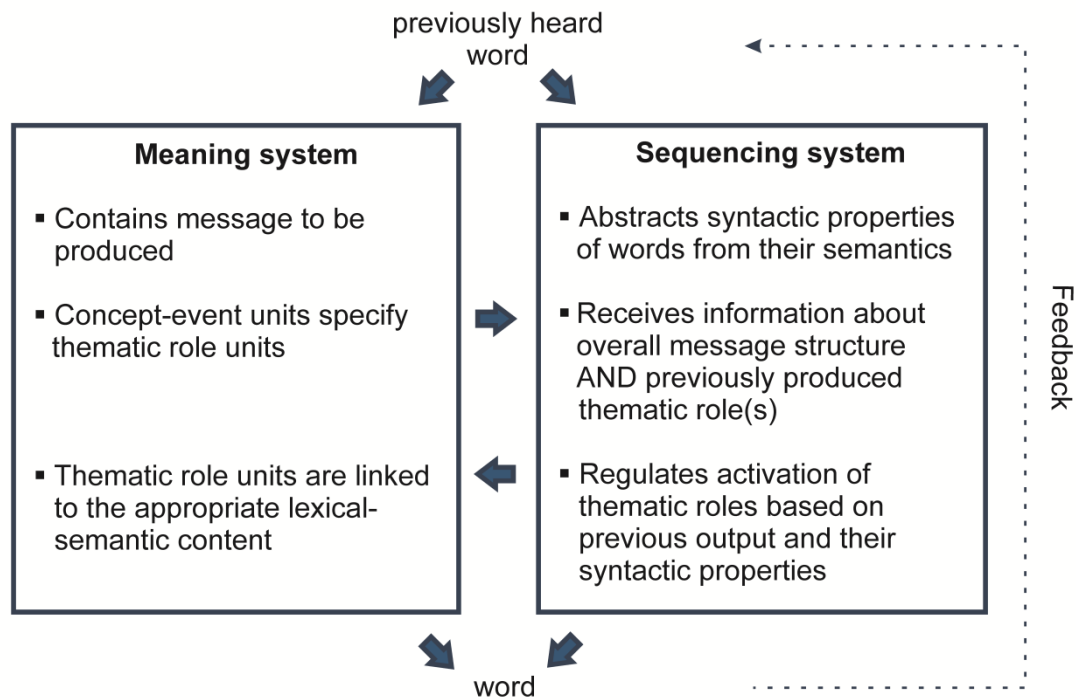
### **Division of Labour Theories**

So far, we have seen that models of sentence production differ to the extent to which they assume an interaction between structure and content retrieval processes during speech production. In some models, structure generation and lexical content retrieval processes occur completely independently (e.g., Garrett, 1975). In other models, there is an interactive flow of information between lexical selection processes and structure generation during sentence planning (e.g., Bock, 1982, Stemmerger, 1985). However, in a third class of models, there are distinct syntactic and conceptual components which independently contribute to the activation of candidate lexical elements, but which are capable of sharing the labour involved in lexical element retrieval. One example of such a model is the Dual-path (DP) model of Chang, Dell, and Bock (2006). In the DP model, the grammatical properties of words and their rules of combination are not specified by the researchers, but rather are learned by the model after many exposures to a wide range of sentences. The model learns to produce sentences incrementally, one word at a time, selecting each successive word on the basis of two sources of input: 1) information about the meaning of the current message; and 2) information about the sequencing rules that apply, given the grammatical class of the element being produced and that of the previously-produced items. Crucially, the model has access to information about the last few words produced (this information is fed back into the model

after each word production event), which it uses to further constrain production of the next upcoming word.

The broad structure of the Dual-path model is shown in Figure 1.6 (Chang et al., 2006). Because the model produces sentences incrementally, each word constitutes a single action in the model. During sentence production, each word at the output layer is activated via two different systems: the meaning system and the sequencing system. The meaning system specifies the overall conceptual-thematic structure of the planned message in the form of a set of concept-event units. They consist of conceptual entities (semantic descriptions) linked to their thematic roles. So, for the sentence “*The dog is chasing the cat*”, the concept-event units would include AGENT-*cat*, PATIENT-*dog*, and so on. This linking process is implemented by a temporary increase between the connection weights that bind the two elements. This feature enables the model to flexibly assign different roles to the same concept, or the same role to different concepts, to account for the endless possibilities of messages to be conveyed.

Similar to other models of word production, the concept units are linked to their corresponding verbal labels (e.g., *cat*). Importantly, in the model, the strength of activation of each concept-event unit can vary, and this can influence the ultimate word order in the sentence. For example, if the patient unit is more activated than the agent unit (for example, because the patient is the most important element of the sentence), it will be more likely to be linked to an early position in the sentence and signal for a passive construction.



*Figure 1.6.* Simplified representation of the Dual-path model (adapted from Chang et al., 2006). This figure represents how a single word is selected for production in a sentence, depending upon the meaning representations that are currently activated (concept-event units) and the grammatical properties of preceding words that belong to the same sentence.

The sequencing system has only limited access to the meaning system. Its role is to modulate each word production event so that it conforms to the grammatical rules of the language. As noted earlier, it learns these grammatical rules via an error-driven learning process – if an element is produced that is grammatically incorrect in the current sentence context, adjustments will be made to the relevant connections to reduce the likelihood of such an error in the future. Crucially, the sequencing system does not learn to link specific words to their appropriate grammatical positions, but rather their associated thematic roles (which are themselves linked to the specific words). This gives it the capacity to generalise its learning to new words that occupy the same thematic role in the future.

Again, the sequencing mechanism works in an incremental, word-by-word manner. For example, at the start of the sequencing process, it selects the most highly activated thematic role unit from the meaning system whose grammatical class conforms to the sequencing rules, and begins to generate predictions as to the next permissible grammatical category. The information about the model's output is then fed back into the system so that it "knows" about that recently produced item. The model then selects the next most highly activated thematic role unit whose grammatical class is consistent with its predictions, and then produces it, and so on. As sentence construction progresses, information about previously produced words is continually fed back into the system, providing information about the current context and further constraining the word selection process at each juncture. Importantly, the sequencing system ensures that each successive word fits the grammatical sequencing rules by activating thematic roles whose grammatical properties are appropriate for that position, and activating roles whose properties are. This is an idea similar to Stemberger's (1985) proposal that phrase-structure units send activation to all appropriate lexical candidates and inhibit all others.

An important aspect in the DP model is that the sequencing system has no direct contact with the meaning system (only to the thematic roles), resulting in an abstract representation of syntactic knowledge that has no associations with conceptual or lexical variables. The model can account for a number of empirical phenomena. For example, it can account for the syntactic priming effect: the finding that the structure chosen for a particular sentence can be influenced by the structure of a previously-uttered sentence, even if it contains none of the same actual word content (e.g., Bock & Griffin, 2000; Saffran & Martin, 1997). Chang and colleagues (2006) argue that the same mechanism that helps the DP model learn sequences is responsible for this effect: Connection weights in the model change accordingly to the structures they are exposed to, making the use of these structures more

likely to occur in the future. In fact, results from computer simulation showed that the Dual-path model reliably produced structural priming data similar to human data (Chang et al., 2006).

Nevertheless, the DP model does have difficulty accounting for the enhanced syntactic priming effect that is observed when there are shared lexical elements between the prime and target sentence (e.g., Branigan & Pickering, 1998; Branigan, Pickering, & Cleland, 2000; Cleland & Pickering, 2003; Konopka & Bock, 2008). As mentioned above, these findings challenge the proposal of a complete separation of the meaning and syntax component in general, and specifically, in the DP model.

### **The Scope of Planning in Sentence Production**

This review has considered in detail some of the major theoretical approaches to sentence production. However, little has been said about the *scope* of the planning processes involved in sentence production. How much do speakers plan ahead when they produce a sentence? Do they plan the entire sentence in its entirety before speaking? Or only one or two words at a time? Many creative paradigms have been developed to measure and to tap into the scope of sentence planning, and findings are discussed below.

Some evidence from language experiments suggests a clausal scope at higher levels of sentence planning (e.g., Bock & Cutting, 1992; Ford & Holmes, 1978; Meyer, 1996). For example, Ford and Holmes (1978) interviewed participants on certain topics and let them simultaneously react to a tone. The time it took participants to react to the tone while talking was thought to be indicative of the degree of language processing happening at that time. Results showed that reaction times increased if the tone was presented at the end of a clause. The authors interpreted this finding as an indicator that the next clause was in preparation, which would require processing resources that were then not available for responding to the tone, and supporting the idea of a clausal unit of planning in speech production. A similar

conclusion was reached by Meyer (1996) using quite a different experimental paradigm. Meyer presented participants with pictures displaying objects that either had to be described in a sentence (e.g., *the snail is next to the hill*) or as a noun phrase conjunction (e.g., *snail and hill*). The pictures were accompanied by an auditory distractor. Meyer found that if the distractor was semantically related to either of the nouns (e.g., *worm* or *meadow*), latencies for both sentence production and conjoined noun phrase production increased. This finding suggests that there is advance planning of both lexical elements in the utterance before speech onset, and supports a clausal scope of planning.

In line with these findings, evidence from early speech error studies also suggests a relatively wide scope of planning, whose exact size depends on the current stage of sentence planning (Dell, 1986; Garrett, 1975, 1980). Garrett (1980) observed that in slips of the tongue, word exchange errors usually occur within one clause (e.g., *Put the fridge back into the cheese*), suggesting that this level of planning extends across at least a whole clause. In contrast, sound errors almost exclusively occurred within a single phrase, suggesting that advance phonological planning may extend only to the current phrase. This evidence suggests that higher levels of planning have a bigger scope and more elaborated frames than lower levels, but that there is at least some planning beyond the single word even at later planning stages (Damian & Dumay, 2007).

However, the picture is further complicated by recent evidence that supports a far narrower scope of sentence planning, consisting of a single phrase or even a single word (Allum & Wheeldon, 2007; F. Ferreira, 1991; Martin, Crowther, Knight, Tamborello II, & Yang, 2010; Martin & Freedman, 2001; Martin, Miller, & Vu, 2004; Smith & Wheeldon, 1999; Wheeldon, Ohlson, Ashby, & Gator, 2013). For example, F. Ferreira (1991) asked participants to produce sentences that consisted of two noun phrases of different complexity. Participants showed shorter latencies for speech onset when the initial noun phrase was less

complex (e.g., *The large and raging river...*), than when the sentences began with a more complex noun phrase (e.g., *The river that stopped flooding...*). The complexity of the final noun phrase did not influence speech onset, indicating that only the initial noun phrase was planned before speech onset. Several other studies are also consistent with the phrase-level view. Smith and Wheeldon (1999) let their participants describe the movement of objects on a computer screen. This required either the production of conjoined noun phrases at the start of the description, followed by a simple noun phrase (complex-simple condition, e.g., *The dog and the foot move above the kite*), or a simple phrase consisting of only one noun, followed by a complex noun phrase (simple-complex condition, e.g., *The dog moves above the foot and the kite*). They found longer speech onset times for the complex-simple condition when compared to the simple-complex condition, suggesting that only the first phrase of the sentence was fully planned prior to sentence initiation (see also Wheeldon, Smith, & Apperly, 2011, and Wheeldon et al., 2013, for further studies consistent with this view).

Further, evidence from eye movement studies has suggested an even narrower scope of planning (Griffin & Bock, 2000; Meyer, Sleiderink, & Levelt, 1998; Meyer & van der Meulen, 2000). For example, in an eye-tracking study, Griffin and Bock (2000) asked participants to describe a simple pictured event with either an active or a passive sentence. They found that the time it took from the last fixation of the element in the picture to the naming of it resembled the time that it normally takes to name a single object. Griffin and Bock therefore argued for an incremental, word-by-word scope of planning. However, they asked participants to describe the pictured events as quickly as possible, which might have forced them to plan incrementally. Therefore, the results may not reflect the planning scope under normal conditions.

The conflicting findings discussed above may reflect a scope of planning that varies depending upon the demands posed by different language tasks. In line with this, more recent



proposals postulate a flexible scope of advance planning whose unit size is dependent on the availability of cognitive resources. A study by Wagner, Jescheniak, and Schiefers (2010) used a picture-word interference design: Participants were asked to name an object or to describe a scene in a picture. An irrelevant auditory distractor word was presented just prior to the picture, and participants were simply told to ignore the distractor word. Such auditory distractors commonly elicit a *semantic interference effect*: Naming latencies for the picture are longer when it is preceded by a semantically related distractor than by an unrelated one (e.g., Levelt et al., 1999). In their study, Wagner and colleagues found an equally strong semantic interference effect for both nouns of a simple utterance (e.g., *The frog is next to the mug*), indicating that lexical access for both nouns had started before speech onset. When participants were required to produce a more complex sentence (e.g., *The blue frog is next to the blue mug*), a stronger interference effect was found for the first noun. This finding indicates that as processing demands increase, the scope of planning becomes reduced to a single noun phrase. The authors concluded that if the task does not require a lot of processing capacity, advance planning seems to involve later parts of the utterance as well, while when certain cognitive demands are high, planning units are smaller. A similar conclusion was reached by F. Ferreira and Swets (2002), using an arithmetic task. They presented participants with simple arithmetic problems (for example,  $24 + 23$ ). In one condition, participants had to produce the answer in a sentence (e.g., *The answer is forty-seven*). The authors found that response initiation times were influenced by the difficulty of the arithmetic task, indicating that participants planned the whole utterance in advance. In a second experiment, a production deadline was introduced: Participants had to give their answer in two to four seconds, represented by a timing bar that was displayed on the screen. In this experiment, participants initiated their answers significantly faster than in the first experiment, indicating that production had become more incremental. Taken together, these

findings suggest that instead of a fixed scope of planning, advance planning may depend on current processing demands, in the form of task complexity and time constraints present. This approach could account for the contradicting findings in the literature, since the complexity of tasks used in the above-mentioned experiments vary.

### **Comment**

To summarise, the sentence production models described in this review differ with regards to specific details about the production process. However, they agree on the main stages involved in the production process. Firstly, all models postulate a level that contains nonverbal representations of the message to be conveyed. Furthermore, all models make a distinction between those processes involved in lexical retrieval and those involved in syntactic encoding (that is, those that apply grammatical knowledge to determine the linear structure of the sentence to be produced). However, there are significant disagreements as to whether these stages are distinct (Chang et al, 2006; Garrett, 1975; Levelt, 1989; Levelt et al., 1999) or interactive (Bock, 1982; Dell, 1986; Stemmer, 1985). Other controversies include whether content and function words are processed separately (Garrett, 1975) or in the same way (Dell, 1986; Stemmer, 1985); whether activation of representations is excitatory only (Dell, 1986) or both excitatory and inhibitory (Chang et al., 2006; Stemmer, 1985). There are also unresolved questions about the scope of advance planning during sentence production, although most current research favours a phrasal scope of planning or at least one that is of flexible rather than of fixed size.

Now that we have reviewed processes involved in sentence production of healthy speakers, the following section will examine how these processes may go awry in nonfluent aphasic sentence production. This section will open with a discussion of the major (sub-) types of aphasia, and will then consider theories of how sentence production breaks down in aphasia, particularly nonfluent aphasia.

## **Aphasia**

Aphasic language disorders vary on a number of dimensions, but one of the major sources of variation concerns the fluency of an individual's spontaneous speech (Berndt, 2001; Goodglass, Kaplan, & Barresi, 2001). On the basis of this dimension, aphasic disorders have been divided into two broad subtypes, fluent and nonfluent. The fluent/nonfluent distinction is fundamental in many current aphasia classification systems, including the Boston Diagnostic Aphasia Examination (BDAE; Goodglass et al., 2001). The main discriminatory features of nonfluent/fluent aphasic speech as defined in the BDAE are: a) phrase length (generally less than four words per utterance in nonfluent aphasia/normal utterance length in fluent aphasia); b) melodic line (intonation may be low or absent in nonfluent aphasia/normal use of intonation in fluent aphasia); c) grammatical form (reduced variety of grammatical constructions and function words in nonfluent aphasia/normal range of grammatical structures in fluent aphasia). In the BDAE, each of these fluency factors is specified based on spontaneous speech samples elicited by free conversation and a complex picture description task ("the cookie theft"), and is rated on a scale ranging from one (highly impaired) to seven (normal speech characteristics). Furthermore, in nonfluent aphasia, the overall information content of speech is high with a good range of content words, whilst in fluent aphasia, the number of content words may be low and the speech marked with word finding difficulties (Goodglass et al., 2001). In addition, articulation in nonfluent aphasia may be effortful, whilst individuals with fluent aphasia show no articulatory difficulties (Goodglass et al., 2001).

With the BDAE classification system as our example, the following sections review the typical profiles associated with these two major aphasia subtypes in more detail, and the types of lesions usually associated with each. An illustration of the brain and its main gyri

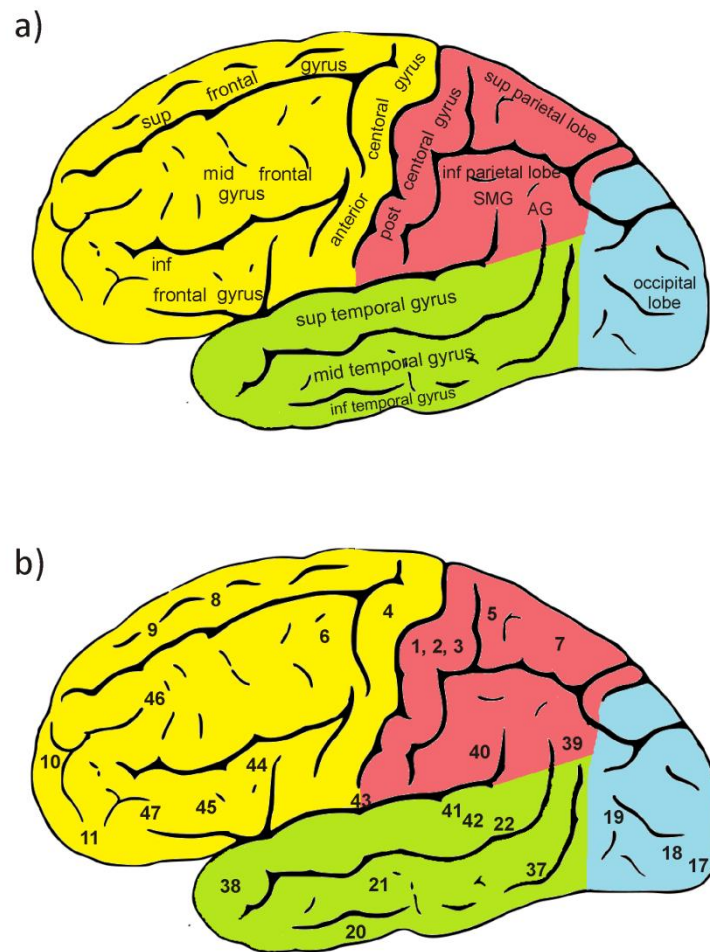
and Brodmann areas is given in Figure 2.1. Figure 2.2 displays overlay images of brain lesions in individuals with fluent and nonfluent aphasia.

### **Fluent Aphasia**

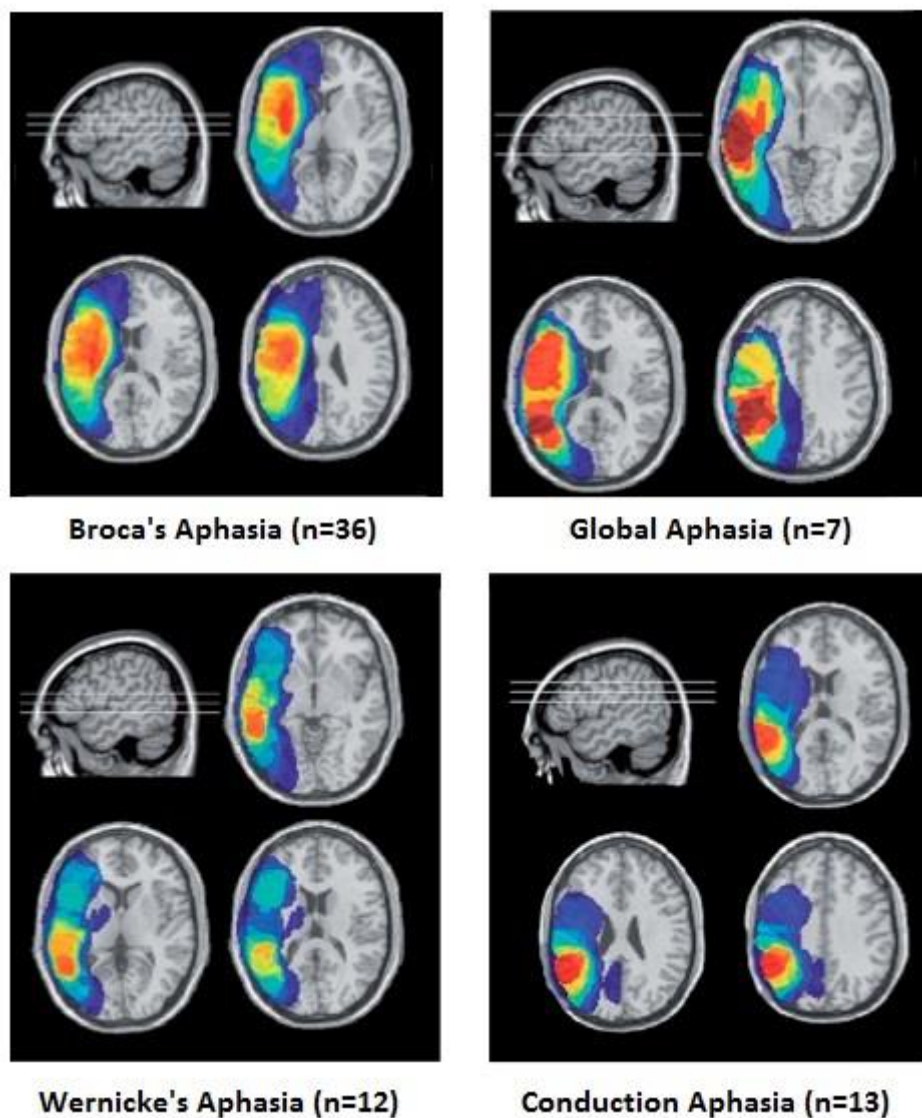
The term fluent aphasia is used in the BDAE and other similar classification systems (e.g., the Western Aphasia Battery; Kertesz, 1982) to describe individuals whose speech consists of long utterances produced at a relatively normal rate. Typically, articulation is unimpaired and there is little reduction of syntactic complexity, with a good range of function words and inflections (Berndt, 2001; Goodglass et al., 2001). However, individuals with fluent aphasia make frequent errors on content words. In some cases, this can result in “empty speech” that is low in informational output; in others, the speech might contain numerous outright errors, making it difficult or impossible to interpret. In standard aphasia classification systems such as the BDAE, fluent aphasia can further be divided into subtypes that differ on a more fine-grained scale: *Wernicke’s aphasia*, *conduction aphasia* and *anomic aphasia*. Wernicke’s aphasia is characterised by fluent speech output with a good range of function words and inflections. However, speech is characterised by many errors, mainly on content words in the form of whole word substitutions, phonemic paraphasias (sound errors) or neologisms (nonwords). Grammatical morphemes may be substituted as well, sometimes even to the extent that speech may become meaningless (so-called jargon). Repetition ability and auditory comprehension in individuals with Wernicke’s aphasia is generally poor. Conduction aphasia is also characterised by fluent speech, although it can contain neologisms and phonemic paraphasias. However, in contrast to individuals with Wernicke’s aphasia, individuals with conduction aphasia generally try to correct their errors and have preserved comprehension. They further show great difficulty with repetition, especially with multi-syllabic words. A third subtype, anomic aphasia, is characterised by good comprehension and repetition skills, and grammatically well-formed speech, although the speech frequently lacks

content words due to word finding difficulties. This may result in an extensive use of circumlocutions, the attempt to describe the intended word rather than to substitute it.

Fluent aphasia is usually associated with damage to posterior left hemisphere regions, involving temporal and temporoparietal areas (see Figure 2.1 and Figure 2.2). More specifically, deficits with semantic-lexical retrieval and comprehension in fluent aphasia have been linked to extensive lesions of the middle temporal gyrus and the underlying white matter, as well as to the superior temporal gyrus (Bates et al., 2003; Borovsky, Saygin, Bates, & Dronkers, 2007; Dronkers & Larsen, 2001). Specifically, the posterior portion of the superior temporal gyrus (Wernicke's area, Brodmann area (BA) 22) seems to play an important role in processing language and assigning meaning to sound, and thus is frequently, but not always, damaged in Wernicke's aphasia (Goodglass, 1993; Goodglass et al., 2001; Hickok & Poeppel, 2000). Repetition deficits, as observed in conduction aphasia, have been associated with damage to the arcuate fasciculus (a fibre pathway connecting Broca's and Wernicke's area) and to the posterior superior temporal gyrus (Dronkers & Larsen, 2001; Goodglass et al., 2001). Lesions encompassing the temporoparietal junction have been linked to the occurrence of phonemic speech errors, as frequently observed in conduction and Wernicke's aphasia (Hickok & Poeppel, 2007).



*Figure 2.1.* Left lateral view of the surface of the brain. Colours are indicative of the four brain lobes: Yellow = frontal lobe, red = parietal lobe, green = temporal lobe, blue = occipital lobe. Panel a) includes labels for the major gyri (sup = superior, mid = middle, inf = inferior, SMG = supramarginal gyrus, AG = angular gyrus). Panel b) includes Brodmann areas.



*Figure 2.2.* Lesion overlay maps of individuals with different subtypes of nonfluent (top row) and fluent (bottom row) chronic aphasia (adapted from Dronkers et al., 2009). The colour scheme indicates the degree of lesion overlap (dark blue = 0%, dark red = 100%).

### **Nonfluent Aphasia**

In nonfluent aphasia, by definition, spontaneous speech is slow and effortful, and many individuals present with distorted articulation. Utterances tend to be of reduced length and are separated by long and frequent pauses. Furthermore, spontaneous speech is often

syntactically impoverished; function words and grammatical morphemes may be frequently missing. However, key content words are often produced accurately. Based on the BDAE, nonfluent aphasia can be further distinguished into *Broca's aphasia*, *transcortical motor aphasia* and *global aphasia*. In Broca's aphasia, speech is sparse and fragmented. Utterances may consist of only one or two words at a time, and articulatory agility may be poor. Sentences, if they do occur, are usually limited to very basic subject-verb-object (SVO) structures and idiomatic phrases. Also, in some individuals, speech may be *agrammatic*, consisting entirely of content words, such as nouns and verbs, but lacking appropriate grammatical function words and inflections (Goodglass et al., 2001; Rochon, Saffran, Berndt, & Schwartz, 2000; Saffran, Berndt, & Schwartz, 1989). Despite their nonfluent spontaneous speech, individuals suffering from Broca's aphasia generally present with good reading and auditory comprehension, while writing abilities may be similar to those of speech production. Comprehension, at least at the single word level and in conversational context, is generally preserved, while repetition is generally poor. In transcortical motor aphasia, individuals usually perform comparatively well in repetition tasks, and individuals may articulate words with greater ease than in Broca's aphasia. Finally, in global aphasia, virtually all language abilities are affected: Spontaneous speech may be reduced to stereotyped utterances, repetition skills are poor and auditory comprehensions is generally impaired.

Nonfluent aphasia, particularly Broca's aphasia, is associated with damage to anterior left hemisphere regions, especially to the left inferior frontal gyrus (LIFG, including Broca's area (BA 44/45)) and/or the insula (Goodglass et al., 2001) (see Figure 2.1 and Figure 2.2). Both areas have long been known to be implicated in nonfluent aphasia, but have more recently been linked directly to a low speech rate and reduced grammatical complexity (Bates et al., 2003; Borovsky et al., 2007). The articulatory-motor deficits associated with Broca's aphasia have been linked to the left precentral gyrus of the insula: Damage to this area



discriminated between individuals displaying apraxia of speech and individuals who did not (Dronkers, 1996, see also Baldo, Wilkins, Ogar, Willock, & Dronkers, 2011). However, research so far has failed to localise an exclusive region of grammatical processing, indicating that it is more likely that a network of brain areas serves different aspects of syntactic processing than one single specialised area (Dronkers & Ludy, 1998; Kaan & Schwaab, 2002).

The following section will focus specifically on the sentence production difficulties in nonfluent aphasia, and will revise and critically discuss current theories put forward to explain nonfluent aphasic sentence production. Furthermore, the role of lexical content on sentence production in nonfluent aphasia will be examined, establishing the theoretical basis for the current research.

### **Theories of Nonfluent Aphasic Sentence Production**

We have seen above that one of the key characteristics of nonfluent aphasia is effortful, fragmented speech. However despite their difficulties producing sentences, many individuals with nonfluent aphasia perform comparatively well on single word production tasks (e.g., Freedman et al., 2004; Schwartz & Hodgson, 2002; Scott & Wilshire, 2010; Williams & Canter, 1982). For example, Williams and Canter (1982) showed that individuals with Broca's aphasia, as a group, were more accurate at producing the names of pictures when presented in isolation than when they were part of a larger scene, which had to be described using sentences. In a case study, Schwartz and Hodgson (2002) report a similar pattern for their participant MP, who was diagnosed with nonfluent aphasia. She performed relatively well in standard picture naming (73% correct), but struggled to produce these very same words in a sentence context, when asked to describe complex scenes which incorporated several pictures at once (23% correct).

Theoretical accounts of sentence production impairments in nonfluent aphasia have focussed primarily on individuals with classic "agrammatic" speech. Overall, these theories can be broadly grouped into structure-based and processing-based accounts. Structure-based accounts focus on the way in which sentence structure is represented in these individuals. They view agrammatic aphasia as a selective difficulty establishing the structural relations amongst elements in the sentence. Processing-based accounts focus more on the role of processing limitations and timing deficits in sentence production difficulties in agrammatic aphasia and sometimes nonfluent aphasia more generally. The following paragraph will discuss each group of accounts and their contribution to explain sentence production in nonfluent aphasia.

## Structure-based Accounts

One proposal is that individuals with agrammatism have difficulties mapping thematic roles (e.g., agent, patient) onto an appropriate syntactic frame (Maher, Chatterjee, Gonzalez Rothi, & Heilman, 1995; Saffran, Schwartz, & Marin, 1980; Schwartz, Saffran, & Marin, 1980). Evidence for this suggestion comes from findings that individuals with agrammatism have difficulties expressing syntactic relations based on word order. For example, Saffran, Schwartz and Marin (1980) used pictured event to elicit sentences containing subject and object nouns of varying animacy. The authors were able to show that five individuals with agrammatism were impaired when producing sentences that contained either two animate (e.g., *the boy carries the girl*) or inanimate nouns (e.g., *the pencil is in the sink*). They were more likely to confuse the agent and patient in such sentences compared to sentences that contained an animate agent and an inanimate patient (e.g., *the boy pulls the wagon*), presumably because they could not rely on a strategy to assign the animate item to the agent role and placing it in the subject-noun phrase. The authors proposed that the fundamental deficit in these cases is an inability to identify the thematic roles of the various participating entities, a problem that affects sentence comprehension as well as production (Schwartz et al., 1980). However, others have reported conflicting evidence from agrammatic individuals who were able to identify thematic roles (e.g., Caramazza & Hillis, 1989; Nespoulous et al., 1988).

Other researchers have proposed that impaired retrieval of closed-class vocabulary may play a key role in agrammatic speech, since grammatical morphemes – such as function words and inflections – are often systematically omitted (Bradley, Garrett, Kean, Kolk, & Zurif, 1979; Bradley, Garrett, & Zurif, 1980; Zurif, 1982). Alternatively, Kean (1977, 1979) proposed that the omission of function words and inflections in agrammatic speech might rather reflect a phonological deficit, based on the observation that individuals with

agrammatism primarily omit a certain class of grammatical morphemes, namely clitics, which do not influence stress.

Other theories attribute agrammatic speech patterns to a difficulty constructing an appropriate syntactic frame for a sentence – one which specifies its grammatical constituents and their interrelationships, and also specifies their sequential order and any necessary grammatical function words or inflections (e.g., Caramazza & Hillis, 1989; Goodglass et al., 1994; see Bastiaanse & Jonkers, 2012, for a recent review). For example, Thompson and colleagues have proposed that agrammatic individuals with nonfluent aphasia have difficulty accessing information about verb argument structure (e.g., Lee & Thompson, 2004; Thompson, 2003; Thompson, Lange, Schneider, & Shapiro, 1997). If this information is unavailable, an appropriate syntactic frame for the sentence cannot be constructed. Support for this proposal comes from the finding that individuals with Broca's aphasia, considered as a broad group, are generally less accurate at producing verbs than nouns (e.g., Breedin, Saffran, & Schwartz, 1998; Chen & Bates, 1998; Miceli, Silveri, Villa, & Caramazza, 1984; Zingeser & Berndt, 1988, 1990), and are particularly poor at producing verbs that carry multiple arguments (e.g., Dragoy & Bastiaanse, 2010; Kim & Thompson, 2000; Thompson & Choy, 2009). However, others report that some individuals with fluent aphasia also show impairments of verb production, and some studies have failed to find dissociations between verb and noun production in individuals with agrammatism (Berndt & Haendiges, 2000; Berndt, Haendiges, & Wozniak, 1997; Berndt, Mitchum, Haendiges, & Sandson, 1997). It also remains unclear why some agrammatic individuals have more difficulties producing semantically light verbs (that carry only a few semantic features, e.g., to *have*) than semantically heavy verbs (that carry many semantic features, e.g., to *fly*) (Bencini & Roland, 1996).

Other proposals characterise agrammatic speech in terms of the specific syntactic operations that are compromised. For example, Bastiaanse and van Zonneveld (2005) make a primary distinction between sentences that follow the “base” word order for the language (e.g., SVO sentences in English, such as *The dog chased the cat*), and “derived” sentences, where the base order is violated (e.g., Wh-questions, such as *Who is the dog chasing?*). They predict that derived sentences involve additional syntactic computations, and are therefore more likely to be failed by agrammatics, a prediction that has been supported in a number of studies (e.g., Abuom & Bastiaanse, 2013; Bastiaanse & Thompson, 2003; Burchert, Meißner, & De Bleser, 2008).

Alternatively, Friedmann and Grodzinsky postulate a hierarchy of syntactic relations that need to be specified in a sentence (Friedmann, 2006; Friedmann & Grodzinsky, 1997). Some of these operate over low-level constituents, such as individual noun phrases (e.g., noun number agreement: *those books*), and others operate over larger units such as entire clauses (e.g., tense agreement: *Yesterday, the boy wrote*), or even across multiple clauses (e.g., *The children whose work\_ the teacher\_ read were proud*). A severe agrammatic impairment could compromise the specification of syntactic relations at all levels, even the lowest ones. A milder impairment, on the other hand, might impact on higher levels but leave the lowest levels intact - for example, agreement within a noun phrase. In other words, the authors propose that damage at one level in the hierarchy leads to an inability to access all higher levels, so they become unavailable in the production process, while lower levels are unimpaired. This prediction has been supported in some studies (e.g., Friedmann, 2006; Friedmann & Grodzinsky, 1997). However, it has to be noted that the data taken to support this prediction is contradictory: While it was true that agrammatic individuals showed poorer performance for tense inflections than agreement inflections (confirming the prediction that damage to a higher level left access to lower levels intact), they were still able to produce

tense inflections correctly from time to time, as indicated by error rates of less than 100%. In fact, according to Friedmann and Grodzinsky's theory, damage to one level leads to the full unavailability of the corresponding syntactic relation, and individuals would be expected to never produce such relations correctly. Additionally, several studies failed to confirm Friedmann and Grodzinsky's predictions (e.g., Burchert, Swoboda-Moll, & De Bleser, 2005; Lee, Milman, & Thompson, 2008).

Recently, Bastiaanse and colleagues have suggested that syntactic/thematic relations that span more than one sentence may be particularly severely compromised in agrammatism. Examples are pronominal references (e.g., *The man was fat. He walked with a stick*), and references to tense/time frames (e.g., *I saw the doctor yesterday. He told me to rest up*). These types of relations may even operate across different speakers within the same discourse (e.g., *Q: Which of these cakes are gluten-free? A: The pink ones are*; Bastiaanse, 2008; Dragoy, Stowe, Bos, & Bastiaanse, 2012).

### **Challenges to Structure-based Accounts**

According to these various structure-based accounts, nonfluent – or more specifically, agrammatic – speech is seen as a selective difficulty establishing the structural relations amongst elements in the sentence. The processes involved in generating the actual lexical *content*, on the other hand, are considered to be relatively preserved. Such a proposal accounts nicely for the disproportionate difficulty with the production of well-formed sentences when compared to that of single words in individuals with nonfluent aphasia. Further, all these accounts have in common that they focus on issues of representation rather than process: That is, they aim to describe the types of relational representations that cannot be generated, while putting to one side the exact nature of the cognitive processes that may be engaged to support these operations. Consequently, they are not well-equipped to deal with the kind of variability in performance that can occur in real life speech production situations,

in which time constraints and working memory demands vary widely. In other words, they face two major challenges.

First, there is considerable variability amongst individuals with nonfluent aphasia in the features they exhibit. For example, some nonfluent individuals omit function words consistently, while others succeed in producing them in some contexts but not others (Saffran et al., 1989). Also, some individuals systematically omit all grammatical morphemes, but others omit mainly free-standing ones (e.g., *to*, *for*), and yet others omit mainly bound ones (e.g., *jumps*; Miceli, Silveri, Romani, & Caramazza, 1989; Rochon et al., 2000). Further, while many nonfluent individuals are less accurate at producing verbs than nouns, others show the opposite pattern (Berndt & Haendiges, 2000; Berndt, Haendiges, et al., 1997). Any complete account of the sentence production difficulty in nonfluent aphasia needs to be able to explain this inter-individual variability.

The second issue that arises is that often there is also substantial variability *within* individuals (e.g., Bastiaanse, 1995; Beeke, Wilkinson, & Maxim, 2007; Hofstede & Kolk, 1994; Kolk, 2006, 2007; Kolk & Van Grunsven, 1985; Nespoulous, 2000; Sahraoui & Nespoulous, 2012). Individuals with nonfluent aphasia may find it difficult to produce certain syntactic structures most of the time, but seem to be able to produce them in some situations. For example, Sahraoui and Nespoulous (2012) found that agrammatic participants were more likely to produce complete, grammatically well-formed sentences in a picture description task than they were during an open-ended interview (for additional supporting evidence, see Beeke, Maxim, & Wilkinson, 2008; Beeke, Wilkinson, & Maxim, 2003; Heeschen & Schegloff, 1999; Salis & Edwards, 2004). In a case study, Beeke and colleagues (2007) report the same performance pattern for their agrammatic participant Roy. Remarkably, they also found variability of performance even within each task: For example, in a picture description

task, the majority of Roy's attempts to produce subject-verb-object sentences were unsuccessful. Nevertheless, he succeeded on three out of 20 such attempts.

Another issue that challenges current structure-based accounts is that non-grammatical factors also seem to play an important role in nonfluent aphasic language production. For example, poor performance in multi-word utterances has been reported for tasks that do not require any syntactic operations, like the production of two or three target words in one utterance, for example "*red ear horse*" (Freedman et al., 2004; Schwartz & Hodgson, 2002). Furthermore, the nonfluent case of BM reported by Wilshire and McCarthy (2002) showed impairments in another non-grammatical task: In addition to his preserved single word production and multiword naming deficit, he made more errors when required to successively name semantically related items than unrelated ones (a method known as blocked cycling). Other studies have shown the same effects (e.g., McCarthy & Kartsounis, 2000; Schnur, Schwartz, Brecher, & Hodgson, 2006).

### **Processing-based Accounts**

Several processing-based accounts have been put forward to explain some of these challenges. For example, the framework of Kolk and colleagues emphasises the importance of timing in sentence production. These researchers have suggested that timing may play a central role in the sentence production difficulties in nonfluent aphasia. They suggest that the fundamental problem may be a lack of synchronicity between the activation of lexical elements and their associated syntactic slots in the syntactic frame (Hartsuiker & Kolk, 1998; Kolk, 1995, 2006; Kolk & Van Grunsven, 1985). If the activation for a given lexical element does not peak during the period in which its associated syntactic slot is available, the dominant syntactic constituent may be incomplete, or entirely omitted (Kolk, 1995). This lack of synchronicity may be due to a number of factors, including slow retrieval of syntactic frame information, slow retrieval of lexical content items, or fast decay of either of these



types of representations (Kolk & Van Grunsven, 1985). As mentioned above, structure-based approaches to sentence production in nonfluent aphasia are not well-equipped to account for variability in participants' performance. However, Kolk's approach offers an elegant way to do so. Within this framework, between-participant variability may be attributed to different decay and/or retrieval functions in different individuals (Kolk, 2006). Within-participant variability may be attributed to differences in the timing demands imposed by different conversational contexts, and to the specific structure and/or content of the sentences being planned. Importantly, this type of explanation is not limited to classic agrammatic speech – in fact, Kolk and colleagues suggest that at least some of the features of agrammatism are not themselves indications of an underlying grammatical impairment, but rather reflect the use of compensatory strategies (called adaptation) designed to maximise the amount of information the speaker can produce in a limited time frame (see also de Roo, Kolk, & Hofstede, 2003; Ruiter, Kolk, & Rietveld, 2010; Salis & Edwards, 2004). Kolk argues that the timing problem is not restricted to nonfluent aphasia, but also underlies fluent aphasia. He believes that the differing speech patterns of individuals with fluent aphasia only reflect a missing adaptation strategy, possibly due to a lack of self-monitoring (Kolk, 2006).

Even though the timing account can elegantly explain between- and within-participant variability in nonfluent aphasic speech production, it does not account for some of the more specific dissociations that have been observed between individuals with fluent and nonfluent aphasia. For example, while many nonfluent individuals are more accurate at producing nouns than verbs, many individuals with fluent aphasia show the converse pattern (e.g., Breen & Warrington, 1994; Chen & Bates, 1998; Miceli et al., 1984; Zingeser & Berndt, 1990; Zingeser & Berndt, 1988). If the same deficit underlies both types of aphasias, even with Kolk's broadly applicable adaptation theory it is difficult to explain why there is a difference in the production of words that belong to different grammatical categories.

Furthermore, it is not clear exactly which processes cause the timing deficit, as Kolk (2006) just assumes that overall processing capacity is reduced in general after brain damage. Moreover, like structure-based accounts, Kolk's timing account has no explicit explanation for difficulties in multiword utterances that have no grammatical structure and therefore do not rely on lexical and syntactic synchronicity.

One account that addresses some of these challenges is that of Randi C. Martin and her colleagues. They suggest that some of the sentence production difficulties observed in nonfluent aphasia may be attributable to deficits in a system they call *semantic short-term memory* (e.g., Martin & Freedman, 2001). Martin and colleagues postulate that semantic short-term memory is a limited-capacity buffer that operates to maintain ordered lexical elements in a temporarily heightened state of activation, for example, during the planning of a larger utterance. Semantic short-term memory is itself contentless, consisting merely of a series of placeholders (or "slots"), which can be tied to particular lexical elements as needed. During production, representations of the lexical elements to be produced are connected to their respective placeholders in the buffer system. Interactive flow of activation between these units and the lexical elements ensures that information in the buffer is maintained over the course of production. Once it is time to produce a particular word, an attention shift is directed to the relevant "slot" unit which boosts its activation, and consequently that of its associated lexical element (Freedman et al., 2004). Abnormally rapid decay between the units in the buffer system and the lexical elements could result in a failure to maintain the elements of the sentence long enough to structure them into well-formed sentences.

Such a deficient capacity for retention of lexical elements could be involved in the sentence production difficulties in individuals with nonfluent aphasia when compared to that of single words, leading to difficulties when more than one word needs to be kept in the buffer (Martin & Freedman, 2001; Martin & He, 2004; Martin, Lesch, & Bartha, 1999). In

line with this, Martin and her colleagues were able to show that two patients with nonfluent aphasia had difficulty producing adjective-noun phrases (e.g., *short straight hair*), even though they were able to produce the words in isolation as well as in utterances that could be planned incrementally, only containing phrases that comprised one single content word (e.g., *the hair is short*) (Martin & Freedman, 2001; Martin & He, 2004). Interestingly, in a patient with fluent aphasia, accuracy in these production tasks was within the normal range. These findings indicate that if maintenance demands in the semantic short-term memory are high, for example, when planning phrases that contain multiple content words, production accuracy drops only for the nonfluent aphasic individuals.

Like that of Kolk (1995), Martin and colleagues' (2001) framework offers an elegant way to deal with variability in language performance of individuals with nonfluent aphasia. Although not explicitly stated by the authors, varying degrees of decay rates between the units in the buffer system and the lexical elements could potentially account for differences between individuals, while varying demands on retention capacity in different tasks and in different situations may explain within-participant variability. Furthermore, the reduced retention capacity approach accounts nicely for difficulties with grammatical and non-grammatical multi-word utterances in nonfluent aphasia. A deficient capacity for retention of lexical items could be involved in the non-grammatical multiword naming deficit observed in individuals with nonfluent aphasia, making it difficult to keep more than one word in the buffer. However, like that of Kolk, Martin and colleagues' theory does not address specific dissociations that have been observed in nonfluent aphasic speech, such as their superior production of nouns when compared to verbs, and that of content words compared to function words.

One model of aphasic sentence production that may speak to these issues is the *Division of Labour* (DoL) model (Gordon & Dell, 2003). The DoL model is a connectionist

model which embodies the main idea of Chang and colleagues' Dual-path model (2006), that lexical elements receive activation from separate syntactic and semantic components. More specifically, in the DoL model, the activation of different lexical elements, including determiners, nouns and verbs, depends on the combined input from two types of *nodes*, which are capable of sharing the labour involved in lexical element retrieval: *conceptual-semantic* and *syntactic-sequential* nodes. Both types of input nodes receive activation from a higher-level message that remains completely activated throughout the whole production process. The conceptual-semantic input nodes activate the lexical element that represents the correct meaning of a word in an utterance, whereas the syntactic-sequential input nodes ensure that the right word in the utterance is produced at the right time, activating lexical elements that are congruent with the structure that needs to be produced. Gordon and Dell (2003) used a simple learning algorithm to train the connection strength between these two sets of input nodes and various types of lexical items. After training, nouns and semantically rich verbs, (e.g., to *fly*) were primarily activated by semantic-conceptual input nodes, whereas determiners and semantically impoverished verbs (e.g., to *have*) were primarily activated by syntactic-structural input nodes.

By lesioning the connection strength from syntactic input nodes to lexical elements, Gordon and Dell (2003) were able to simulate nonfluent agrammatic speech, represented by the omission of determiners or semantically impoverished verbs, with a good range of noun production and semantically rich verbs<sup>4</sup>. This pattern occurred because function words and semantically impoverished verbs received only weak, if any, activation from the lesioned syntactic input nodes. Moreover, there was also enhanced lexical competition, because all the

---

<sup>4</sup> Conversely, lesioning the semantic input nodes to lexical elements led to fluent anommic speech patterns, represented by a good production of determiners and semantically impoverished verbs but a lack of nouns and semantically rich verbs, in agreement with previous findings (e.g., Breedin et al., 1998).

planned sentence elements remained concurrently activated from the message, and no structural representation was in place to enhance activation of the right word for the right position.

In the DoL model, variability in performance amongst individuals with nonfluent aphasia can be accounted for by varying the degree to which the relevant connection strengths are weakened. Moreover, within-individual variability may be attributed to the context in which lexical elements need to be produced, and the kind of lexical competition that arises from it. Although it has not been explicitly stated by the authors, the DoL may also offer an explanation for the non-grammatical factors that seem to play an important role in nonfluent aphasic language production: If the input of syntactic-sequential nodes is necessary to establish a serial order in an utterance by activating the right word at the right time, a reduction of syntactic connection strength may account for the observed difficulties with the production of multiple nouns in the same utterance (e.g. Freedman et al., 2004; Schwartz & Hodgson, 2002). However, the DoL model simulated speech under very simplified conditions, by only using subject-verb sentences, and it does not make any specific assumptions about the production of bound-morphemes like verb inflections. It therefore remains unclear how more complex structures in the model would be generated, and how specifically these structures would be affected by lesions to any of the input nodes.

### **The Role of Lexical Content in Nonfluent Aphasic Sentence Production**

We have seen that, while structure-based accounts do not offer an explanation for variability in language performance between and within individuals, they are effective at explaining many of the structural features of nonfluent aphasic speech. Conversely, processing-based theories do not address structural aspects of speech in any detail, but they elegantly account for the different sources of variability. However, a crucial component of *all* processing-oriented frameworks is that sentence production success in real time depends not

just on the utterance's syntactic structure, but also on its *lexical content* – even when lexical content retrieval is not itself compromised in that individual. Even if the process of frame construction takes place relatively independently of lexical content retrieval, the lexical content items still need to be available at the appropriate time for integration into the sentence plan. As mentioned before, several models of normal sentence planning go further, suggesting that the processes of structure generation and lexical content retrieval can directly influence one another. In several such models, lexical content elements are capable of activating representations of their respective syntactic constituents (Bock, 1982) and vice versa (Chang et al., 2006; Stemberger, 1985). This means that if a lexical item is particularly highly activated, it can have a direct impact on the syntactic structure of the resultant utterance. If so, then an individual with a difficulty constructing a syntactic frame in a top-down manner may be particularly reliant on current activated lexical elements to help support frame construction.

There are at least three existing studies that may speak to the issue of how lexical availability impacts nonfluent aphasic sentence production. The first is a study by Faroqi-Shah and Thompson (2003) that examined the effect of providing written lexical prompts to participants just prior to producing a sentence. A group of seven individuals with Broca's aphasia was asked to describe pictured actions using a single sentence (e.g., *A girl is tickling a boy*). Arrows and other types of prompts were used to indicate the type of sentence to be produced (active or passive). The Broca's group failed to demonstrate reliably improved accuracy when lexical information about the key nouns and the verb was provided in uninflected written form in the picture. Unfortunately, the authors did not report individual data. Also, the authors scored the participants' best effort instead of their first response, a procedure that may not reflect the challenges of online sentence production where time restrictions may not allow for multiple attempts at the sentence.

Two further studies provide some informal evidence that improving lexical availability may improve nonfluent sentence production accuracy more generally. One examined the effectiveness of a therapy software package called *SentenceShaper*®, which allows participants to record, replay and even reorder words and phrases they want to produce (Linebarger et al., 2007; Linebarger, Schwartz, Romania, Kohn, & Stephens, 2000). Individuals with nonfluent aphasia have shown improvements when using this treatment device, both in the length and the grammatical well-formedness of their sentences. For example, nonfluent participant DD's description of a scene changed from "*Ohh! A fish! Ah, water and....uh mmmm and attendant, here, and bumped his head. Oh boy, oh my hand, my hand, my hand*" to "*The boy and the fishmonger is taking the fish. The boy hit his hand*" (Linebarger et al., 2000, p. 422). It is likely that one of the primary benefits of *SentenceShaper*® is that it enables the participant to make relevant lexical content items more readily available during sentence construction. Another recent therapy study trained a single individual with nonfluent aphasia to produce a specially selected set of nouns and verbs (Raymer & Kohen, 2006). After training, this person was more accurate at producing sentences featuring both the trained nouns and verbs, than those featuring untrained items. Importantly, the training not only affected accuracy of the trained words themselves, but it also improved overall syntactic well-formedness, including the use of correct inflections and determiners.

### **Semantic Interference Effects**

More evidence for the role of lexical content in sentence production in nonfluent aphasia comes from recent research suggesting that individuals with nonfluent aphasia may be particularly susceptible to semantic interference effects when they are planning to produce two words within the same utterance (e.g., Freedman et al., 2004; Scott & Wilshire, 2010). In many contemporary models of normal speech production (e.g., Dell, 1986; Levelt, 1989;

Levelt et al., 1999), the activation of a single lexical item leads to partial activation of semantically related words. Within this framework, if two meaning-related words need to be produced in one sentence, both words will activate one another, making selection of each more difficult. For example, in the sentence “the cat is chasing the dog”, the selection of *cat* will be slower and/or more prone to error because *dog* is concurrently highly activated at the same time, and therefore competes strongly for selection. Recent research has shown that normal speakers are indeed sensitive to the semantic similarity of the words they produced in an utterance. For example, Smith and Wheeldon (2004) asked participants to describe the movement of objects and written words presented on a computer screen. The nouns describing the objects and the written words were either produced in one phrase (e.g., *The saw and the axe move down*) or in separate phrases (e.g., *The saw moves towards the axe*), and the nouns were either semantically related (e.g., *saw and axe*) or unrelated (e.g., *saw and cat*). Participants exhibited slower onset latencies for sentences containing semantically related nouns than for sentences containing unrelated nouns, independent of whether they were in the same or in different phrases (for similar findings using different experimental paradigms, see, for example, Kroll & Stewart, 1994, and Sass et al., 2010).

In interactive sentence production models, competition between lexical elements planned for the same utterance (called *syntagmatic interference*) is resolved by the structural representations, which selectively activate their respective lexical content elements (e.g., Stemmer, 1985). In other words, each slot in the syntactic plan is capable of activating its corresponding lexical element, thereby minimising lexical competition between sentence elements. In a recent computer simulation study, Dell and his colleagues explicitly specified a mechanism responsible to resolve competition between lexical elements during speech production: the *syntactic traffic cop* (Dell, Oppenheim, & Kittredge, 2008). The syntactic traffic cop is developed through learning and constitutes both inhibitory and excitatory links



from syntactic input nodes to lexical output units (in the same fashion as in the DoL model of Gordon & Dell, 2003). In an intact model, the syntactic nodes are guided by the syntactic traffic cop, which ensures that the right word is produced at the right time, as the nodes virtually “move” through the utterance and activate lexical elements from the syntactic category that is required for successful structure building at that particular moment (e.g., the head of a subject noun phrase would require a lexical element from the noun category). Importantly, the syntactic traffic cop can help to resolve competition between lexical elements from *different* syntactic categories (e.g., nouns, verbs) during sentence production. However, because this model was implemented in the context of simple subject-verb sentences, Dell and colleagues do not specifically address how it would deal with competition between words from the *same* syntactic category (e.g., subject noun, object noun). A model that elegantly manages syntagmatic competition between words from the *same* class is the DP model of Chang and colleagues (2006).

As mentioned earlier, the model incorporates the main principle of the DoL model (and the syntactic traffic cop), that activation of lexical elements depends on the combined input from distinct meaning and syntactic systems. Further, the DP model was trained to produce a variety of different sentence structures, for example, SVO sentences. During the production of such an SVO sentence, the sequencing system in the model selects the most highly activated thematic role from the meaning system (e.g., AGENT) whose syntactic category (e.g., noun) is consistent with its predictions about the next appropriate syntactic category. In this example, competition between the subject and the object noun is managed effectively because the object noun is linked to a different thematic role (PATIENT) than the subject noun (AGENT), and this thematic role will only get selected from the sequencing system after production of the AGENT and ACTION role has been completed. In this way, the sequencing system of the DP model successfully manages competition between words

from the same word class that form part of the same utterance. Consequently, if the process of structure generation helps to limit syntagmatic interference by ordering lexical elements, then individuals with structure generation difficulties may be particularly susceptible to syntagmatic interference when they are planning a sentence, because lexical elements are freer to compete with one another for selection to a particular position.

Although to our knowledge no studies have explicitly manipulated lexical competition in aphasic *sentence* production, several studies have investigated its effect in *noun phrase* production. First, Freedman and colleagues (2004) showed two individuals with nonfluent aphasia pairs of pictures, and asked them to produce their names in a conjoined noun phrase. These two individuals not only made more errors in this task than in simple single picture naming, but they were also slower to name noun phrases containing semantically related nouns (e.g., *nose and mouth*) than they were to name phrases containing unrelated nouns (e.g., *nose and hat*). Controls showed a similar, though weaker effect in the same direction. Interestingly, this effect was not observed in a case with fluent aphasia, which in fact showed a trend in the opposite direction.

Another study examined semantic interference effects in a single individual. Scott and Wilshire (2010) used a similar task to that of Freedman and colleagues (2004). They asked a participant with nonfluent aphasia (JHM) to name two consecutively presented pictures in one single utterance (e.g., *goat and pig*), keeping their original order. The pictures were visible for 2000ms, then disappeared, encouraging advance planning of the utterance and therefore increasing lexical competition. JHM was significantly slower to name semantically related picture pairs (e.g., *goat and pig*) than unrelated ones (e.g., *goat and ball*). Although controls showed a trend in the same direction, the semantic interference effect was reliably larger for JHM.

### **Introduction to the Current Study**

In sum, there is considerable evidence that lexical content variables are of vital importance in sentence production in nonfluent aphasia, be they in the form of availability of lexical elements in a sentence or the semantic features that they share with each other. However, no study so far has systematically manipulated different aspects of lexical content in sentence production tasks in individuals with nonfluent aphasia. The current research aims to address this gap in the literature. As mentioned above, lexical aspects of sentence production in nonfluent aphasia have received relatively little attention. The theoretical focus has been squarely on syntactic and structural aspects of speech planning, and consequently, most studies have explored the effects of syntactic structure on production while minimising the influence of lexical content – for example, comparing different types of structures while keeping lexical content constant and/or minimising the impact of lexical variables by using very easily accessible words. In the current study, we aim to do the exact opposite and investigate how manipulations of lexical content influence sentence production accuracy. In this approach, we keep the syntactic structure of our target sentences constant, and manipulate the properties of the content words that must be included in the sentence. If we take the view that content and structure generation processes are interactive, then sentence production accuracy in nonfluent aphasia may be particularly powerfully influenced by the availability of lexical content. Furthermore, if the structure generation processes are not operating effectively, this may lead to enhanced lexical competition effects in individuals with nonfluent aphasia.

In contrast to many previous studies on sentence production in aphasia, we classified our participants as nonfluent or fluent based on standard measures of fluency derived from the BDAE, which include utterance length, melodic line and grammatical form (see above). We did not select our nonfluent participants on the basis of the presence or absence of

agrammatism, because we did not wish to make specific assumptions about the relationship between agrammatic features of speech and the underlying sentence production difficulties. Individuals with fluent aphasia were also included. If, as proposed, the sensitivity to lexical content in sentence production is specific to individuals with nonfluent aphasia due to selective underlying impairments, individuals with fluent aphasia should show no (or at least different) effects of lexical content in our sentence production tasks. Finally, healthy control participants were included to serve as a baseline of how normal sentence production is influenced by the same experimental manipulations. These participants were selected based on their age (60 years and older) to roughly account for the age range of the aphasic participants.

The data is analysed in two ways. First, we adopt a case series approach, analysing both accuracy and latency data separately for each aphasic individual. Second, if the patterns of individual performance justify it, we will compare performance across our broader participant groups (the nonfluent, fluent, and control groups, respectively). The combination of a case series approach and a group study approach provides two benefits. First, it has been argued that when using a group approach, it is impossible to know whether all individuals within that group show the exact same performance patterns in the task administered (Caramazza & McCloskey, 1988; Nickels, 2002). As mentioned above, although many symptoms are shared between individuals from the same aphasia group, there may be a considerable source of variability. Consequently, group averages may not reflect the performance pattern of each individual, whose patterns of performance even go in the opposite direction in some cases (e.g., Schnur et al., 2006). However, the statistical power of a single case analysis is inevitably limited, especially when the magnitude of the investigated effect is small (Crawford, 2004; Crawford, Garthwaite, & Gray, 2003). Since statistical power is dependent on the size of the sample (Tabachnick & Fidell, 2012), analysing

individuals at the group level can be an effective strategy to improve statistical power, given that they show a similar performance pattern.

In the presented research, we used picture events to elicit simple SVO sentences (e.g., *The cat is chasing the dog*). The simple canonical SVO structure was chosen to avoid floor effects: All of the participants could produce simple subject-verb-object sequences under at least some conditions. Further, the key nouns used in the sentences were well within the vocabulary of all the participants. Indeed, as we document below, although these participants made very occasional errors on the key nouns during single naming, no single lexical item was *consistently* failed by any of the participants. In order to mimic naturally occurring speech as closely as possible, participants were asked to describe pictured events without any further instructions or external prompts such as arrows. Also, the pictures were removed after the commencement of the sentence response, to prevent participants from falling back on a simple labelling strategy. Finally, we ensured a high response agreement to the pictures used in the sentence production tasks by extensively piloting the materials on normal, healthy controls before commencement.

In Experiment 1, we systematically varied the lexical availability of sentence elements by manipulating the frequency of the noun items that needed to be included in the sentence. By varying their frequency, it is possible to manipulate the speed with which the lexical elements can be retrieved, making them more or less readily available. Many studies have shown that normal speakers are faster to produce very common words than less frequent ones (e.g., Barry, Morrison, & Ellis, 1997; Kittredge, Dell, Verkuilen, & Schwartz, 2008; Oldfield & Wingfield, 1964). Based on the premise that individuals with nonfluent aphasia may rely abnormally heavily on the currently activated lexical elements to drive the sentence structure planning process, we hypothesised that the frequency of the lexical elements, particularly those appearing early in the sentence, would influence the production accuracy of the other

sentence elements. In contrast, individuals with fluent aphasia were not expected to show these effects.

In Experiment 2, we varied the semantic relationship between the subject and object nouns. If individuals with nonfluent aphasia cannot make effective use of a structural representation to maintain lexical elements in their correct order during sentence planning, those lexical elements may be freer to compete with one another for selection to a particular position. If so, then sentences containing semantically related subject and object nouns would be more prone to error overall, and may also be associated with longer sentence initiation times. This effect would again not be expected to occur in individuals with fluent aphasia, who might actually show a facilitative semantic priming effect when lexical content elements share semantic features, for both accuracy and response time.

An additional aim of the research was to examine whether certain specific patterns of performance in Experiment 1 and 2 were associated with damage to particular cortical regions. To address this question, we obtained structural MRI scans for each aphasic participant, which were first co-registered to a standard brain template and then analysed using simple nonparametric statistics. Although few studies have examined the neural correlates of sentence production deficits, Thothathiri, Schwartz, and Thompson-Schill (2010) recently reported that Brodmann's area 44/6, at the junction between Broca's area and the left premotor cortex, was consistently damaged in two individuals who showed particular difficulty in the sequencing of lexical elements in phrase production. Such a lexical sequencing problem could in principle arise from the same types of cognitive impairment we hypothesised here. Further, studies focussing on single word production have associated the LIFG (including Broca's area) with resolving competition between simultaneously activated representations (e.g., Schnur, Lee, Coslett, Schwartz, & Thompson-Schill, 2005; Schnur et al., 2009; Thompson-Schill et al., 1998). For example, using a cyclic naming task in which a

small set of pictures is repeatedly presented for naming, Schnur and colleagues (2009) found that damage to the LIFG (including Broca's area) was reliably associated with a pattern of increasing errors across successive presentations, particularly when the items in the set were all semantically related. The authors concluded that the LIFG may be important for resolving competition in situations where competition is high, and the naming of repeated successive semantically related items could be just such a situation<sup>5</sup>. Nonetheless, it is not clear whether the processes and brain areas involved in single word production transfer to the ones involved in sentence production, which poses many different demands on the language system.

---

<sup>5</sup> For a proposal that Broca's area manages competition even between non-linguistic representations and serves a more general cognitive control function, see Novick, Kan, Trueswell, & Thompson-Schill, 2009, and Novick, Trueswell, & Thompson-Schill, 2010.

### **Experiment 1: Effects of Lexical Frequency on Sentence Production**

The purpose of Experiment 1 was to investigate the effect of lexical availability on sentence production in nonfluent aphasic, fluent aphasic and control participants. One simple way to manipulate the availability of lexical items is by varying their frequency of occurrence. Frequency is a variable known to influence the time course of lexical retrieval in normal individuals: In single word production tasks, normal speakers produce common words more quickly than less common ones (e.g., Barry et al., 1997; Kittredge et al., 2008; Oldfield & Wingfield, 1964), and this effect appears to be robust across a range of tasks. Also, in connected speech, normal speakers initiate the production of conjoined noun phrases more quickly when the first noun is of high than of low frequency (Griffin, 2001). Several theories of language production propose higher resting levels for frequently used words than for less common ones (e.g., Dell, 1986; Stemmer, 1985), so that frequently occurring words can reach high levels of activation more rapidly than less frequent ones, speeding up the retrieval process. Furthermore, recent research suggests that more common words may be processed more rapidly than less common ones at a number of stages of processing during word production (Kittredge et al., 2008).

In the current experiment, participants described simple pictured events in a single SVO sentence (e.g., *The cat is chasing the dog*). The frequency of the key nouns in the sentence was manipulated. The picture stimulus disappeared as soon as the participants commenced their response. This was done in order to prevent a strategy of simply labelling the pictured items, and also to promote advance planning of the entire sentence, thereby more closely replicating the demands of everyday speech.

If, as hypothesised, individuals with nonfluent aphasia rely more strongly on the availability of the content elements to drive sentence production than normal controls or fluent aphasic individuals, they should be very sensitive to a frequency manipulation of those



content elements. Further, there may be a particular advantage if a sentence begins with a high frequency word, because frame construction may be faster and downstream sentence elements will have less time to decay before they can be produced. Therefore, lexical frequency of the subject noun would be expected to have a particularly strong effect on the chances of producing the remainder of the sentence correctly. Furthermore, this effect should be specific to individuals with nonfluent aphasia, when compared to individuals with fluent aphasia and nonaphasic controls.

Prior to the sentence production task, a pilot study was conducted to ensure sufficient response agreement to the pictures later used. Therefore, data for two experiments is reported. The first describes the pilot study procedure and the construction of the stimulus pictures, while the second reports the aphasic and control production data from the sentence production task proper.

## **Response Agreement Pilot Study**

### **Method**

The School of Psychology Ethics Committee of Victoria University of Wellington gave ethical approval for this study in 2011.

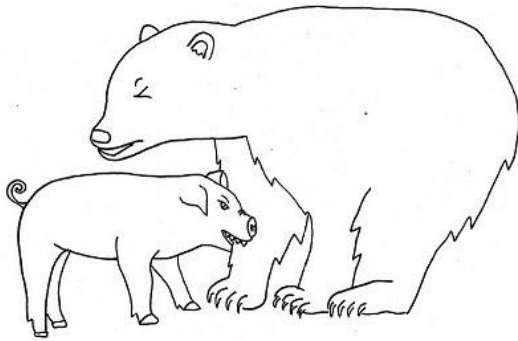
#### **Participants**

Forty-four psychology students enrolled in a first-year course at Victoria University of Wellington took part in this study for course credit. Nine were male (20%). All participants were native speakers of English, ranging in age from 17 to 42 years ( $M=20.06$ ,  $SD=6.12$ ). Three participants were excluded: One did not follow the instructions, and the remaining two had an overall accuracy score below 1.5 standard deviations of the mean accuracy score calculated across all participants. This left a total of 41 participants for the analysis, with the same age range as before (17 to 42 years,  $M=19.63$ ,  $SD=5.47$ ). Eight participants (20%) were male.

#### **Materials**

A set of 120 line drawings was drawn specifically for the study. Every picture displayed a scene that could be described in an SVO sentence (e.g., *The pig is biting the bear*). Examples are shown in Figure 5.1. The noun elements depicted in the scenes were drawn from a pool of 18 monosyllabic nouns denoting people or animals (see Appendix A for a full list). All items were early-acquired: that is, under the age of seven according to Carroll and White (1973) and/or Morrison, Chappell, and Ellis (1997). The 18 nouns included nine low frequency items (frequencies of 30 or less occurrences per million in the Subtitle Analysis Project; Brysbaert & New, 2009), and nine high frequency items (frequencies in excess of 30 occurrences per million; see also Schnur et al., 2006), for a similar division).

High frequency subject - high frequency object

*'The pig is biting the bear'*

High frequency subject - low frequency object

*'The nurse is shooting the clown'*

Low frequency subject - high frequency object

*'The clown is shooting the nurse'*

Low frequency subject - low frequency object

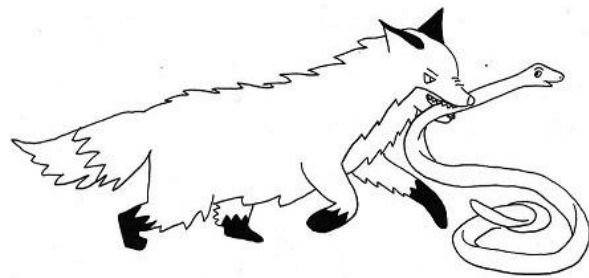
*'The fox is biting the snake'*

Figure 5.1. Examples of the pictured scenes from different experimental conditions.

Each target noun element appeared twice, in two different picture scenes, once as the sentence subject (e.g., *The pig is biting the bear*), and once as the object (e.g., *The bear is biting the pig*). The noun elements in a given sentence were always from the same category: Animals only appeared with animals, and people only with other people. In all pictures, the subject was displayed on the left side of the picture to encourage the production of active rather than passive sentences (e.g., Hartsuiker & Kolk, 1998). As shown in Figure 5.1, there were four conditions representing the four different possible combinations of high and low

frequency nouns, and 30 sentence exemplars for each condition. Fifteen different transitive verbs were depicted in the scenes and their occurrence was balanced, each verb appearing twice in each of the four frequency conditions.

## **Procedure**

Participants were tested in groups with a maximum number of 12 people per testing room, widely spaced so as to prevent them from seeing one another's responses. Immediately prior to the testing session, participants were given an information sheet which contained background information about the study, gave instructions about the upcoming task and asked them for their age and gender. After reading the information sheet, participants were given the instructions a second time, this time orally from the experimenter. They were asked to write down one sentence that would best describe what is happening in each single picture and that if they did not know the answer they should write down what first came into their mind. It was emphasised that they should not go back and edit their answers, nor confer with anyone during the test. After going through an example picture with the experimenter, the participants were each handed out a booklet containing prints of the 120 target pictures (two pictures per A4 page), and they were allowed to start writing the sentences in a space below each picture once they had received the booklet.

**Scoring procedure.** To be considered as correct, a response was allowed to vary from the target sentence only in the following ways: a) It could include additional words that did not change the grammatical properties of the sentence (e.g., *The clown is shooting the nurse in the chest*); b) it could contain an appropriate particle (e.g., *lifting up* instead of *lifting*); or c) it could contain an alternative verb that was meaning-appropriate and did not change the grammatical properties of the sentence (e.g., *eating* instead of *biting*). The criterion for selection of a picture for the main experiment was that it elicited the target construction (as defined above) in at least 80% of participants. Since each target item appeared twice, in two

different picture scenes – once in subject and once in object position – both pictures needed to reach this criterion to be selected.

### **Results**

Agreement for each individual picture ranged from 70 to 100 per cent ( $M=91.95$ ,  $SD=11.39$ ). Seventy-one pictures reached name agreement of 80%, 48 of which scored 90% or higher. However, five of these pictures had to be excluded because the reversed partner picture did not reach the required amount of agreement. This left a total of 66 pictures: 16 pictures containing two items with high frequency, 18 pictures containing two items with low frequency, 16 pictures with a high frequency subject and a low frequency object, and the remaining 16 pictures with a low frequency subject and a high frequency object. All of those pictures were selected to serve as the material for the sentence production task proper (see Appendix A for a full list), which will be reported in the following section.

## Sentence Production Task

The School of Psychology Ethics Committee of Victoria University of Wellington gave ethical approval for this experiment in 2011.

## Method

### Participants

**Control participants.** Eight control participants were recruited from a database of individuals from the Wellington region who had indicated their willingness to be contacted about participating in research. These eight participants were all aged over 60 (range: 63 to 84 years;  $M=67.50$ ,  $SD=6.82$ ), six had a university education and three were male (38%). All were native speakers of English, and none reported any significant neurological history. All control participants gave written consent, and were reimbursed for their participation in each session (NZD20 if they travelled to university for participation, and NZD10 if they were visited at home). None of the participants was excluded from the study.

**Participants with aphasia.** Nine participants with aphasia due to stroke were either recruited from a register of volunteers willing to be contacted about research or from regional stroke groups and via speech language therapists. All met the following inclusion criteria: a) Their stroke occurred more than 12 months prior to the commencement of the study; b) their aphasia was classified as mild to moderate-severe according to the BDAE severity scale (Goodglass et al., 2001); and c) they correctly named at least 50% of the pictures in the Boston Naming test (long form; Goodglass et al., 2001). Further, all had normal or corrected-to-normal vision, and English was their native language. According to the BDAE and based on the measures outlined above, five of the participants were classified as having nonfluent aphasia (all had a diagnosis of Broca's aphasia), and the remaining four as having fluent aphasia (two were diagnosed with anomic aphasia, one with conduction aphasia and one with

Wernicke's aphasia). Diagnostic and background information for the aphasic participants is presented in Table 5.1.

Each of the nine aphasic individuals had completed extensive language assessment within the three years prior to the experiment, the results of which were used to compile individual language profiles for each aphasic participant to complement the BDAE measures (test records are presented in Table 5.2). Wherever individuals scored below two standard deviations of the mean for normal controls, performance was interpreted as below the normal range. A detailed description of the tests administered is included in the language profile of the first aphasic case. Examples of spontaneous speech for each aphasic individual are given in Table 5.3. Graphical illustrations of each participant's brain lesion (where available) are displayed in Figure 5.2 (showing nonfluent participants), and Figure 5.3 (showing fluent participants). Figure 5.4 illustrates the lesions of all aphasic participants on the rendered surface of the brain.

In the following, individual language profiles for each aphasic participant based on their test records, as well as descriptions of participants' lesion sites (where available), are presented.

***Participants with nonfluent aphasia.***

*BY.* BY suffered a subarachnoid haemorrhage as a consequence of a motorbike accident he had 37 years prior to testing and further suffered a cerebrovascular accident (CVA) eight years ago. His BDAE diagnosis was mild Broca's aphasia. BY's speech was halting and slow, and utterances were separated by pauses. As illustrated in Table 5.3, his sentences were of simple syntactic structure and generally well-formed, containing both content and function words. A quantitative assessment of BY's spontaneous speech was conducted using the Quantitative Production Analysis (QPA; Saffran, Berndt, & Schwartz, 1989). The QPA is based on the participant's narration of a well-known story (usually the

Table 5.1

*Background, medical and diagnostic information for each aphasic participant.*

Participant	Age	Gender	Education	Years post CVA	Lesion Site/aetiology	BDAE measures					
						Diagnosis	Severity level	Phrase length	Melodic line	Grammatical form	Fluency classification
BY	59	male	HS	37/8	Subarachnoid haemorrhage, subsequently operated upon, large lesion extending from anterior horn of L lateral ventricle to L parietal lobe	Broca's	3	5	4	4	nonfluent
DA	71	male	HS	11	Medical notes not available	Broca's	2	4	3	5	nonfluent
JG	73	female	HS	6	Isch. CVA, L MCA region	Broca's	2	4	4	3	nonfluent
JHM	52	female	UD	10	Isch. CVA, extensive L MCA	Broca's	2	3	2	3	nonfluent
RP	66	male	HS	9	Unspecified CVA, extensive L frontal and parietal and also R medial frontal infarct	Broca's	2	3	4	4	nonfluent
NP	73	male	HS	13	Isch. CVA, several foci in L occipital and temporal lobes	Anomia	2	7	7	7	fluent
STR	81	female	HS	13	Isch. CVA, possibly multiple, infarcts in R occipital and L parietal lobe	Anomia	3	7	7	7	fluent
SW	82	female	HS	4	Haem. CVA, L posterior temporal lobe	Wernicke's	1.5	7	6	6	fluent
WL	64	male	UD	2	Isch. CVA, L parietal and L posterior temporal lobe	Conduction	3	7	7	7	fluent

*Note.* HS = high school; UD = university degree; Isch. = ischaemic; Haem. = haemorrhagic; CVA = cerebrovascular accident; MCA = middle cerebral artery; L = left; R = right.



Table 5.2

*Performance of participants with aphasia on language and related cognitive tests*

	Nonfluent aphasia					Fluent aphasia				Control
	BY	DA	JG	JHM	RP	NP	STR	SW	WL	2 SD cut-off
<b>QPA measures</b>										
Words per minute	<b>67</b>	<b>19</b>	<b>69</b>	<b>48</b>	<b>22</b>	103	153	96	126	86.82 <sup>a</sup>
Proportion closed-class items	0.64	<b>0.40</b>	0.47	<b>0.37</b>	<b>0.31</b>	0.58	0.61	0.54	0.66	0.46 <sup>a</sup>
<b>Word production</b>										
Boston Naming Test (N=60)	<b>41</b>	53	<b>46</b>	<b>46</b>	<b>38</b>	<b>32</b>	53	<b>37</b>	<b>40</b>	48.1 <sup>b</sup>
Druks and Masterton										
Object and Action naming										
Nouns (N=50)	<b>45</b>	50	<b>43</b>	48	46	48	47	<b>42</b>	46	45.16 <sup>c</sup>
Verbs (N=50)	<b>43</b>	45	<b>38</b>	46	<b>41</b>	44	48	<b>30</b>	<b>43</b>	43.20 <sup>c</sup>
<b>Connected Speech</b>										
Sentence Production Test										
Sentence Score	85.7	82.8	70.8	81.6	64.7	80.7	89.2	59.6	74.0	na
Sentence Residual	-2.7	6.3	-12.8	-2.0	-9.5	-2.9	12.2	-14.7	40.1	na
Closed-class Residual	-5.1	-11.3	-14.5	-4.6	-10.6	2.7	13.3	-27.8	15.3	na
Noun Score	92.5	96.7	76.7	85.0	74.2	75.0	79.2	84.2	77.5	na
Verb Score	80.0	70	60.0	75.0	35.0	90.0	65.0	17.5	40.0	na
<b>Comprehension</b>										
PCB Sentence Comprehension subtest										
Lexical distracters (N=30)	30	<b>29</b>	30	<b>29</b>	<b>29</b>	30	<b>29</b>	<b>27</b>	30	29.1 <sup>d</sup>
Reverse role distracters (N=30)	29	<b>21</b>	<b>25</b>	26	<b>23</b>	30	<b>21</b>	<b>12</b>	<b>18</b>	25.8 <sup>d</sup>
Peabody Picture Vocabulary Test (Form IIIB) Standard Score	90	99	90	91	84	81	96	< <b>40</b>	102	70

Table 5.2 continued.

	Nonfluent aphasia					Fluent aphasia				Control
	BY	DA	JG	JHM	RP	NP	STR	SW	WL	2 SD cut-off
<b>Comprehension</b>										
Pyramids and Palm Trees Three-picture subtest (N=52)	52	48	48	51	50	48	52	<b>43</b>	50	46.5 <sup>e</sup>
<b>Auditory Language Processing</b>										
PALPA same-different discrimination using word pairs	70	<b>54</b>	71	71	67	<b>34*</b>	66	<b>49</b>	70	61.97 <sup>f</sup>
<b>Short-term memory</b>										
PALPA Auditory digit span	<b>4</b>	<b>4</b>	6	<b>3</b>	<b>2</b>	5	<b>4</b>	<b>2</b>	<b>3</b>	na
<b>Verbal Fluency</b>										
Letter fluency: Total of F, A, S	<b>7</b>	16	<b>5</b>	<b>6</b>	<b>3</b>	27	22	10	15	9.08 <sup>g</sup>
Category fluency (1min):										
Animals	11	15	11	9	<b>5</b>	14	17	<b>4</b>	9	8.2 <sup>h</sup>
Fruits	6	5	9	8	3	9	7	5	7	na

Note. na=not available

Scores in bold font indicate performance below the normal range.

PCB: Philadelphia Comprehension Battery (Saffran, Schwartz, Linebarger, Martin, & Bochetto, 1988)

QPA: Quantitative Production Analysis (Saffran et al., 1989)

PALPA: Psycholinguistic Assessment of Language Processing in Aphasia (Kay, Lesser, & Coltheart, 1992)

<sup>a</sup> Control data from Rochon et al., 2000

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

<sup>c</sup> Control data are based on Druks and Masterson (2000), who report normative data for a more extensive version (N=100)

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

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

<sup>f</sup> Control data from Kay et al., (1992)

<sup>g</sup> Control data from Donovan, Siegert, McDowall, and Abernethy (1999)

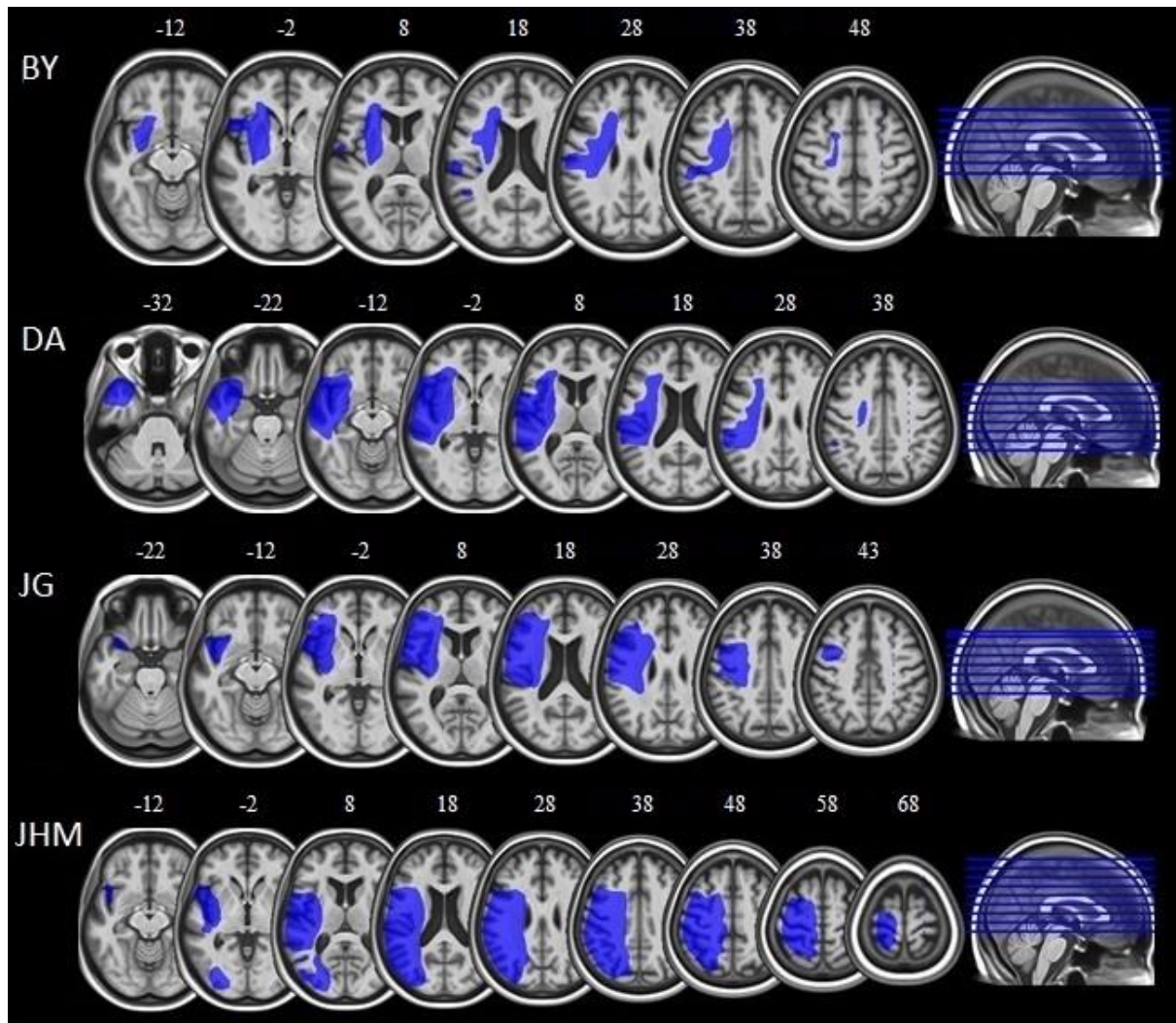
<sup>h</sup> Control data from Loonstra, Tarlow, and Sellers (2001)

\* score not representative because participant NP had difficulties with his hearing aid during testing

Table 5.3

*Spontaneous speech sample of aphasic participants describing the Cookie Theft Scene in the BDAE (Goodglass et al., 2001)*

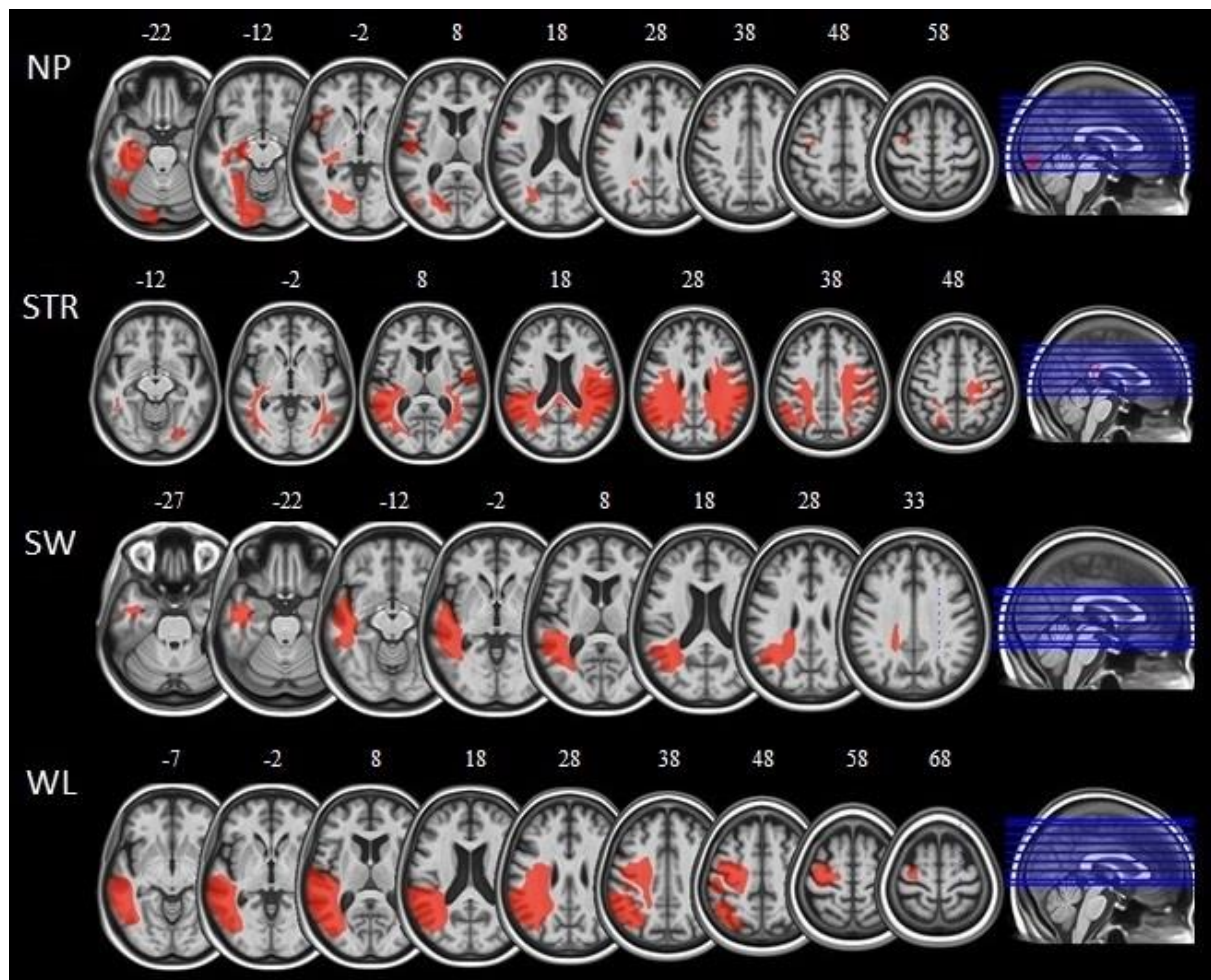
<b>Nonfluent</b>	
BY	Um...  oh god, I dunno...  it's it's waiting for it to come...  uh...  and that ...  (sigh)... god, I can't say it  it's overflowing... [long pause]  and she's washing dishes and...  with her back to the...  child  getting the..  cookies...  [truncated]
DA	Um...  the the the the /tʌd/  the tub was on yeah...  um yeah...  oh...  behind /bi:hænd/ behind...  her children...  um...  um... [long pause]  were reaching... [long pause]  reach reaching...  into a cookie jar...  in the...  in the... um...  the cupboard  [truncated]
JG	Um, a child...  and, cookie- jars...  um...  washing /wɪt/ um the tap...  /wɒʃssss/...  I dunno  yeah (laughs)...  /n-/  spilling /ɪ- ʌnn/...  /lɪ/ floor...  um...  garden  [truncated]
JHM	OK...  the the woman ah ah ah  dreaming ah dreaming uh uh ah um...  she uh d:ry a plate  um...um...  the water...  over:flow uh...  to the floor um...[long pause]  um  the- the boy...  stealing a cookie...  and um...  uh the boy is...  giving a... girl /g-/ giving  a cookie a /g-, gɔ:/ [truncated]
RP	Cookie jar...  boy...  girl...  /t/ tip over...  washing the dishes... [long pause]  pill...  floors is spill...  and gardening...  and /tu:/...  uh...  uh...  kitchen...  and...  cups...  plates...  cupboards...  uh curtains...  trees... [long pause]  curtains again...  shrubs...  uh lawns
<b>Fluent</b>	
NP	Uh that's  that's his mother...  the water's...  overflowing shhhh so I don't know what the hell she's doing (laughs)  that's a- an apron...  over a frock and shoes and she's got a dinner plate and tea towel...  um...  out the window- that's window's uh curtains  [truncated]
STR	Um...  there is a um  there is a mother and two children and they're in the kitchen and they are doing the washing doing washing up but the  unfortunately the tap has over- /θ/  the sink has overflowed the /w/ the...  /ovəflɒ/ with the water so there's  there's  there's water on the floor  [truncated]
SW	An /æprɒn...  æpərən/...  and he's got a little uh...  car, jar...  I mean cookie...  I think she's going to eat something or drink something...  he going fly up...  you see...  I can't see the word! You got it...  why is it like that...  him going on the...  skull...  Why is that? ...  [truncated]
WL	Ah, the man /n/ or the the boy  the boy is um...  trying to get, uh  the cookie jar, uh  to give uh  the girl a cookie and the the the the boy /f/-fell down the  the uh  the uh um  The uh...  the stool  and uh the wife  the mother  the mother- was washing the dishes and uh /tu:/  /w/ ah overflowed the sink  [truncated]



*Figure 5.2.* Lesion maps for participants with nonfluent aphasia showing axial slices of the brain on a standard template (Rorden, Bonilha, Fridriksson, Bender, & Karnath, 2012). Slices were selected according to representative display of individual lesions (corresponding Montreal Neurological Institute (MNI) Z coordinates are reported above each slice). Further details of imaging methods and image preparation are described in Chapter 7.

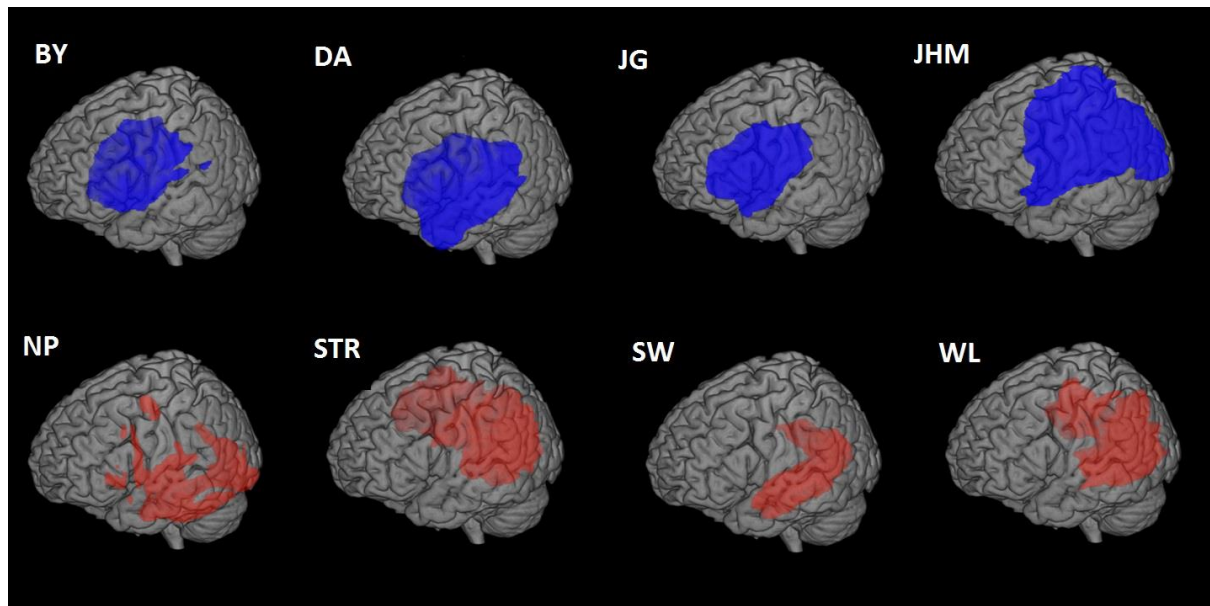
story of “Cinderella”) and gives information about various measures of speech production, including the number of words per minute and the proportion of closed-class words produced. In this test, BY’s rate of words per minute fell within the higher range of scores for a group of individuals with Broca’s aphasia reported by Rochon et al. (2000), and his proportion of closed-class words measure was even within the fluent range.

Single word production was assessed using the Boston Naming Test (BNT; Goodglass et al., 2001) and the Object and Action Naming Test Battery (Druks & Masterson, 2000). Both tests involve confrontation naming of pictures: The BNT includes pictures of nouns that increase in difficulty, whilst the Object and Action Naming Test Battery contains two separate sets of pictures showing nouns and verbs, making the test useful to dissociate between noun and verb naming ability. BY's performance in both naming tests was only very mildly impaired, and he showed no difference between the production of nouns and verbs.



*Figure 5.3.* Lesion maps for participants with fluent aphasia showing axial slices of the brain on a standard template (Rorden, et al., 2012). Slices were selected according to representative display of individual lesions (corresponding MNI Z coordinates are reported above each slice). Further details of imaging methods and image preparation are described in Chapter 7.





*Figure 5.4.* Lesions maps for individuals with aphasia showing the left lateral side of the brain, using a standard template (Colin27; Holmes et al., 1998). The top row shows lesions of individuals with nonfluent aphasia (blue), the bottom row shows lesions of individuals with fluent aphasia (red). For further details of imaging methods and image preparation, see Chapter 7.

The Controlled Oral Word Association Task (COWAT; Loonstra et al., 2001) was used to measure verbal fluency. The test, as administered in the current study, requires the production of as many words as possible within one minute each either starting with a given letter (F, A, and S) or belonging to a certain category (fruits and animals). BY's letter fluency score was below the normal range; however, his performance in the category fluency task was unimpaired.

BY's sentence production ability was examined using the Sentence Production Test for Aphasia (SPT; Lukkien, Burmester, & Wilshire, 2012), which contains 30 pictured events that need to be described in one sentence. The SPT offers a range of measures taken from the sentence analysis, some of which are similar to measures used in the QPA. However, unlike the QPA, the SPT is designed to measure accuracy of sentence production, and thus is a

useful test to complement fluency measures gained from open-ended speech. The specific SPT measures chosen for the current study included the sentence score, which credits for the correct production of the target nouns, verbs, as well as for their appropriate inclusion into large phrasal constituents. BY's sentence score was the second highest in the group. Further, the sentence residual score indicates how well participants produce sentences in comparison to their single picture naming (as assessed on a single noun test presenting pictures of the nouns used in the SPT sentences). BY's negative score on this measure showed that he performed worse on the sentence production test than would be expected given his performance at naming the sentence nouns in isolation. In the same way, the SPT also offers a residual score for closed-class elements, which quantifies production of closed-class elements relative to open-class elements. BY's negative closed-class residual score indicated that he produced fewer closed-class elements than would be predicted based on his production of open-class elements. Finally, the SPT includes separate verb and noun scores that give the percentage correct with which these elements were produced within the target sentences. BY's scores in these measures indicated that he was slightly more accurate at producing nouns than verbs.

Auditory sentence comprehension was assessed using the Philadelphia Comprehension Battery (PCB; Saffran et al., 1988). In the sentence comprehension subtest, participants are read a sentence (e.g., *The boy washes the girl*) and asked to select the corresponding picture in the presence of either a reverse role distracter (e.g., *The girl washes the boy*; indicating the ability to comprehend syntactic relations) or a lexical distracter (e.g., *The boy washes the dog*; indicating lexical comprehension). Single word comprehension and semantic processing were also examined using picture-matching tasks. The Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 1997) requires participants to point to one out of four pictures that corresponds to a spoken target word, whilst the Pyramids and Palm Trees

Test (PPT; Howard & Patterson, 1992) is completely nonverbal and only requires participants to match a target picture with a semantically related alternative out of two presented pictures. Phonological input processing was examined using PALPA Test 2 (Kay et al., 1992), which requires the participant to indicate whether phonemically similar word pairs are the same or different. BY showed no evidence of difficulty in auditory comprehension, semantic processing, or phonological processing difficulties: he scored in the normal range for tasks of single word and sentence comprehension, and also for auditory minimal pair discrimination. Verbal short-term memory span was assessed by asking him to repeat auditorily presented strings of digits that increased in length (PALPA Test 13; Kay et al., 1992). BY was able to remember a four-digit span, indicating that his verbal short-term memory span was slightly below the normal range.

BY's structural MRI scan revealed damage exclusively to the left hemisphere, including the inferior parietal cortex, the parietal operculum and the left amygdala. Anteriorly, the lesion encompassed the primary motor cortex (BA 4), and small portions of BA 44 and the insula. More posteriorly, the lesion encompassed the primary somatosensory cortex (BA 2 and BA 3) and a small portion of the superior temporal gyrus, along its superior aspect (see Figure 5.2 and Figure 5.4 for graphical depictions of BY's lesion). However, the remainder of the temporal lobe was spared.

*DA.* DA suffered a CVA 11 years ago. His BDAE diagnosis was moderate Broca's aphasia. DA's spontaneous speech was severely nonfluent and characterised by long pauses and articulation difficulties. He produced mainly content and some function words, and was able to produce grammatically well-formed sentences when there was no time constraint. When analysed using QPA measures (Saffran et al., 1989), DA's speech rate was within the lower range of Rochon et al.'s (2000) group of individuals with Broca's aphasia, while his proportion of function words fell within one standard deviation of the group mean. DA's



single picture naming abilities were well preserved: He scored in the normal range or near ceiling in all tests, and showed no dissociation between his ability to produce nouns and verbs. His performance on verbal fluency measures was also unimpaired.

On the SPT, DA sentence production performance exceeded expectations based on his single picture naming, indicating that he was able to produce grammatically well-formed sentences when given enough time and the chance to correct himself. Further, DA's noun score was near ceiling, and he also produced the majority of the verbs in the SPT correctly. However, his production of closed-class elements in the sentences was lower than expected based on his open-class word production. DA further presented with intact single word comprehension and semantic processing abilities. In contrast, comprehension of sentences was impaired in the presence of a reverse role distracter. Furthermore, DA showed poor auditory-phonological input processing. On the digit span task, he presented with a slightly reduced short-term memory span of four digits.

DA's structural MRI scan revealed a lesion extending from BA 44 anteriorly to the left inferior parietal cortex and the parietal operculum, and encompassing the primary somatosensory cortex (BA 3) and the left insula. Inferiorly, the lesion encompassed almost the entire left superior and middle temporal gyri (see Figure 5.2 and Figure 5.4 for graphical depictions of the lesion).

*JG.* JG suffered a CVA approximately six years ago. Her BDAE diagnosis was moderate Broca's aphasia. JG's spontaneous speech was nonfluent and fragmented, and characterised by articulatory difficulties. As shown in Table 5.3, her utterances were limited to only one or two words, and there were few function words. JG's speech rate as assessed in the QPA (Saffran et al., 1989) was at the higher end of the range for a group of individuals with Broca's aphasia reported by Rochon et al. (2000), while the proportion of closed-class items fell within one standard deviation of the group mean. JG's picture naming scores fell

consistently below the normal range in the word production tests administered. Her performance in letter fluency was also well outside the normal range. On the SPT, JG performed well below the expectations from single picture naming, showing that it was more difficult for her to produce well-formed sentences than producing the sentence nouns in isolation. Further, she produced much fewer closed-class elements than expected based on her open-class word production, and was slightly more accurate in producing nouns than verbs in the sentences.

In contrast to production measures, JG presented with normal performance in semantic processing, single word comprehension and phonological input processing tasks. However, on sentence level comprehension measures, she performed slightly below the normal range when sentences were presented with a reverse role distracter, indicating a mild syntactic comprehension deficit. Her short-term memory span of six fell within the lower normal range.

JG's structural MRI scan revealed damage exclusively to the left hemisphere: Anteriorly, the lesion encompassed a large portion of the left inferior frontal gyrus (including BA 44 and BA 45), the premotor and primary motor cortex (BA 6 and BA 4), and the insula. There was also some damage to the primary somatosensory cortex (BA 1, BA 2, and BA 3), the inferior parietal cortex and the parietal operculum. Posteriorly, the lesions further extended to the anterior and middle superior temporal gyri, but sparing the posterior superior frontal gyrus (BA 22) and the middle and inferior temporal gyrus (see Figure 5.2 and Figure 5.4 for graphical depictions of JG's lesion).

*JHM.* JHM suffered a CVA 10 years ago at the age of 42. Her BDAE diagnosis was moderate Broca's aphasia. JHM's spontaneous speech was severely nonfluent and fragmented, characterised by mild articulation difficulties. As can be seen in Table 5.3, utterances consisted mostly of one or two words, and there were few function words and

inflections. When assessed using the QPA (Saffran et al., 1989), JHM's speech rate and the proportion of closed-class items fell within one standard deviation of the mean reported for individuals with Broca's aphasia (Rochon et al., 2000). Her single word production was mildly impaired: She scored slightly below the normal range in the Boston Naming Test, but just within the normal range on the Object and Action Naming Test Battery, and she showed no clear dissociation between noun and verb production. JHM's performance on letter fluency tasks was poor, while she scored within the normal range when producing words from the same semantic category instead.

On the SPT, JHM performed more poorly than expected based on her single picture naming, and her closed-class element production was also lower than expected from open-class word production. She further produced slightly more nouns correctly than verbs. JHM's comprehension of both single words and sentences was well preserved, and she presented with relatively intact auditory phonological processing. In contrast, her verbal short-term memory span was significantly impaired; she was only able to remember three digits.

JHM's structural MRI scan revealed a large left hemisphere lesion, which encompassed BA 44, the premotor and primary motor cortex (BA 6 and BA 4) and the left insula anteriorly, extending posteriorly to encompass a large portion of the left parietal lobe, including the primary somatosensory cortex (BA 1, BA 2, and BA 3). The lesion extended inferiorly to encompass the posterior two-thirds of the superior temporal gyrus, including BA 22 (see Figure 5.2 and Figure 5.4 for graphical depictions of the lesion).

*RP.* RP suffered a CVA nine years ago at the age of 57. His BDAE diagnosis was moderate Broca's aphasia. As illustrated in Table 5.3, RP's speech was severely nonfluent and fragmented, with long pauses and articulatory difficulties. His utterances consisted of mainly one or two words, and there were very few function words and inflections. When assessed using the QPA (Saffran et al., 1989), RP's speech rate and his proportion of closed-

class items fell within one standard deviation of the mean of Rochon et al.'s (2000) group of individuals with Broca's aphasia. On measures of single word production, RP's performance was impaired: He scored below the normal range in the Boston Naming Test. However, in the Object and Action Naming Test Battery, he only scored below the normal range in action naming, indicating a better production of nouns relative to verbs. Further, RP's performance in letter and category fluency tasks was also impaired; he scored well below the normal range on both tasks.

On the SPT, RP performed very poorly. His sentence score was the second lowest in the group. His overall performance fell short of expectations based on his single picture naming, and his closed-class element production was also much lower than expected given his open-class word production. Furthermore, his verb production was much more impaired than his noun production, in agreement with his poor performance in confrontation action naming. RP showed normal semantic processing and single word comprehension and normal auditory-phonological discrimination. In contrast, he showed impaired syntactic comprehension of sentences in the presence of a reverse role distracter. RP's short-term memory span was severely impaired; he was only able to remember a two digit-span.

No MRI scan was available for RP, so detailed lesion locations cannot be described. However, according to his medical notes, he had extensive damage to left frontal and parietal regions as well as to right medial frontal areas.

### ***Participants with fluent aphasia.***

*NP.* NP suffered a CVA 13 years ago. His BDAE diagnosis was moderate anomic aphasia. Table 5.3 shows that NP's speech was fluent and well-formed, containing elaborated sentences, and was characterised by occasional pauses due to word finding difficulties. On the QPA (Saffran et al., 1989), his speech rate and his proportion of closed-class items fell within the normal range (Rochon et al., 2000). On tests assessing single word production, NP

showed a moderate naming deficit: he scored below the normal range in the Boston Naming Test. However, his performance was normal for both nouns and verbs in the Object and Action Naming Test Battery. Further, NP scored in the normal range in both letter and category fluency tasks.

On the SPT, NP performed slightly worse than would be expected from his picture naming. In contrast, his closed-class element production fell in excess of the expectation based on his open-class word production. Further, NP was the only one in the group who showed a higher verb than noun production accuracy. NP displayed intact comprehension at both the single word and sentence level. In contrast, his phonological input processing performance was severely impaired. However, at the time of testing he had just received a new hearing aid and had problems with adjusting it, so this score is unlikely to be an accurate representation of his abilities. When he performed the same test in a previous session some 10 years earlier, his performance was significantly more accurate (65/75; Cloutman, 2005, unpublished). NP's verbal short-term memory span of five was within the low normal range.

NP's structural MRI scan revealed several foci of damage in the left hemisphere. The largest lesion encompassed posterior portions of the superior, middle and inferior temporal gyri, a small portion of the inferior parietal lobe, and the lateral occipital lobe, including the primary visual cortex (BA 17), the secondary visual cortex (BA 18), and the associative visual cortex (BA 19). There were also several smaller lesion foci: One was located in the region of BA 44; another in premotor and primary motor cortices (BA 6 and BA 4); and a third affected small portions of the insula (see Figure 5.3 and Figure 5.4 for graphical depictions of the lesions).

*STR.* STR suffered multiple CVAs approximately 13 years ago; her best fit diagnosis on the BDAE was mild anomic aphasia at the time of testing. However, initially after her stroke, she had been diagnosed with conduction aphasia. As shown in Table 5.3, STR's

speech was fluent with grammatically correct and elaborated sentences; however, occasionally she displayed word finding difficulties and produced phonemic and formal paraphasias. On the QPA (Saffran et al., 1989), both her speech rate and proportion of closed-class items fell within the normal range (Rochon et al., 2000). STR's naming abilities were relatively well preserved, as was her performance on letter and category fluency tasks.

On the SPT, her sentence score was the highest in the group. STR performed well in excess of the expectation from single picture naming, indicating it was *less* difficult for her to produce well-formed sentences than to produce the sentence nouns in isolation. Further, she produced more closed-class elements than expected based on her open-class element production, and was more accurate at producing nouns than verbs in the sentences. STR exhibited normal semantic processing and comprehension at the single word level and good auditory-phonological discrimination, but she performed below the normal range at sentence comprehension when a reverse role distracter was present. She was able to remember a four-digit span, indicating a slightly reduced verbal short-term memory span.

STR's MRI scan revealed damage to both the left and the right hemisphere of the brain. The left hemisphere lesion encompassed a large portion of the inferior parietal lobe, extending anteriorly to encompass portions of the primary somatosensory and motor cortices (BA 3 and BA 4) and the dorsolateral prefrontal cortex, and inferiorly to encompass posterior portions of the superior and middle temporal gyri. Damage to the right hemisphere encompassed the right inferior parietal lobe, the parietal operculum, and the occipital lobe, including the associative visual cortex (BA 19) (see Figure 5.3 and Figure 5.4 for graphical depictions of STR's lesions).

SW. SW suffered a CVA four years ago. Her BDAE diagnosis was moderate to severe Wernicke's aphasia. Table 5.3 shows that SW produced fluent speech that was sometimes empty of content, with occasional phonemic and formal paraphasias and grammatical errors.

On the QPA (Saffran et al., 1989), SW's speech rate and proportion of closed-class items both fell within the normal range (Rochon et al., 2000), however, her speech rate was at the lower end of that range. Across all tests assessing single word production, SW performed poorly, indicating a significant naming deficit. While her letter fluency score was unimpaired, she scored well below the normal range in the category fluency task.

On the SPT, SW's performance fell well below the expectation from single picture naming. Her sentence score was the lowest in the group. In addition, SW's verb production ability was greatly impaired compared to her noun production ability. Further, she was the only individual in the fluent group to produce fewer closed-class elements than expected based on her open-class word production. At first glance, this seems to conflict with her unimpaired ratio of closed-class items in the QPA. However, the QPA does not penalise for the incorrect use of closed-class elements (e.g., *For that when she left*), but the SPT does. This indicates that although SW produces a normal proportion of closed-class items during spontaneous speech, she sometimes produces them in an inappropriate way (see also Table 5.3).

In line with her diagnosis of Wernicke's aphasia, SW exhibited a severe auditory comprehension deficit, as evidenced by her severely impaired performance on tests of auditory word discrimination, picture-word matching and sentence-picture matching. She was even impaired on tasks with no auditory processing requirement, such as picture-picture matching on the PPT. Further, her short-term memory span of two digits was well outside the normal range.

SW's structural MRI scan revealed extensive damage to the left posterior temporal lobe, including the posterior two-thirds of the superior (BA 22) and middle temporal gyri, further extending into the left inferior parietal lobe (see Figure 5.3 and Figure 5.4 for graphical depictions of the lesion).

WL. WL suffered a CVA two years ago. His BDAE diagnosis was mild conduction aphasia. As can be seen in Table 5.3, WL's speech was fluent with grammatically correct and elaborated sentences; however, occasionally he had word finding difficulties and also produced phonemic paraphasias. Both his proportion of closed-class items and speech rate were within the normal range (Rochon et al., 2000). Further, WL scored within the normal range on single noun naming in the Object and Action Naming Test Battery; however, his verb score was slightly lower and fell below the normal range, as did his BNT score, indicating a moderate single word production deficit. WL performed normally in letter and category fluency tasks.

On the SPT, WL performed in excess of the expectation from single picture naming, indicating that he was much more accurate at producing well-formed sentences than he was at producing the sentence nouns in isolation. Further, he produced more closed-class elements than expected based on his production of open-class elements, and was much more accurate to produce nouns than verbs in the sentences. WL exhibited normal semantic processing and comprehension at the single word level, and good auditory phonological discrimination abilities, but he exhibited some difficulties on sentence-picture matching tasks when a reverse role distracter was present. His verbal short-term memory span of three was well below the normal range.

WL's MRI scan revealed extensive damage to the left inferior parietal lobe, and the parietal operculum. His lesion also encompassed the primary motor and premotor cortices (BA 4 and BA 6) as well as the primary somatosensory cortex (BA 1, BA 2 and BA 3) and posterior parts of the superior (BA 22) and middle temporal gyri (see Figure 5.3 and Figure 5.4 for graphical depictions of the lesion).



**Summary and comment.** Overall, there are a few differences in the demographic characteristics of the control group and the two aphasia groups that are worth mentioning. First, the control group is considerably higher educated than the two aphasia groups. Second, the fluent group is considerably older than both the control and the nonfluent group. In order to take into consideration these potentially confounding factors, below we informally examine the relationships between demographic factors and error score ranges in greater detail.

There are also a few points to note about the performance characteristics of the nonfluent and fluent aphasic participants. First, speech rate, as assessed using the QPA (Saffran et al., 1989), was consistently lower for the nonfluent than for the fluent participants, and for the nonfluents, all scores were within one standard deviation of the mean reported by Rochon and colleagues for a group of individuals with Broca's aphasia (Rochon et al., 2000). Second, on the QPA proportion of closed-class words measure, scores for the nonfluent individuals varied widely, and while the majority of nonfluent participants exhibited lower scores than the fluent participants, one exception was BY. The fluent participants scored in the normal range on these measures. However, naming and single word production abilities varied widely among all aphasic individuals, as did short-term memory impairment. Importantly, severity levels were similar across both aphasia groups, so that potential performance differences in the experimental tasks are unlikely to be attributable to this characteristic.

On sentence production, as assessed with the SPT (Lukkien et al., 2012), each of the aphasic individuals was more accurate at producing nouns than verbs in the sentences, with the exception of fluent participant NP. Further, the nonfluent participants (except DA) were consistently more accurate at producing the sentence nouns in isolation than producing correct and well-formed sentences, and showed lower scores for the production of closed-

class elements than expected based on their production of open-class elements. The fluent participants showed a more heterogeneous pattern in these measures.

The only individual exhibiting a significant comprehension deficit at the single word level was fluent aphasic SW (who was diagnosed with Wernicke's aphasia). It seems likely that her comprehension deficit was also reflected in her low scores in the SPT, as the results stand in sharp contrast to her sentence production abilities in spontaneous speech (many of her responses in the SPT indicated that she did not understand or that she had forgotten the instructions).

The examination of structural MRI scans showed that nonfluent individuals primarily presented with anterior lesions to the left hemisphere, while fluent participants showed lesions mainly to posterior regions, as expected. Within the nonfluent group, all individuals presented with lesions to the posterior LIFG, including BA 44 (Broca's area); in contrast, only one fluent participant, NP, had any damage to this region, and this affected only a small portion of BA 44. Aside from RP, whose medical notes described a right medial frontal lesion (no MRI scan was available), fluent participant STR was the only individual whose MRI scan revealed damage to the right hemisphere. Two fluent individuals presented with lesions to the occipital lobes: NP had slight damage to the left primary and secondary visual cortex, whilst STR's lesion encompassed the right associative visual cortex. However, they presented with a good performance in the Peabody Picture Vocabulary Test (Dunn & Dunn, 1997) and the Pyramids and Palm Trees Test (Howard & Patterson, 1992), indicating that visual object identification was intact in these individuals; this is important, since both experiments require the analysis of complex visual stimuli. Visual attentional biases were not systemically measured, but the descriptions of the BDAE cookie theft scene indicate that all were able to attend to visual information on the right-hand side of a complex scene when no

time limit was imposed – similar conditions to the ones used in the current set of experiments.

## **Materials**

Target pictures were the 66 pictures selected from the response agreement pilot study. To avoid a priming of sentence structures (Bock, 1986b), 66 distracter pictures which elicited sentences of different syntactic structure than the SVO-target pictures were interspersed amongst the target pictures (see Appendix A for a full list), thus giving a total of 132 pictured events. Four pictures that just failed the response agreement criterion in the pilot study and four distracter pictures served as practice trials for the main experimental block. Three different fixed randomised versions of the experiment were created, each of which contained the same set of 132 pictured events, and differed only in the order of presentation of the items.

In addition, 18 single pictures of the target nouns depicted in the sentences were chosen from freely accessible picture-pools. All had been previously used in the Neuropsychology laboratory at Victoria University of Wellington, and all had elicited name agreement of 80% or higher when tested on normal controls. These single pictures were used to create a single object naming pre-test, to be administered prior to the sentence production task proper.

## **Procedure**

**Control participants.** The experiment was programmed using the *Psyscope* software (Cohen, MacWhinney, Flatt, & Provost, 1993). Participants were tested individually in a testing room at Victoria University of Wellington. They were given a questionnaire asking about demographic details and also about their neurological history. Participants also were informed that answering these questions was completely optional. However, all of them

chose to provide the information. They were then seated in front of an Apple iNotebook and asked to adjust the distance of the screen according to their needs.

Testing began with the pre-test naming task. Before the task began, written instructions appeared on the screen, informing the participants that they would see a series of pictured objects and that they were to name each one as quickly and accurately as possible. The instructions were repeated orally by the experimenter and once the participants indicated they were ready, the experiment was started. Each naming trial began with a fixation cross that appeared in the centre of the screen, followed 1000ms later by the target picture, accompanied by a tone. The picture remained on the screen until the participants had completed their response. If the participants named one of the pictures incorrectly, they were asked to think of a different word for it. They were given feedback when they had produced the target word. Each participant named all 18 pictures in this way.

Following the naming pre-test, a new set of written instructions appeared on the screen informing participants that this time, they would see a series of pictured events, and needed to describe in one sentence what was happening in each picture as quickly and accurately as possible. The instructions were orally repeated by the experimenter and the participants were told that there would be a few practice trials so they could familiarise themselves with the task. Once the participants indicated that they understood the task, the practice trials commenced. Each trial began with a fixation cross that remained on the screen for 1000 milliseconds, followed immediately by the picture. As soon as the participants began to vocalise their response to the picture, the experimenter pressed a key and the picture immediately disappeared from the screen. If there was no response within 5000 milliseconds, the picture automatically disappeared. The experimental block commenced after participants had completed the eight practice trials; participants were then asked whether they had any questions before the main part of the task began. Participants were informed that if they felt

tired during the experiment, they could have a break. At the end of the experiment, participants were thanked for their participation.

Control participants completed all three versions of the sentence production task, in different testing sessions, separated by at least a week. Each session began in the same way, with the pre-test naming task. At the end of their last session, all participants were orally debriefed about the purpose of the study and were given a written debriefing sheet to take home. All sessions were tape-recorded using a SONY MZ-M10 Hi-Mini Disc recorder for later transcription and analysis.

**Participants with aphasia.** Procedure for participants with aphasia was exactly the same as for the control participants, the only difference being that all participants except nonfluent participant JHM were tested at home (she was tested in the same testing room as the controls). Both nonfluent participants JHM and DA were not able to complete the full version of the experiment in one session due to exhaustion and slow responses, respectively. The three different versions were therefore split in half, resulting in six testing sessions for both JHM and DA. Fluent participant SW was unable to produce any of the target sentences once the picture had disappeared from the screen at speech onset, so the procedure was modified to enable her to view the picture throughout her response.

### **Response Scoring**

**Naming pre-test.** For the naming pre-test, each individual's first response to each picture was scored as either correct (identical to the target) or incorrect. Dysfluencies (e.g., *d-* *dog*) were not penalised. Naming latencies were measured manually from the digitised recordings, from the onset of the tone that accompanied the picture until the onset of the response. Responses which included errors were not submitted to the latency analysis, except if the only error was a phonological error.

**Sentence production task.** For the sentence production task, only responses to the target pictures were scored; the distracter pictures were not analysed. A response was scored as correct if it included the two target nouns and the target verb (or a permissible variant: see below), and these were incorporated into a thematically and grammatically correct sentence (complete with appropriate inflections and other closed-class morphemes). For example, passives like “The bear is being bitten by the pig” or possessive phrases like “The pig is biting the bear’s leg” were not counted as errors. However, such responses were not included in the latency analysis. Where there was more than one attempt at a sentence, only the first attempt was scored, so that self-corrections did not affect the score. Each of the constituent nouns was also scored correct or incorrect. Dysfluent utterances that involved several attempts to produce a target word (e.g., “*the ho...ho...horse*”) were not penalised. Substitution of the target noun with a pronoun (e.g., *it* instead of *pig*) was counted as an error. Verbs were scored using the same criteria, except that substitution of the target verb with a meaning-related one was allowed when it did not substantially change the meaning of the sentence (e.g., *The pig is eating the bear* instead of *The pig is biting the bear*).

Sentence production latencies were obtained manually from the digitised recordings of the sessions in the same way as for the naming pre-test data. Responses which included errors were not submitted to the latency analysis, except if the only error was a phonological one, involving either the verb or one or both nouns. Also, responses that were preceded by a comment about a previous picture (e.g. “...*Oh, that was a tricky one*”) were excluded from the latency analysis. Three different sentence production latencies were measured: *sentence initiation time*, *subject-noun initiation time*, and *verb-object noun initiation time*, which are illustrated and described in Table 5.4.

## General Statistical Methods

The analyses were conducted across items – that is, rather than collapsing across all trials of the same kind, each individual trial was treated as a separate measure. For the analysis of accuracy data for each participant group and each individual aphasic participant (correct/incorrect responses), the general approach was to use simultaneous logistic regression (Liang & Zeger, 1986). To perform these analyses, we used the Proc Genmod procedure in SAS 9.3 (SAS Institute Inc., 2011, Cary, NC). In this procedure, the regression model is built using Generalised Estimating Equations (GEEs), which enable the researcher to model the effect of a repeated measure – in this case, participant and/or target item. The values reported in the text are the Chi-Square-Score statistics associated with each predictor variable based on empirical standard error estimates.

In analyses at the group level, participant and target item were included as the repeated measures, while analyses at the individual level only included one repeated measure: target item. In all analyses, we employed the standard approach of removing non-significant predictor variables ( $p > .10$ ) from the model in a stepwise manner, always beginning with the interactions and then moving on to the main effects. In some cases, when error rates were very low and the vast majority of cells in the logistic regression analysis contained a zero, we supplemented the logistic regression analyses with an analysis of the aggregated error totals for each group in each condition (under these conditions, logistic regression has been shown to not detect effects reliably; see Peduzzi, Concato, Kemper, Holford, & Feinstein, 1996).

Table 5.4

*Description of the three different latency measures, as indicated by the arrows. The bell represents the tone that accompanies the onset of each picture.*

Latency measure	Description
<p><i>Sentence initiation time</i></p> <p>🔔 ‘The pig is biting the bear’ →</p>	<p>Time from the tone accompanying picture onset to the initiation of speech.</p> <p>If the participant started the sentence with a filler (e.g., “<i>Um...the maid is washing the clown</i>”), the filler was ignored.</p>
<p><i>Subject-noun initiation time</i></p> <p>🔔 ‘The pig is biting the bear’ →</p>	<p>Time from the tone to the onset of the first attempt at the <i>subject</i> name.</p> <p>In the case of dysfluent attempts (e.g., “<i>the ho...ho...horse</i>”), the first fragmented attempt at the picture was treated as its onset. Note that if a response lacked the subject determiner (e.g., <i>Dog is chasing a cat</i> instead of <i>The dog is chasing a cat</i>), the sentence initiation times and the subject-noun initiation times were identical.</p>
<p><i>Verb-object noun initiation time</i></p> <p>🔔 ‘The pig is biting the bear’ →</p>	<p>Time from the <i>offset</i> of the verb response to the <i>onset</i> of the object name.</p> <p>If the object determiner was missing (e.g., “<i>The dog is chasing cat</i>” instead of <i>the dog is chasing the cat</i>), verb-object initiation time was considered missing.</p>



For these aggregated error analyses, we used the Q' test (Michael, 2007), a nonparametric test designed for the analysis of nominal data from small samples. It is preferable to the Chi-Square test in that it does not assume independence of the measures.

For the analysis of latency data, we submitted the data for each item and each individual to a General Linear Mixed Model analysis (or *Mixed Effects model*; Diggle, 1988) using the SAS Proc Mixed procedure. In analyses of a single individual, the model included only one random effect: target item. In analyses of multiple individuals of the same type, the model also included participant as a random factor (see Baayen, Davidson, & Bates, 2008, for more information on the application of this technique to psycholinguistic data). In all Mixed Effects model analyses, non-significant interactions ( $p > .10$ ) of predictor variables and main effects were removed in a step-wise fashion. Covariance structure selected for all models was variance components, which gave the best model fit in the majority of the cases. Fit indices used to guide selection of the best model were Akaike's Information Criterion (Akaike, 1976) and Bayesian Information Criterion (Schwarz, 1978).

## Results

Significance is reported to an alpha level of .05 and marginal significance to an alpha level of .10. Effects of lexical frequency on accuracy and response times were of main interest for the current experiment. Therefore, effects of other variables included in the statistical models are not reported here.

### Naming Pre-test

**Analysis of response accuracy.** Table 5.5 shows the incidence of naming errors for each aphasic individual and the three participant groups (nonfluent, fluent and control). Error rates were extremely low for all participants. These data were analysed using logistic regression: The initial model included the predictor variables frequency condition (high vs. low), session (coded as an ordinal variable: one vs. two vs. three) and position in session

(first vs. second half). All possible two-way interactions were entered into the initial model.

Group analyses using logistic regression revealed no significant frequency effect for either of the two aphasia groups [ $p = .50$  for the nonfluent group and  $p = .20$  for the fluent group], nor the control group [ $p = .52$ ].

Table 5.5

*Pre-test naming errors for aphasic individuals and the three participant groups in the different frequency conditions, collapsed across all three sessions*

		Total errors across all three sessions		Items consistently failed	
	Participant	high	low	high	low
Nonfluent					
	BY	0	1	0	0
	DA	0	0	0	0
	JG	1	1	0	0
	JHM	1	0	0	0
	RP	2	0	0	0
Total		4	2	0	0
Fluent					
	NP	0	3	0	0
	STR	1	0	0	0
	SW	3	2	0	0
	WL	1	6	0	0
Total		5	11	0	0
Controls					
Total		3	5	0	0

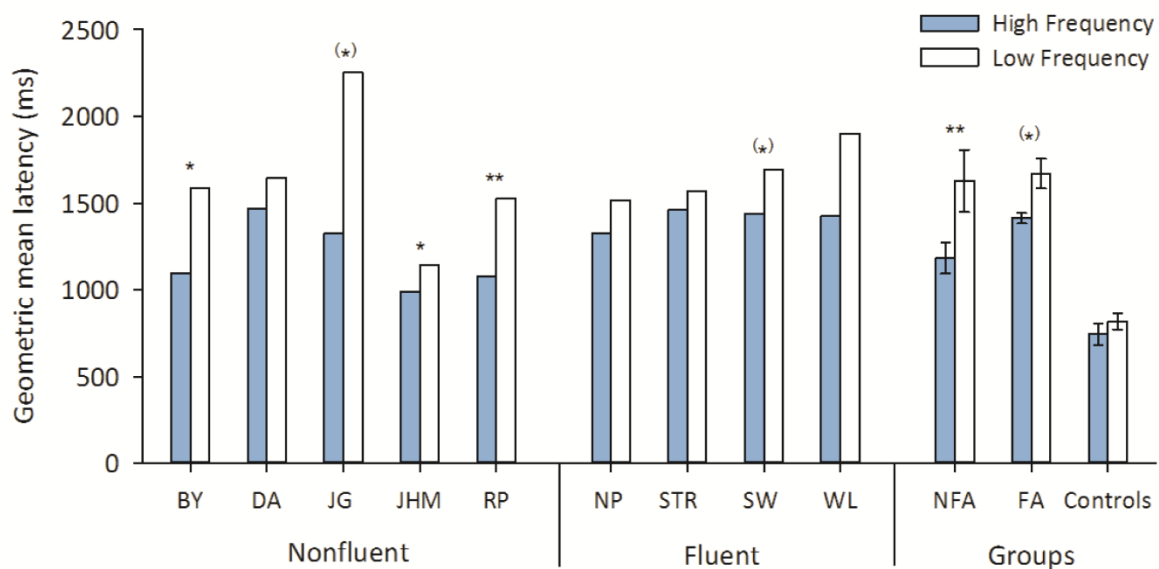
To compare performance directly across the various groups, the additional predictor variable group (nonfluent vs. fluent vs. controls) was added to the initial model. These analyses revealed that the nonfluent and the control group did not differ significantly from each other in their overall error rates [ $p=.21$ ], or in their performance across the two frequency groups [ $p=.29$ ]. The analysis of aggregated error scores using the Q' test supported this finding [ $p=.79$  and  $p=.33$  respectively]. The nonfluent group made significantly fewer errors overall than the fluent group [ $\chi^2(1)=5.17, p<.05$ ], but there was no significant group by frequency interaction [ $p=.16$ ]. Finally, the fluent group made significantly more errors than the control group [ $\chi^2(1)=6.23, p<.05$ ], but again, there was no significant group by frequency interaction [ $p=.59$ ], and results from the Q' test supported this latter finding [ $p=.24$ ]. Due to the very low error rates, logistic regression analyses of individual participants' data could not be performed (the model could not be estimated).

**Analysis of response latency.** Before the naming latency data were analysed, outliers were removed (those above or below three standard deviations of the winsorised mean for that participant/condition). This resulted in the removal of an average of 3.2% of data for the aphasic participants; no data were removed for the controls<sup>6</sup>. Latency data, particularly for aphasic participants, were highly positively skewed, the vast majority displaying a skewness measure greater than one. Consequently, the data were log-transformed in all analyses (see Biegler, Crowther, & Martin, 2008, for advantages of this procedure). Figure 5.5 shows geometric mean latencies and significance levels for the three participant groups and also individual data for each participant with aphasia (for full descriptive details, see Appendix B). The latency data for the nonfluent, fluent and control groups were submitted to a General Linear Mixed Model analysis. For both group and individual analyses, the original model contained the following predictor variables: frequency condition (high vs. low), position in

---

<sup>6</sup> An additional 3.0% of data points (across all groups) were lost due to technical problems.

session (first vs. second half), and session (coded as an ordinal variable: one vs. two vs. three). All possible two-way interactions were entered in the model. In comparisons between groups, the additional predictor variable group (nonfluent vs. fluent vs. control) was added to the initial model, while the predictor variable position in session was omitted.



*Figure 5.5.* Geometric mean latencies in the pre-test for aphasic individuals and the nonfluent, fluent and control group across the frequency conditions. Bars represent one standard error above and below the mean for each participant group. NFA = nonfluent aphasic group; FA = fluent aphasic group. (\*) =  $p < .10$ , \* =  $p < .05$ , \*\* =  $p < .01$ .

The frequency effect did reach significance for the nonfluent group [ $F(1, 227)=9.77$ ,  $p < .01$ ], and was marginally significant for the fluent group [ $p = .08$ ]. However, there was no significant effect of frequency on naming latencies for the control group [ $p = .15$ ]. Direct comparison between groups revealed that both the nonfluent [ $F(1, 625)=46.73$ ,  $p < .0001$ ] and the fluent group [ $F(1, 563)=130.19$ ,  $p < .0001$ ] were significantly slower overall at naming the pictures than controls. Furthermore, for the nonfluent group, there was a significant group by frequency interaction, indicating that frequency had a more marked effect on latencies for the

nonfluents than for the controls [ $F(1, 625)=18.33, p<.0001$ ]. This effect persisted even after the data for each group were transformed into z-scores, based on each group's own grand mean and standard deviation [ $F(1, 625)=6.21, p<.05$ ]. This finding confirmed that the effect was not simply an artefact of the widely differing means and variances of the two groups (also called “overadditivity” effect; Faust, Balota, Spieler, & Ferraro, 1999). The fluent and nonfluent group did not differ significantly in terms of their overall latencies [ $p=.28$ ], and there was only a marginally significant group by frequency interaction [ $p=.06$ ].

Each individual aphasic participant's data were also analysed in the same way as for the groups (Appendix B reports full individual subject analyses). As indicated in Figure 5.5, the effect of frequency trended towards or reached statistical significance for nonfluent participants BY, JG, JHM, and RP, and for fluent participant SW, but not for the remaining aphasic participants.

**Comment.** Overall, the aphasic individuals made very few errors in this simple picture naming task. Indeed, no person failed any item consistently across all three replications. Of course, these high naming accuracy rates may not be representative of the participants' word finding difficulties in everyday life: The words featured in this task had been practiced many times. However, for the purposes of this study, we are able to conclude with confidence that all the target nouns featured in our sentences were “within the vocabulary” of each participant (that is, they could retrieve and produce each noun on at least some occasions). The analysis of naming latencies served as a manipulation check for the frequency manipulations. It revealed a consistent trend towards slower naming latencies for low than for high frequency nouns. At the group level, this effect reached significance for the nonfluent participants and was marginally significant for the fluent participants. Further, despite lower overall levels of naming accuracy, the fluent participants did not show a significantly more marked frequency effect in their naming latencies than the nonfluent

participants. However, there was considerable individual variability within both aphasia groups as to the magnitude of this frequency effect.

### **Sentence Production Task**

Table 5.6 shows the incidence of various types of errors of the aphasic individuals and the three participant groups. Before examining the key hypotheses concerning the effect of lexical frequency, it is worth examining the overall patterns of noun production accuracy across sentence positions, irrespective of frequency.

**Noun production accuracy.** Figure 5.6 shows the overall incidence of noun errors in each aphasic individual and in the three participant groups. Also shown for comparison is the incidence of errors on the same nouns during the pre-test naming task. To compare errors on subject and object position, a multinomial logistic regression analysis was performed. The dependent variable was the total number of errors made on each sentence across the three sessions, which could range from zero to three. The predictor variables entered into the initial model included subject frequency (high vs. low), object frequency (high vs. low) and error position (subject vs. object), and all possible two-way interactions. To compare performance directly across the various groups, the additional predictor variable group (nonfluent vs. fluent vs. controls) was added to the initial model.

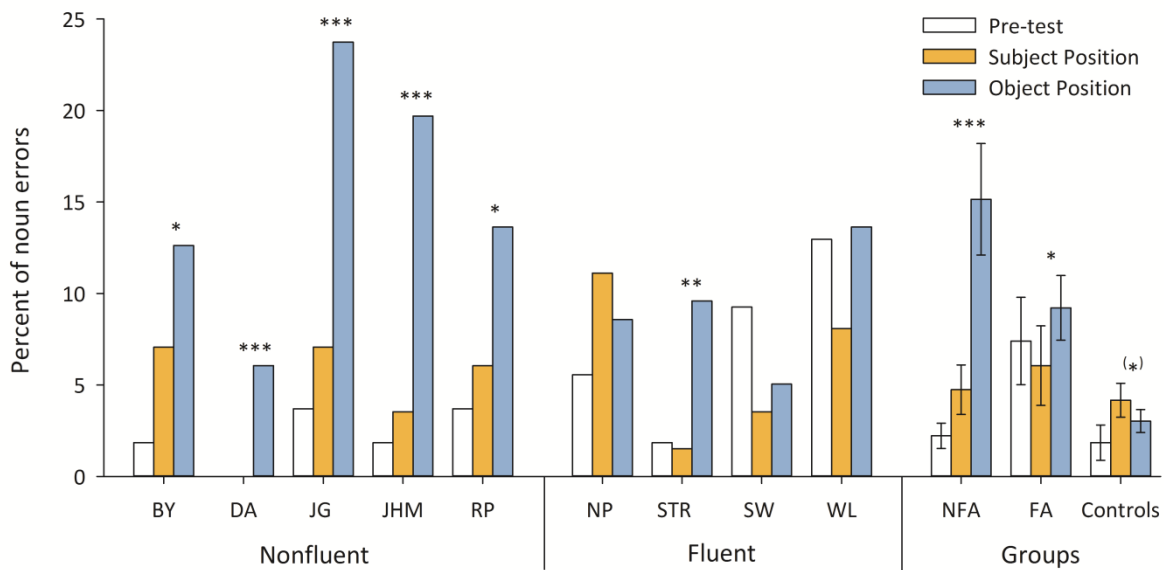
The nonfluent participants, as a group, produced more errors on nouns in object position than those in subject position [ $\chi^2(1)=43.76, p<.0001$ ]. This effect was also significant for every one of the participants with nonfluent aphasia when considered individually (see Appendix B for individual subject analyses). Indeed, if we also consider their rates of noun errors in the pre-test naming task, we observe a consistent linear pattern, in which errors were least common when the nouns were produced in isolation, moderately common when the noun appeared in subject position, and most common when the noun appeared in object position.

Table 5.6

*Percentage of errors for participants with aphasia and the three participant groups for the different frequency conditions*

Participant	Overall Sentence Errors					Subject Errors					Object Errors				
	Subject Freq.		Object Freq.		Total	Subject Freq.		Object Freq.		Total	Subject Freq.		Object Freq.		Total
	high	low	high	low		high	low	high	low		high	low			
Nonfluent															
BY	15.6	26.5	16.7	25.5	21.2	4.2	9.8	6.3	7.9	7.1	10.4	14.7	9.4	15.7	12.6
DA	11.5	18.6	13.5	16.7	15.2	0.0	0.0	0.0	0.0	0.0	5.2	6.9	4.2	7.8	6.1
JG	32.3	49.0	36.5	45.1	40.9	5.2	8.8	4.2	9.8	7.1	19.8	27.5	19.8	27.5	23.7
JHM	42.7	57.8	53.1	50.0	50.5	1.0	5.9	5.2	2.0	3.5	16.7	22.6	17.7	21.5	19.7
RP	39.6	52.9	55.2	39.2	46.5	6.3	5.9	7.3	4.9	6.1	10.4	16.7	11.5	15.7	13.6
Group mean	26.1	42.6	32.3	36.7	34.8	3.1	6.3	4.2	5.1	4.7	11.5	18.4	11.5	18.3	15.2
Fluent															
NP	22.9	20.6	19.8	24.5	21.7	8.3	13.7	11.5	10.8	11.1	11.5	5.9	6.3	10.8	8.6
STR	30.2	20.6	18.8	31.4	25.3	1.0	2.0	1.0	2.0	1.5	10.4	8.8	3.1	15.7	9.6
SW	35.4	27.5	32.3	30.4	31.3	5.2	2.0	2.1	4.9	3.5	5.2	4.9	4.2	5.9	5.1
WL	29.2	17.7	28.1	18.6	23.2	8.3	7.8	10.4	5.9	8.1	17.7	8.8	14.6	11.8	13.1
Group mean	27.1	22.4	23.8	26.5	25.4	5.3	6.6	5.8	6.1	6.1	10.3	7.4	6.5	11.5	9.1
Controls															
Group mean	5.1	9.2	5.1	9.2	7.2	3.1	5.2	2.9	5.4	4.2	2.0	4.0	2.2	3.8	3.0

*Note.* Freq. = Frequency. Error percentages were rounded to the first decimal place.



*Figure 5.6.* Percentage of nouns incorrectly produced in subject and object position for nonfluent and fluent aphasic individuals, as well as for the three participant groups. Also shown for comparison are the percentages of nouns incorrectly produced in isolation in the pre-test naming task (collapsed across all three administrations of the task). Bars represent one standard error above and below the mean for each participant group. (\*) =  $p < .10$ , \* =  $p < .05$ , \*\* =  $p < .01$ , \*\*\* =  $p < .001$ .

For controls, noun accuracy for subject and object nouns did not differ significantly [ $p = .09$ ], but it did for the fluent aphasic individuals as a group [ $\chi^2(1) = 5.46, p < .05$ ], the trend being in the same direction as for the nonfluent aphasia group. However, when considered individually, the error patterns of fluent aphasic individuals were inconsistent, and the linear pattern observed in the nonfluent participants was never observed in the fluent participants. Individual analyses revealed a significant effect of noun position for only one fluent participant, STR (see Appendix B for individual subject analyses).

Finally, direct comparisons between the nonfluent and control groups revealed a significant main effect of group [ $\chi^2(1) = 37.20, p < .0001$ ], indicating that the nonfluent group



produced more errors overall than controls. There was also a significant group by error position interaction [ $\chi^2(1)=30.67, p<.0001$ ], indicating that the nonfluents' and controls' production accuracy was differently affected by the position in which the noun appeared in the sentence. When directly comparing the fluent and the nonfluent group, there was no difference in their overall error rates [ $p=.27$ ], but a significant interaction indicated that the effect of the position in which the noun appeared in the sentence was more marked for the nonfluent than the fluent participants [ $\chi^2(1)=7.35, p<.01$ ]. The performance of the fluent group was also compared directly to the control group, and again, the analysis revealed a main effect of group [ $\chi^2(1)=26.43, p<.0001$ ], and a significant group by error position interaction [ $\chi^2(1)=8.04, p<.01$ ].

**Overall sentence production accuracy.** The percentage of incorrect sentences for each key individual/group as a function of frequency is reported in Table 5.6 (see above). To analyse these data, logistic regression was again used, but in this case, the dependent variable was binary (whether the sentence was correct or incorrect). The predictor variables entered into the initial model were subject frequency (high vs. low), object frequency (high vs. low) and session (one vs. two vs. three, coded as an ordinal variable), and all possible two-way interactions. For direct comparison between groups, the additional predictor variable group (nonfluent vs. fluent vs. controls) was added to each of the initial models.

Of primary interest in the current study was the effect of subject noun frequency on sentence errors. The relevant data is shown in Panel a) of Figure 5.7. Appendix B reports supplementary details of the statistical model, such as object noun frequency effects.

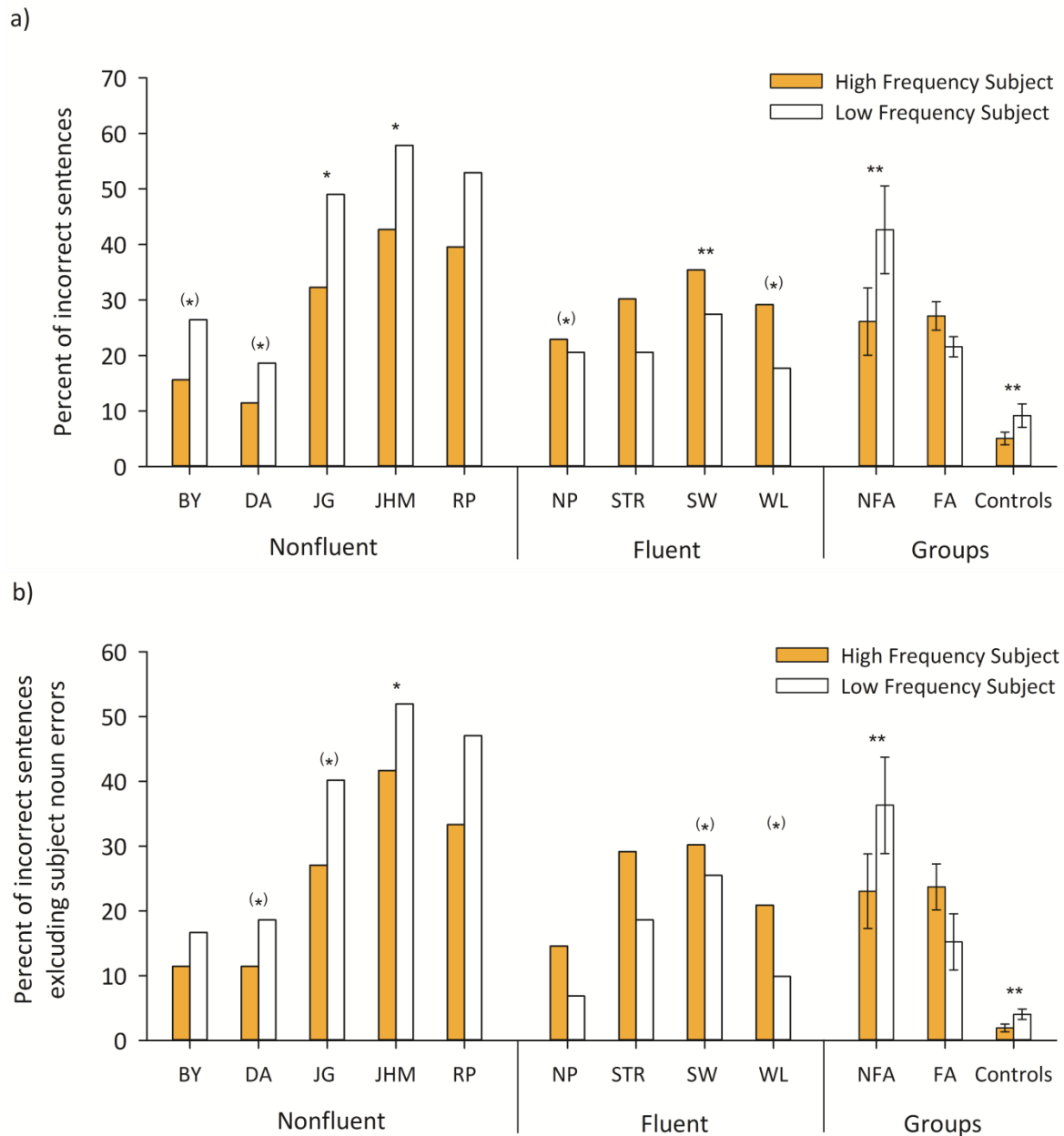


Figure 5.7. Percentage of overall sentence errors (Panel a) and sentence errors not penalising for subject errors (Panel b) as a function of subject frequency for nonfluent and fluent aphasic individuals, as well as for the three participant groups. Bars represent one standard error above and below the mean for each participant group. NFA = nonfluent aphasic group; FA = fluent aphasic group. (\*) =  $p < .10$ , \* =  $p < .05$ , \*\* =  $p < .01$ .

As can be seen in the figure, the nonfluent group produced more incorrect responses overall on sentences with a low frequency subject than on those with a high frequency subject [ $\chi^2(1)=9.89, p<.01$ ]. For the fluent group, there was a trend toward a *reverse* subject frequency effect (they made *fewer* errors on sentences with low than with high frequency subjects), but this failed to reach significance for the group considered as a whole [ $p=.33$ ]. The control group showed the same pattern as the nonfluent group, producing more incorrect responses for sentences with a low frequency subject than on those with a high frequency subject [ $\chi^2(1)=7.10, p<.01$ ].

Direct comparison between the nonfluent and control groups revealed a main effect of group, indicating that the nonfluent participants produced more incorrect sentences overall [ $\chi^2(1)=27.16, p<.0001$ ], but the interaction between group and subject frequency was not significant [ $p=.65$ ]. However, an analysis of aggregated error scores using the Q' test did reveal a significant interaction [ $Q'(1)=15.45, p<.001$ ], suggesting that the subject frequency effect was more marked in the nonfluent group than in controls. When directly comparing the nonfluent and fluent groups, there was no main effect of group [ $p=.48$ ], indicating no reliable difference in the overall number of sentences produced incorrectly. However, there was a significant group by subject frequency interaction [ $\chi^2(1)=8.04, p<.01$ ], suggesting that the reversal of the direction of the subject frequency effect in this aphasia group was statistically reliable. When compared directly to the control group, the fluent group produced more incorrect responses overall [ $\chi^2(1)=96.35, p<.0001$ ], and showed a significant group by subject frequency interaction [ $\chi^2(1)=7.43, p<.01$ ].

At the individual level, the frequency difference trended towards or reached significance for nonfluent participants BY, DA, JG and JHM, while for the fluent participants it reached significance for SW and was trending for NP and WL (see Appendix B for full individual subject analyses).

To ensure that the subject frequency effects were not just due to localised errors on the subject nouns themselves, but were rather affecting downstream sentence planning, Panel b) of Figure 5.7 gives overall sentence errors not penalising for subject errors. As can be seen in the figure, the error patterns stay the same for all individuals and groups, demonstrating that the effect of subject frequency on sentence accuracy cannot be driven by local effects (full statistical details of the group and individual analyses are reported in Appendix B).

**Analyses of response latency.** Latency data for each individual was trimmed of outliers using the same procedure as for the naming pre-test. Collapsed across all individuals, this resulted in the removal of an average of 2.0% of data for the aphasic participants, and 0.9% for the control group<sup>7</sup>. As in the pre-test, latency data were then log-transformed to reduce skewness. Appendix B reports geometric mean latencies for all latency measures and frequency conditions for individuals with aphasia and the three participant groups. Subject and object frequency effects on sentence initiation and verb-object noun initiation times for the three participant groups and aphasic individuals are illustrated in Figure 5.8. Since the results for subject-noun initiation times were very similar to results obtained from the analysis of sentence initiation time, they are only reported where relevant in favour of conciseness (full statistical details of the subject-noun initiation time analysis for aphasic individuals and the three participant groups are reported in Appendix B).

As in the naming pre-test, sentence production latency data were submitted to a General Linear Mixed Model analysis. The initial statistical model included the following predictor variables: subject frequency (high vs. low); object frequency (high vs. low); session (coded as ordinal, one vs. two vs. three); and the number of times the nouns in the target picture had already occurred in the same session either in the subject or object position

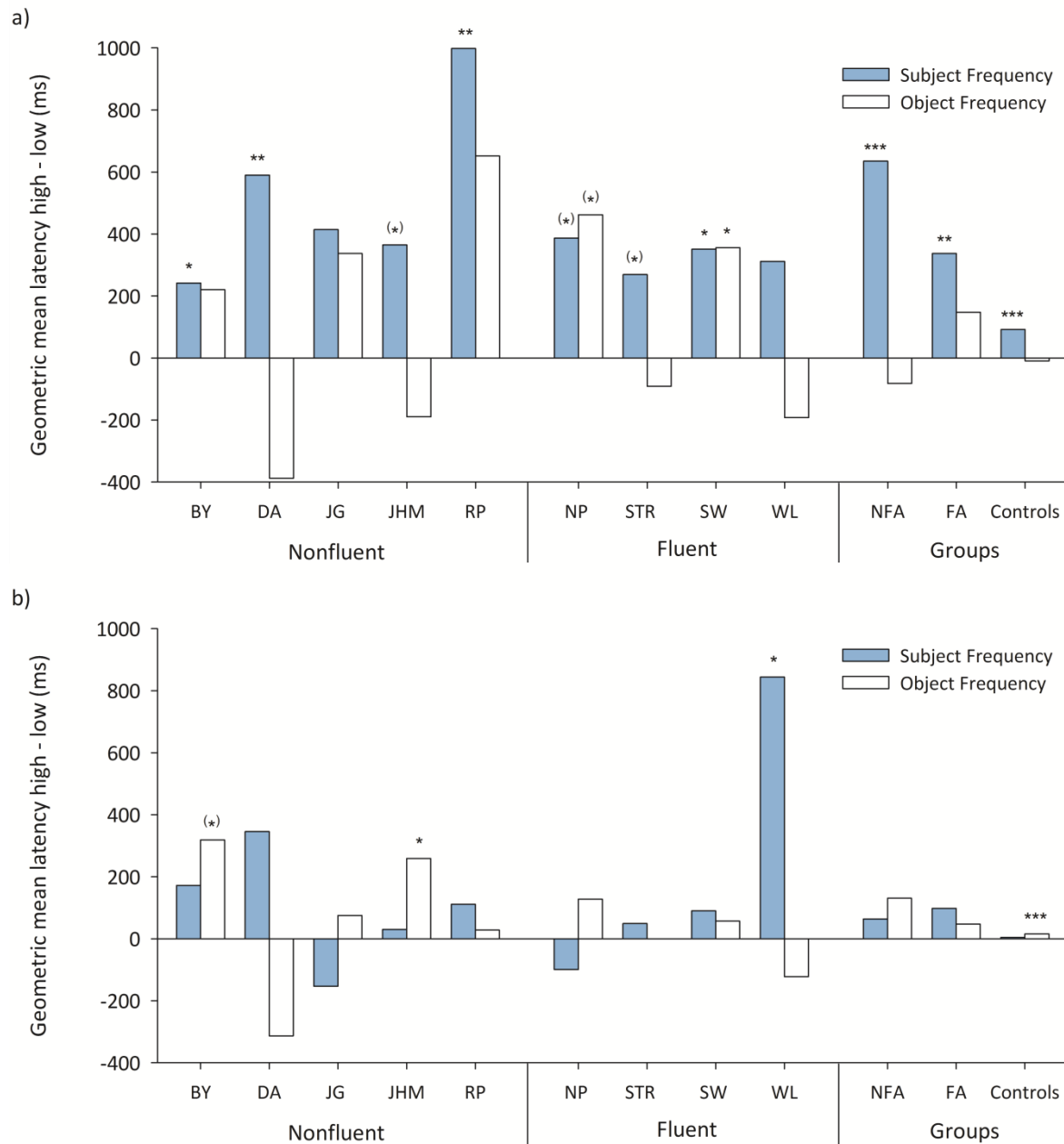
---

<sup>7</sup> An additional 2.8% of data points (across all groups) were lost due to technical problems or other disturbances (like background noise).

(coded as a scale, with values ranging from zero to 20). Two-way interactions of all predictor variables were also included. To compare performance of both aphasia groups with each other and the control group, the predictor variable group (nonfluent vs. fluent vs. controls) was added. The predictor variable that indicated the previous repetition of the target nouns was dropped from the model, while the remaining variables were the same.

**Sentence initiation times.** Panel a) of Figure 5.8 shows the difference in sentence initiation times as a function of subject and object frequency for each aphasic individual and the three participant groups. At the group level, the findings were very similar for each of the three participant groups: Sentence initiation times were significantly influenced by subject frequency [nonfluents:  $F(1, 672)=17.89, p<.0001$ ; fluents:  $F(1, 440)=10.02, p<.01$ ; controls:  $F(1,1337)=16.21, p<.0001$ ].

The trend illustrated in Panel a) of Figure 5.8 towards more marked frequency effects for the nonfluent participants than for controls appeared to be statistically reliable: The nonfluent group not only took longer overall to initiate sentences when compared to controls [ $F(1, 2066)=75.83, p<.0001$ ], but also, there was a significant group by subject frequency interaction [ $F(1, 2066)=12.60, p<.001$ ]. However, when this analysis was re-run with z-scores to check for “overadditivity” effects, the interaction disappeared [ $p=.36$ ]. The nonfluent and the fluent groups themselves did not differ significantly from one another in their sentence initiation times [ $p=.26$ ], and there was no significant group by subject frequency interaction [ $p=.48$ ]. When compared to the controls, the fluent participants showed a similar effect than the nonfluent participants: They also took longer overall to initiate sentences than controls [ $F(1, 1836)=149.34, p<.0001$ ], and there was also a significant group by subject frequency interaction [ $F(1, 1836)=4.74, p<.05$ ]. However, again, this interaction disappeared when the analysis was re-run using z-scores [ $p=.98$ ].



*Figure 5.8.* Subject and object frequency effects on sentence initiation time (Panel a) and verb-object noun initiation time (Panel b) depicted as the difference in latencies between low and high frequent nouns for nonfluent and fluent aphasic individuals, as well as for the three participant groups. A positive value indicates a faster response time for high frequency nouns. NFA = nonfluent aphasic group; FA = fluent aphasic group. (\*) =  $p < .10$ , \* =  $p < .05$ , \*\* =  $p < .01$ , \*\*\* =  $p < .001$ .

At the individual level, there were significant main effects of subject frequency for nonfluent participants BY, DA and RP (and a marginally significant trend for JHM), and also for fluent participant SW (and a marginally significant trend for participants NP and STR; see Appendix B for full details of individual subject analyses).

It is perhaps not surprising that participants were slower to initiate responses when the subject noun was lower in frequency. Of greater theoretical interest for examining planning strategies is whether *object* frequency had any effect on sentence initiation times: Such an effect would indicate significant substantial advance planning of the object noun prior to initiation of the response. Panel a) of Figure 5.8 shows that this variable had a considerably weaker effect than subject frequency. Indeed, at the group level, neither the nonfluent, nor the fluent or the controls participants showed a reliable object frequency effect [ $p=.70$ ,  $p=.27$ , and  $p=.76$  respectively], or a significant subject frequency by object frequency interaction [ $p=.69$ ,  $p=.59$ , and  $p=.54$  respectively].

Analyses across groups revealed that there were no significant group by object frequency interactions [nonfluent vs. controls:  $p=.43$ ; nonfluents vs. fluents:  $p=.58$ ; fluents vs. controls  $p=.08$ ]. However, despite these negative group effects, it can be seen in Panel a) of Figure 5.8 that some individuals, when considered on their own, revealed significant object frequency effects. Fluent cases NP and SW both exhibited significant effects of object frequency on sentence initiation times [ $p=.05$  and  $p<.05$  respectively]. Nonfluent cases BY, JG and RP revealed a nonsignificant trend in this direction. However, these individuals exhibited a pattern that suggested that much of their sentence planning was being delayed until the interval between the determiner and the subject noun (a strategy which is effective, given that every sentence begins with the same determiner “The”). Indeed, when taking results from *subject-noun* initiation times into account, object frequency had a marginally significant effect in RP [ $p=.07$ ] and was approaching a trend for JG [ $p=.17$ ], indicating that it

took them longer to produce the subject when the object of the sentence was of low frequency. However, this was not confirmed for BY, who showed no object frequency effect on subject-noun initiation time [ $p=.75$ ].

**Verb-object noun initiation times.** Another measure of production latency that might shed light on the sentence planning processes is verb-object noun initiation time, the time between offset of the target verb and onset of the object noun. Panel b) of Figure 5.8 presents these data. At the group level, neither the nonfluent, nor the fluent or the controls participants showed a reliable subject frequency effect on object noun initiation times [ $p=.48$ ,  $p=.14$  and  $p=.95$  respectively]. Across group analysis revealed that verb-object noun latencies were longer in the nonfluent group than in the control group [ $F(1, 1942)=107.30$ ,  $p<.0001$ ], but there was no significant group by subject frequency interaction [ $p=.41$ ]. When directly comparing the two aphasia groups, there was no significant difference in verb-object initiation times [ $p=.27$ ], and again, no significant group by subject frequency interaction was found [ $p=.37$ ]. When compared to the controls, the fluent group took significantly longer to produce the object noun [ $F(1, 1722)=22.12$ ,  $p<.0001$ ]. The comparison further revealed a significant group by subject frequency interaction [ $F(1, 1722)=4.12$ ,  $p<.05$ ], however, this effect disappeared after re-running the analysis with z-scores [ $p=.88$ ].

At the individual level, only fluent participant WL showed a significant subject frequency effect: He took reliably longer to produce the object when the sentence began with a low frequency subject than when it began with a high frequency subject (see Appendix B for full details of individual subject analyses).

Turning now to the effect of *object* frequency on verb-object noun initiation times, the analysis failed to reveal any reliable object frequency effects for the nonfluent and the fluent group [ $p=.27$  and  $p=.55$  respectively], and no significant interactions between subject and object frequency were found [ $p=.16$  and  $p=.14$  respectively]. Finally, the analysis of control



participants revealed a significant main effect of object frequency on verb-object noun initiation times [ $F(1, 1232)=28.05, p<.0001$ ]. Curiously, there was also a significant interaction between subject and object frequency [ $F(1, 1232)=4.17, p<.05$ ], indicating that object frequency had a stronger influence on verb-object noun initiation times when the sentence began with a high frequency than with a low frequency subject.

When compared across groups, verb-object noun latencies were longer in both aphasic groups than in the control group [nonfluents vs. controls:  $F(1, 1942)=107.30, p<.0001$ ; fluent vs. controls:  $F(1, 1722)=22.12, p<.0001$ ], but there was no significant difference between the two aphasic groups [ $p=.27$ ]. Further, there were no significant group by object frequency interactions [nonfluent vs. controls:  $p=.66$ ; nonfluents vs. fluents:  $p=.70$ ; fluents vs. controls  $p=.58$ ].

At the individual level, there were only effects of object frequency for nonfluent participants BY and JHM, and for JHM, a significant interaction between subject and object frequency (verb-object latencies were shortest when the sentence contained *both* a high frequency object *and* a high frequency subject, and were longest when the sentence subject was of high frequency and the sentence object of low frequency).

**Comment.** Consistent with the latency analysis of single noun naming, all participants were faster to initiate sentences beginning with a high than a low frequency subject, indicating that lexical retrieval of the subject noun had already commenced at the time of utterance initiation. Not surprisingly, the pattern was also evident when initiation time was measured to the onset of the first subject noun, rather than from the initial word “The”. However, object frequency did not consistently influence sentence initiation times, suggesting most participants had not begun to retrieve the lexical label for the object element at that time. Indeed, controls showed no object frequency effect, which is consistent with previous studies of normal speakers, in which lexical elements have been found to influence

sentence initiation times only when they appear in the initial noun phrase (e.g., F. Ferreira, 1991; Griffin, 2001; Martin et al., 2004; Smith & Wheeldon, 1999). Nevertheless, while there was no significant effect of object frequency at the group level for either aphasia group, some specific individuals did show statistically reliable effects. Fluent participants NP and SW were affected by object noun frequency on their sentence initiation times, and results from subject-noun initiation times indicated a similar tendency for nonfluent participants RP and JG. This pattern, which curiously cut across the fluent/nonfluent distinction, suggests a strategy of more extensive advance planning in these participants when compared to normal speakers. We refer to these individuals as *whole sentence planners*, as distinct from *phrase-by-phrase planners*, a term which characterises the other aphasic participants and the controls.

In the latency analyses, we also examined the time interval between the verb offset and the onset of the object noun. In controls, this interval was reliably longer for low frequency than for higher frequency object nouns, suggesting that object noun retrieval may still not have been complete at the time the verb was produced. Again, this pattern suggests an incremental, phrase-by-phrase planning strategy in controls. Few of the aphasic individuals showed this effect, but two exceptions were nonfluent cases BY and JHM. Interestingly, for controls and for JHM, there was also a significant subject frequency by object frequency interaction, characterised by larger object frequency effects when the subject of the sentence was of high frequency. This latency pattern indicates a temporal overlap between the retrieval of the subject noun and the object noun in JHM, which was only observable when the subject noun was produced so quickly that retrieval of the lexical label for the object noun had not yet been completed.

## Discussion

Experiment 1 yielded two important findings. First, when we considered participants' accuracy in producing the key nouns themselves, the nonfluent participants exhibited a very distinctive and consistent pattern: Irrespective of frequency, their noun production accuracy was higher when the nouns were produced in isolation (in the naming pre-test) than when they were produced in the context of a sentence. This finding is consistent with several other studies that showed that individuals with nonfluent aphasia perform significantly better when naming pictures that depicted a single object than when they were required to integrate the named objects into a sentence (e.g., Schwartz & Hodgson, 2002; Williams & Canter, 1982). Further, nonfluent individuals were more accurate at producing the nouns when they appeared in subject position than when they were in object position. These patterns were not consistently observed in the control participants or the fluent aphasic participants, whose performance across the three context conditions was more variable. In fact, the majority of the fluent participants were *less* accurate at producing the target nouns in isolation than as the subject of a sentence. The consistent error pattern within the nonfluent group indicates remarkable effects of context on word production in nonfluent aphasia. These effects are all the more notable given that the target nouns were within every participant's range of vocabulary, meaning that no single noun was consistently failed in the naming pre-test.

Second, in SVO sentence production, the frequency of the subject noun had a powerful influence on the nonfluent participants' sentence production accuracy. These individuals produced more sentences incorrectly when the subject was of low frequency than of high frequency and this effect persisted even when we examined accuracy measures that excluded the subject noun itself. Again, this effect was not observed for the fluent aphasic participants. The control participants showed similar frequency effects in the sentence accuracy measures than the nonfluent individuals, however, not surprisingly, of much smaller

magnitude. Taken together, subject noun accessibility appears to have a profound effect on downstream sentence production accuracy and well-formedness only in individuals with nonfluent aphasia. These findings support the initial hypothesis that the frequency of lexical target items – at least when they appear early in the sentence – has more generalised effects on sentence production accuracy in nonfluent than in fluent aphasia.

Within a Kolk-type framework (e.g., Kolk, 2006), both of the features of nonfluent aphasia displayed in this experiment could be interpreted as an inability to retrieve early sentence elements in sufficient time to enable late-appearing lexical elements to be integrated into the current representation of the sentence frame. However, there was no direct evidence to suggest that lexical retrieval itself was any slower in the nonfluent than in the fluent aphasic participants: Single picture naming latencies did not differ significantly for the two groups, nor did any of the sentence production latency measures. Both participant groups were much slower than the controls, but they were slow to the same extent. An alternative possibility within the Kolk framework is that syntactic frame construction may itself be delayed, so that the activation levels of some lexical elements have already begun to decay by the time their corresponding frame elements become available. However, this would in fact predict higher accuracy for sentences containing low frequency elements, since the more slowly the words are retrieved, the greater their synchronicity with the delayed frame construction process. This was the opposite of what we found.

A more plausible explanation is that in nonfluent aphasia, the structure generation process itself is ineffective, so these individuals are more heavily reliant on the representations of the lexical constituent elements themselves to support a structural plan for the sentence. This idea could in principle be accommodated within a more formal model of multiple word planning, such as that proposed by Randi C. Martin and colleagues (Freedman et al., 2004; Martin & Freedman, 2001; Martin et al., 1999; Martin et al., 2004; Martin &

Romani, 1994; Martin, Shelton, & Yaffee, 1994). As mentioned previously, Martin and colleagues postulate that the order of lexical elements during sentence planning is maintained in semantic short-term memory. During production, lexical elements that form part of the utterance are tied to their respective slots in the buffer system. Interactive flow of activation between the nodes and the lexical elements ensures that information in the buffer is maintained over the course of production. Once it is time to produce a particular word in the utterance, its “slot” receives a boost of activation, thereby heightening the state of activation of its corresponding lexical elements (e.g., Freedman et al., 2004). This model is able to capture some of the ways in which content and structure processes interact in speech – at least in the context of very simple utterances, such as noun phrases.

Within this framework, inefficient binding of content elements to their respective “slots” and/or abnormally rapid decay of this information could result in a failure to produce the utterance correctly. Crucially, the effects of such an impairment would be modulated by lexical factors. Since there is a bidirectional flow of activation between elements and their “slots”, strongly activated lexical elements (such as those for common, frequently used words) may be maintained in semantic short-term memory for longer than would otherwise be possible. Although this framework deals only with planning within a single noun phrase, it could in principle be extended to account for planning over more complex syntactic structures.

The idea that structure generation processes are slow or ineffective in nonfluent aphasia was also explored by Gordon and Dell (2003). In their DoL model, activation of lexical elements within the model depends on both conceptual-semantic and syntactic-sequential input nodes. Lesioning the connection strength of syntactic input nodes to lexical elements in the model led to an agrammatic speech pattern. However, it had other consequences too: When no structural representation was in place to ensure that lexical

elements planned for an utterance were produced in the appropriate order, they were freer to compete for production. So consequently, noun production may be less accurate in the context of a sentence – where syntagmatic competition is high – than in isolation. However, the DoL model has difficulty explaining why the nonfluent participants' accuracy of downstream sentence elements in the current experiment was affected by subject frequency. In fact, although sharing the labour involved in lexical element retrieval, the input processes of syntactic and semantic-conceptual components in the model are distinct. Therefore, the availability of lexical elements may have local effects, potentially making low frequency words harder to retrieve than high frequency ones, but subject frequency would not be expected to have any impact on sentence production accuracy in general.

In conclusion, structure-based accounts do not provide an appropriate framework for the current findings, because they do not offer any explanation for the effect of lexical content on sentence production accuracy. We will consider some alternative accounts in the General Discussion section.

One further aspect of the results from Experiment 1 warrants some discussion. In this experiment, the stimulus picture disappeared from view as soon as the participants commenced their response. The rationale for doing this was to promote advance planning of the entire sentence. However, it remains possible that in some participants, the intended conceptual message might have decayed from working memory before they had an opportunity to encode each element into a linguistic form. If the nonfluent participants were more prone to this problem than the fluent participants, it could account for why they performed more poorly on sentences with low frequency subject nouns. However, the results provide little support for this hypothesis. First, all the nonfluent participants were able to complete the task effectively and could correctly produce at least some of the sentences in all conditions; indeed, the only participant who could not complete the task under the original

conditions was SW, who had fluent (Wernicke's) aphasia. Second, much of the delay induced by the subject noun frequency manipulation is likely to have occurred before the sentence was initiated, and before the picture had been removed from view. Third, there is direct evidence from a number of participants – both fluent and nonfluent – that lexical-level planning of the object element had begun in earnest even before they had begun to initiate their response. This would seem to preclude an explanation in terms of decay of the relevant conceptual representations. We return to this point in the discussion of Experiment 2.

Finally, there is no evidence that the differences in education and age between the three participant groups had a confounding effect on the sentence production performance in this experiment. For example, both control participants with the highest and lowest overall error rates had a university degree. Further, only two participants with aphasia had a university education: JHM, who produced the highest error rates in the nonfluent group, and WL, who produced the second to lowest error rates in the fluent group. Furthermore, we found no effect of age on sentence production performance. Although the fluent aphasia group was on average older than the control group and the nonfluent aphasia group, there were no significant differences in overall accuracy and latency between the two aphasia groups. Compared to the control group, the fluent group was significantly slower and less accurate – however, the same was found for the nonfluent group. Furthermore, all fluent participant showed the same overall error pattern in this experiment, including WL and SW, whose age was comparable to the average age of the control and the nonfluent group. Taken together, there does not appear to be a systematic relationship between age and education and sentence production performance in this experiment.

To summarise, the marked effects of lexical frequency on sentence production accuracy in nonfluent aphasia suggest that structure generation processes in these individuals may be ineffective. Some models predict that a weak structural representation will further

result in higher syntagmatic interference between the lexical elements that form part of the same utterance, because those elements are not effectively tied to their appropriate position in the sentence (Freedman et al., 2004; Gordon & Dell, 2003). Consequently, individuals with nonfluent aphasia may be particularly susceptible to syntagmatic interference effects when they are planning to produce a sentence. This proposal is specifically investigated in Experiment 2.



## **Experiment 2: Effects of Noun Semantic Relatedness on Sentence Production**

The current experiment aims to explore the effect of syntagmatic interference on sentence production in nonfluent aphasic, fluent aphasic and control participants by manipulating the semantic relationship of sentence nouns. It has been suggested that a robust structural representation of the target utterance helps to manage the naturally competitive influence of one word on another when both have similar selection constraints (Dell et al., 2008; Gordon & Dell, 2003; Stemberger, 1985). If this structural representation is absent, lexical items may be freer to compete for selection for the same position in the utterance plan. Consequently, if individuals with nonfluent aphasia have difficulty constructing and/or maintaining a structural representation of the sentence, then the lexical elements planned for the utterance may be particularly prone to confusion.

One way to explore the extent of lexical competition during sentence production is to manipulate the semantic similarity of lexical elements within the utterance being produced. Within current theories of language production (e.g., Dell, 1986; Levelt, 1989), two meaning-related words that need to be produced in one sentence will activate one another via the semantic nodes that they share and, thus, compete with one another for selection to a greater extent than unrelated words. If so, then we might expect individuals with nonfluent aphasia to be disproportionately poor at producing sentences that contain semantically related words. Several recent studies examining phrase production in nonfluent aphasia have found support for this prediction (e.g., Freedman et al., 2004; Scott & Wilshire, 2010). Although semantic interference effects have never been examined across phrases (at least not to our knowledge), there was evidence from Experiment 1 that some aphasic individuals plan the sentences entirely in advance and if so, then semantic interference effects might occur over utterances larger than a single phrase.

In the current experiment, a sentence production task similar to the one used in Experiment 1 was administered to the same aphasic and control participants. However, this time the semantic relationship between the noun items in the SVO sentence was varied: Nouns were either from the same category or from a different category. If, as hypothesised, individuals with nonfluent aphasia are more susceptible to semantic interference effects due to weak structural representations, they should be less accurate and/or slower at producing sentences containing semantically related subject and object nouns, when compared to those containing unrelated subject and object nouns. A similar, but weaker effect was expected for the control participants. In contrast, individuals with fluent aphasia were not expected to show this effect and might actually demonstrate a facilitatory semantic priming effect when the two nouns share meaning.

As in Experiment 1, a pilot study was conducted to ensure sufficient response agreement to the pictures used in the sentence production task. Therefore, data for two experiments is reported. The first describes the pilot study procedure and the construction of the stimulus pictures, while the second reports the aphasic and control production data from the sentence production task proper.

## Response Agreement Pilot Study

### Method

The School of Psychology Ethics Committee of Victoria University of Wellington gave ethical approval for this study in 2011.

### Participants

Thirty-three psychology students enrolled in a first-year course at Victoria University of Wellington took part in this study for course credit. The sample consisted of students ranging in age from 18 to 27 years ( $M=18.70$ ,  $SD=1.63$ ), out of which 14 were male (42%). The participants were all English native speakers. Three participants were excluded from the study: One participant did not follow the instructions, and two participants scored below 1.5 standard deviations of the mean accuracy score calculated across all participants. This left a total of 30 participants for the analysis, their age ranging from 18 to 27 years ( $M=18.72$ ,  $SD=1.65$ ). Twelve participants (40%) were male.

### Materials

Fifty-four new line drawings depicting simple agent-object events were drawn specifically for Experiment 2. Their format was identical to that of Experiment 1: Each picture displayed a scene that could be described in a SVO sentence. The drawings depicted the same noun elements that appeared in Experiment 1, except that *dog* and *fox* were not used, and *snake* was replaced with *skunk* (a noun with a similar lexical frequency that also met our previous age of acquisition criterion, see Appendix C). Twelve different verbs were depicted in the scenes. In the drawings, both noun elements in a sentence were always from the same frequency group: High frequency nouns always appeared with other high frequency nouns, and low frequency nouns always appeared with other low frequency nouns, according to the criteria used in Experiment 1. In the majority of the drawings, the two key nouns were from different broad categories (e.g., one was an animal and one was the name of a type of

person). As in Experiment 1, each target item appeared twice, in two different picture scenes, once as the subject of the target sentence and once as the object, and the subject was always displayed on the left side of the picture. Forty-eight of the drawings depicted subject and object nouns from different categories (24 depicted nouns of high frequency, and 24 depicted nouns of low frequency). The remaining six pictures displayed low frequency nouns that were from the same category.

### **Procedure and Scoring**

The administration and scoring procedure applied was identical to that used in the response agreement pilot study for Experiment 1.

### **Results**

Response accuracy ranged from 50 to 100 per cent ( $M=85.51$ ,  $SD=12.38$ ). Forty-four of the new pictures reached name agreement of 80%. However, six of these pictures had to be excluded because the reversed partner picture did not reach the required agreement level. This left a total of 38 pictures; 36 depicted unrelated items (18 pictures containing high frequency items and 18 pictures containing low frequency items), and two pictures which contained related items (of low frequency). All of these pictures were selected to serve as the material for the sentence production task, which will be reported in the following section.

### Sentence Production Task

The School of Psychology Ethics Committee of Victoria University of Wellington gave ethical approval for this experiment in 2011.

#### Participants

The participant group for this study consisted of the same nine individuals with aphasia and eight older controls who took part in Experiment 1.

#### Materials

All 38 of the pictures that reached the 80% criterion in the response agreement pilot study were selected to serve as target pictures in this experiment. These pictures were combined with 26 pictures from Experiment 1 to create a final set of 64 pictures in total. The 64 final stimulus pictures were composed of: a) 28 *related* pictures, which depicted a subject and an object noun from the same broad category - either animals (e.g., *bear*, *skunk*) or real/imagined persons (e.g., *king*, *ghost*); and b) 36 *unrelated* pictures depicting subject and object nouns from different categories. As mentioned above, noun frequency was balanced so that half of the related pictures depicted two high frequency nouns and the other half depicted two low frequency nouns, and the same was true for the unrelated pictures (see Appendix C for a complete list of the target sentences). Again, the target pictures were combined with 64 additional distracter pictures depicting various sentence structures (for a full list see Appendix A), and again, the entire set of pictures was administered three times across three different testing sessions, each time in a different, fixed randomised order. Practice items were the same as for Experiment 1.

Similar to Experiment 1, a pre-test single picture naming task was also constructed, in which each of the 16 noun elements that appeared in the scenes was depicted in isolation. All pictures were drawn from freely accessible picture pools and had reached name agreement of 80% or higher in previous norming studies.

## Procedure and Response Scoring

The administration and scoring procedures were identical to those used in Experiment 1; however the debriefing sheet that was given to the participants after the last testing session was different.

## General Statistical Methods

Accuracy and latency data were analysed using the same statistical procedures as outlined in Experiment 1. The only exception was that in all models, semantic relatedness (related vs. unrelated) and frequency (high vs. low) were included as the main predictor variables, instead of subject frequency and object frequency.

## Results

Significance is reported to an alpha level of .05 and marginal significance to an alpha level of .10. Effects of semantic relatedness on accuracy and response times were of main interest for the current experiment and are reported in the following. Other variables included in the statistical models are not reported, however, results from the frequency manipulations are presented in Appendix D.

### Naming Pre-test

**Analysis of response accuracy.** As for Experiment 1, accuracy on the naming pre-test was close to ceiling in all participants (Appendix D reports the full descriptive data). There were no significant frequency effects for either of the two aphasia groups considered alone [nonfluents:  $p=.31$ ; fluents:  $p=.09$ ]. Due to the very low error rates, logistic regression analyses of the controls could not be performed (the model could not be estimated). However, the analysis of aggregated scores was possible and revealed no effect of frequency on naming errors [ $p=.38$ ].

Accuracy data was also analysed across groups. The nonfluent group made significantly more errors in the pre-test than the control group [ $\chi^2(1)=4.55, p<.05$ ], but the

model could not estimate the interaction parameters. Again, an analysis using aggregated error scores was possible, and revealed no significant group by frequency interactions [ $p=.40$ ]. A direct comparison of the nonfluent and fluent groups failed to yield a reliable overall difference in errors between the two groups, [ $p=.28$ ], or a significant group by frequency interaction [ $p=.94$ ]. Finally, when directly comparing the fluent group with the control group, the analysis revealed that the fluent group made significantly more errors than the controls [ $\chi^2(1)=8.08, p<.01$ ], but again, interaction parameters could not be estimated in the model. The analysis of aggregated error scores revealed no significant group by frequency interactions [ $p=.57$ ].

Logistic regression analyses of individual participants' data could not be performed due to very low error rates (the model could not be estimated).

**Analysis of response latency.** The naming latency data were trimmed for outliers and log-transformed using the same procedures as described in Experiment 1 (this led to a reduction of an average of 2.5% of data for the aphasic participants and 1.3% for the controls<sup>8</sup>). There was a trend across all groups and the majority of individual aphasic participants towards shorter latencies on high frequency words (see Appendix D for full details of descriptive data). However, the effect of frequency failed to reach significance for the nonfluent group [ $p=.12$ ] but was significant for the fluent group [ $F(1, 154)=6.00, p<.05$ ]. The control group showed only a marginally significant frequency effect [ $p=.06$ ].

Further analysis across groups revealed that the nonfluent group was overall significantly slower to name the pictures than the controls [ $F(1, 560)=30.55, p<.0001$ ], but no significant group by frequency interaction was found [ $p=.15$ ]. When directly comparing the two aphasic groups, there was no significant difference in naming latency [ $p=.28$ ], and no significant interaction between group and frequency [ $p=.94$ ]. Finally, a direct comparison of

---

<sup>8</sup> In addition, technical problems resulted in a further loss of 5.8% of data points (across all groups).

the fluent group and the control group revealed significantly slower naming latencies for the fluent group [ $F(1, 510)=51.91, p<.0001$ ], and there was also a significant group by frequency interaction [ $F(1, 510)=8.04, p<.05$ ]. However, this interaction was nonsignificant after re-running the analysis with z-scores [ $p=.38$ ].

At the individual level, the effect of frequency was trending towards or reached statistical significance for nonfluent participants BY, DA and JHM, and for fluent participant NP and SW, while no effects were found for the remaining aphasic participants (Appendix D reports full details of individual subject analyses).

**Comment.** Individuals with aphasia and the control participants again performed very accurately in pre-test naming. In fact, they were even more accurate than in Experiment 1, most likely because they were more familiar with the items by this point. This time, the nonfluent and the fluent group did not differ in their error rates. Again, no target noun was failed consistently by any of the participants across the three sessions of the pre-test naming. Again, the nonfluent and the fluent groups were slower to name the pictures than the controls and did not differ in their response times when compared to each other.

### **Sentence Production Task**

**Noun production accuracy.** Table 6.1 reports the percentages of different types of sentence production errors for the aphasic individuals and the three participant groups. Figure 6.1 plots the noun production accuracy for each individual and group as a function of utterance context. It can be seen from the figure that the pattern observed in Experiment 1 was replicated in Experiment 2: That is, the nonfluent participants showed an incremental pattern of noun errors, which was lowest for nouns produced in isolation, and highest for nouns produced in object position within the SVO sentences.

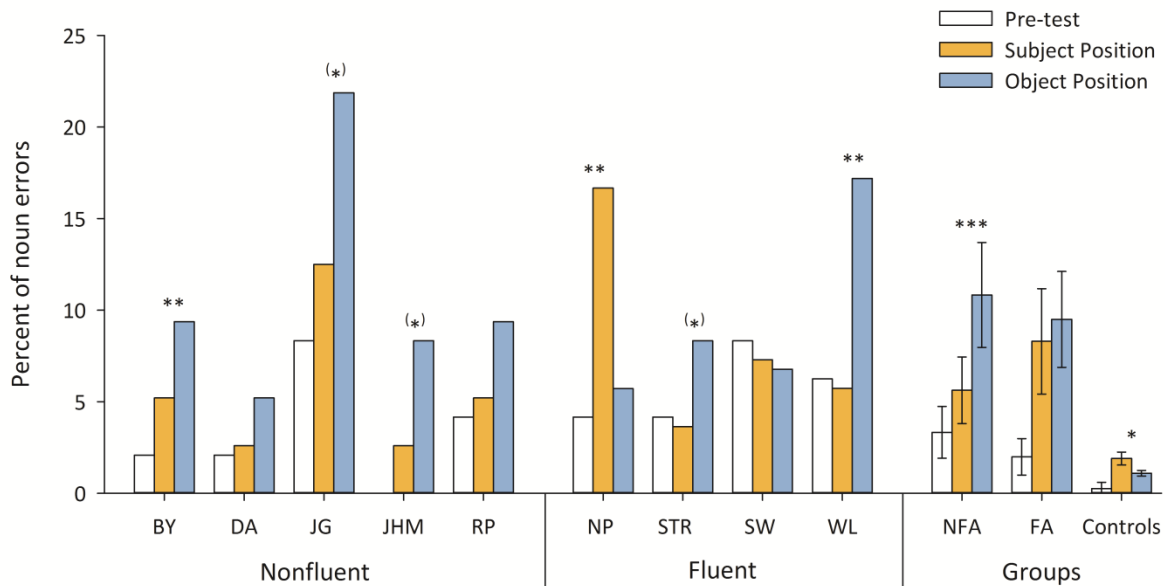


Table 6.1

*Percentage of errors for participants with aphasia and the three participant groups for the different experimental conditions*

	Overall Sentence Errors					Subject Errors					Object Errors				
	Relatedness		Frequency			Relatedness		Frequency			Relatedness		Frequency		
Participant	Rel	Un	high	low	Total	Rel	Un	high	low	Total	Rel	Un	high	low	Total
<b>Nonfluent</b>															
BY	23.8	16.7	13.5	26.0	<b>19.8</b>	6.0	4.6	3.1	7.3	<b>5.2</b>	11.9	7.4	5.2	13.5	<b>9.4</b>
DA	22.6	15.7	19.8	17.7	<b>18.8</b>	2.4	2.8	5.2	0.0	<b>2.6</b>	6.0	4.6	4.2	6.3	<b>5.2</b>
JG	58.3	37.0	50.0	42.7	<b>46.4</b>	16.7	9.3	12.5	12.	<b>12.5</b>	25.0	19.4	29.2	14.6	<b>21.9</b>
JHM	90.5	64.8	75.0	77.1	<b>76.0</b>	0.0	4.6	4.2	1.1	<b>2.6</b>	10.7	6.5	10.4	6.3	<b>8.3</b>
RP	67.9	49.1	51.0	63.5	<b>57.3</b>	7.1	3.7	6.3	4.2	<b>5.2</b>	10.7	8.3	6.3	12.5	<b>9.4</b>
Group mean	52.6	36.7	41.9	45.4	<b>43.6</b>	6.4	5.0	6.3	5.0	<b>5.6</b>	12.9	9.3	11.0	10.6	<b>10.8</b>
<b>Fluent</b>															
NP	23.8	27.8	27.1	25.0	<b>26.0</b>	14.3	18.5	17.7	15.	<b>16.7</b>	4.8	6.5	6.3	5.2	<b>5.7</b>
STR	22.6	32.4	25.0	31.3	<b>28.1</b>	2.4	4.6	4.2	3.1	<b>3.6</b>	7.1	9.3	6.3	10.4	<b>8.3</b>
SW	41.7	54.6	49.0	49.0	<b>49.0</b>	3.6	10.2	6.3	8.3	<b>7.3</b>	6.0	7.4	4.2	9.4	<b>6.8</b>
WL	13.1	35.2	26.0	25.0	<b>25.5</b>	4.8	6.5	2.1	9.4	<b>5.7</b>	7.1	25.0	20.8	13.5	<b>17.2</b>
Group mean	25.3	37.5	31.8	32.6	<b>32.2</b>	6.3	10.0	7.6	9.1	<b>8.3</b>	6.3	12.0	9.4	9.6	<b>9.5</b>
<b>Controls</b>															
Group mean	4.9	2.9	4.1	3.1	<b>3.8</b>	2.8	1.2	2.1	1.7	<b>1.9</b>	1.5	0.8	1.2	1.0	<b>1.1</b>

*Note.* Rel = semantically related; Un = semantically unrelated. Error percentages were rounded to the first decimal place.



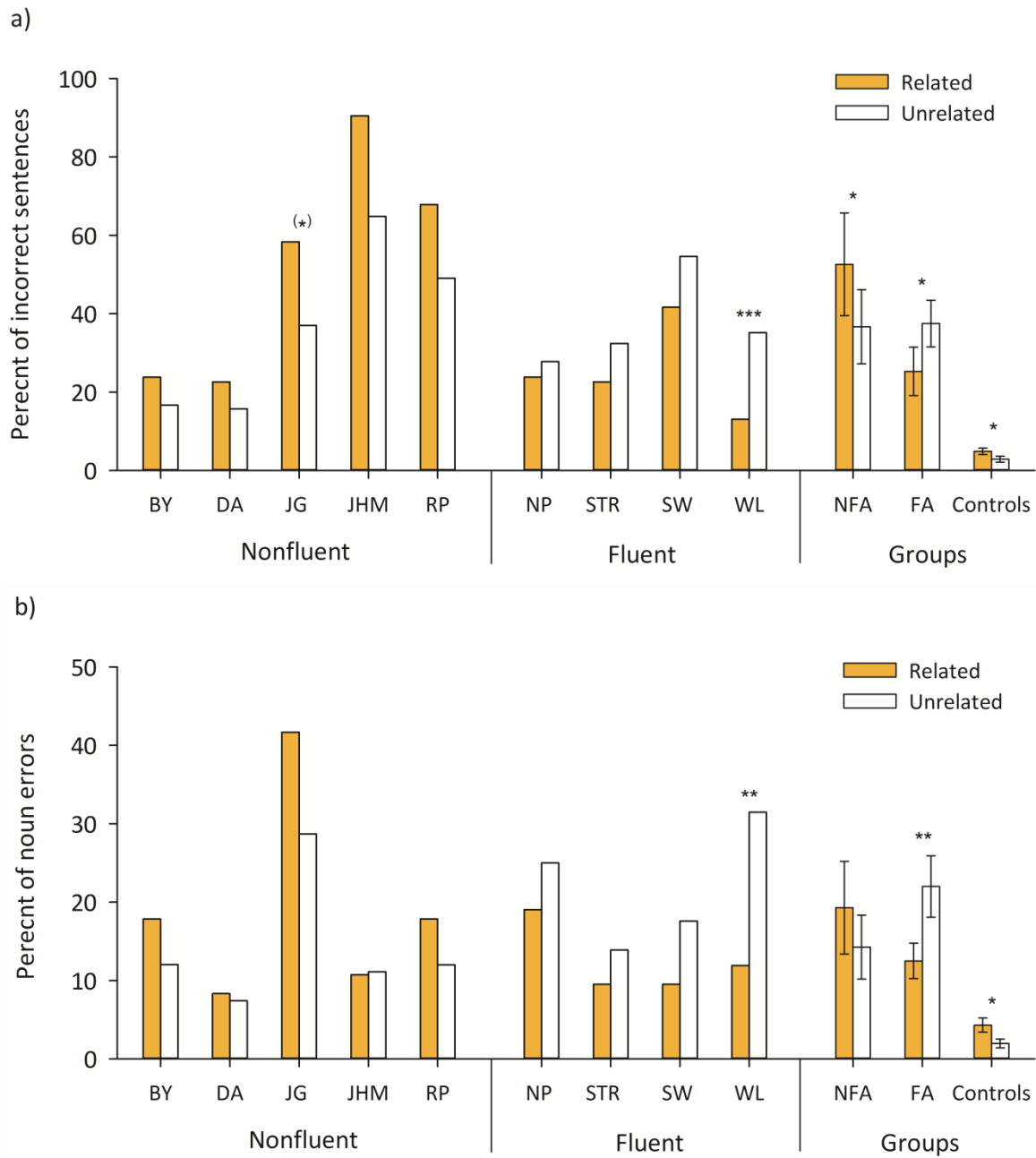
*Figure 6.1.* Percentage of nouns incorrectly produced in subject and object position for nonfluent and fluent aphasic individuals, as well as for the three participant groups. Also shown for comparison are the percentages of nouns incorrectly produced in isolation in the pre-test naming task (collapsed across all three administrations of the task). Bars represent one standard error above and below the mean for each participant group. (\*) =  $p < .10$ , \* =  $p < .05$ , \*\* =  $p < .01$ , \*\*\* =  $p < .001$ .

When analysed at the group level, the difference between subject noun and object noun errors was significant [ $\chi^2(1) = 16.15, p < .0001$ ]. This effect was also significant at the individual level for BY and trending for JG and JHM (Appendix D reports full details of individual subject analyses). No reliable position effect was found for the fluent group [ $p = .51$ ], and when considered individually, error patterns of fluent aphasic individuals were inconsistent. Individual analyses revealed significant effects of noun position for NP and WL, and a trend for STR.

The control participants showed a significant effect in the opposite direction to the nonfluent group (they made more errors in the subject than the object position) [ $\chi^2(1)=3.96$ ,  $p<.05$ ].

Analyses across groups revealed that the nonfluent participants produced significantly more errors than the controls [ $\chi^2(1)=55.89$ ,  $p<.0001$ ], and a significant group by error position interaction indicated that the nonfluent group was differently affected by noun position when compared to the control group [ $\chi^2(1)=13.57$ ,  $p<.0001$ ]. When directly comparing the two aphasia groups, there was no difference in error rates [ $p=.58$ ], but again, there was significant group by error position interaction [ $\chi^2(1)=5.05$ ,  $p<.05$ ], revealing that both groups' accuracy in producing the nouns was differently affected by the position of the nouns in the sentence. Further, the fluent group made significantly more errors when directly compared to the control group [ $\chi^2(1)=44.88$ ,  $p<.0001$ ], and a significant group by error position interaction indicated a more marked effect of noun position for the control group [ $\chi^2(1)=4.11$ ,  $p<.05$ ].

**Analysis of sentence production accuracy.** Turning now to the results that are directly relevant to our hypotheses, Panel a) of Figure 6.2 shows the overall percentage of sentence errors for each participant group and each aphasic individual, broken down by relatedness condition. The participants with nonfluent aphasia, considered as a group, showed a significant semantic interference effect, characterised by poorer accuracy on sentences containing related nouns than on sentences containing unrelated nouns [ $\chi^2(1)=4.69$ ,  $p<.05$ ]. In contrast, the fluent group exhibited a significant trend in the opposite direction, producing fewer errors on sentences containing related nouns than those containing unrelated ones [ $\chi^2(1)=5.75$ ,  $p<.05$ ]. The control group, although they made few errors overall, exhibited a significant trend in the same direction as the nonfluent group [ $\chi^2(1)=5.67$ ,  $p<.05$ ].



*Figure 6.2.* Percentage of overall sentence errors (Panel a) and combined errors on subject and object nouns (Panel b) as a function of relatedness for nonfluent and fluent aphasic individuals and the three participant groups. Bars represent one standard error above and below the mean for each participant group. NFA = nonfluent aphasic group; FA = fluent aphasic group. (\*) =  $p < .10$ , \* =  $p < .05$ , \*\* =  $p < .01$ , \*\*\* =  $p < .001$ .

A combined analysis of the data from the nonfluent and control groups revealed that, overall, the nonfluent group produced significantly fewer sentences correctly than the controls [ $\chi^2(1)=78.81, p<.0001$ ], but there was no significant group by relatedness interaction [ $p=.52$ ]. However, due to the low error rates for controls, logistic regression may not be sufficiently sensitive to reveal any genuine interactions. Consequently, the aggregated data for each group were reanalysed using the Q' test, and a significant group by relatedness interaction was found [ $Q'(1)=17.34, p<.0001$ ]. This suggests a more marked semantic interference effect for the nonfluent than for the control group. A direct comparison between the nonfluent and fluent groups revealed a main effect of group [ $\chi^2(1)=12.09, p<.0001$ ], indicating that the nonfluent group failed more sentences overall than the fluent group, and a significant group by relatedness interaction, but this time it was associated with a difference in the direction of the relatedness effect [facilitation vs. interference:  $\chi^2(1)=10.14, p<.01$ ]

Finally, a combined analysis of the fluent and control groups revealed that, overall, the fluent group produced fewer sentences correctly [ $\chi^2(1)=64.68, p<.0001$ ]. In this analysis, there was also a significant group by relatedness interaction, which demonstrated that the differential effect of relatedness across the two groups was statistically reliable [ $\chi^2(1)=9.10, p<.01$ ].

At the individual level, the effect of relatedness trended for nonfluent participant JG, and reached significance for fluent participant WL, while no effects were found for the remaining aphasic participants (see Appendix D for full statistical details of individual subject analyses).

Panel b) of Figure 6.2 shows the percentage of subject and object noun errors as a function of relatedness for each of the individual aphasic participants and the three participant groups. It can be seen from the figure that the differential semantic relatedness effects exhibited by the nonfluent and the fluent groups were also evident when noun accuracy was

considered on its own. The participants with nonfluent aphasia, as a group, showed poorer accuracy for producing the subject or the object noun for sentences containing meaning-related nouns than for sentences containing unrelated nouns; however, this difference failed to reach significance [ $p=.20$ ]. Conversely, the fluent participants showed a trend in the opposite direction, exhibiting more errors on the subject and object nouns when they were *not* related in meaning, a difference that was significant at the group level [ $\chi^2(1)=9.31, p<.01$ ]. The control participants displayed a significant effect in the same direction as the nonfluent group [ $\chi^2(1)=6.36, p<.05$ ].

A direct comparison of nonfluent participants with controls revealed poorer noun production accuracy for the nonfluent group [ $\chi^2(1)=33.03, p<.0001$ ]. Further, there was no significant group by relatedness interaction for the nonfluent participants and the controls [ $p=.17$ ], which was further confirmed when analysing aggregated error scores with the Q' test [ $p=.31$ ].

When directly comparing the fluent and nonfluent groups, no effect of group was found [ $p=.75$ ], indicating that there was no difference in the accuracy for producing the subject and the object nouns. However, there was a significant group by relatedness interaction, which demonstrated that the relatedness had a reliably different effect across the two groups [facilitation vs. interference:  $\chi^2(1)=9.96, p<.01$ ]. Finally, a direct comparison of fluent participants with controls revealed that the fluent group made more noun errors than the controls [ $\chi^2(1)=42.61, p<.0001$ ], and a significant group by relatedness interaction indicated that the groups were differently affected by the relatedness manipulation [facilitation vs. interference:  $\chi^2(1)=13.61, p<.001$ ].

At the individual level, only fluent participant WL showed a significant relatedness effect on noun production accuracy, indicating semantic facilitation (Appendix D reports full details of the individual analyses).

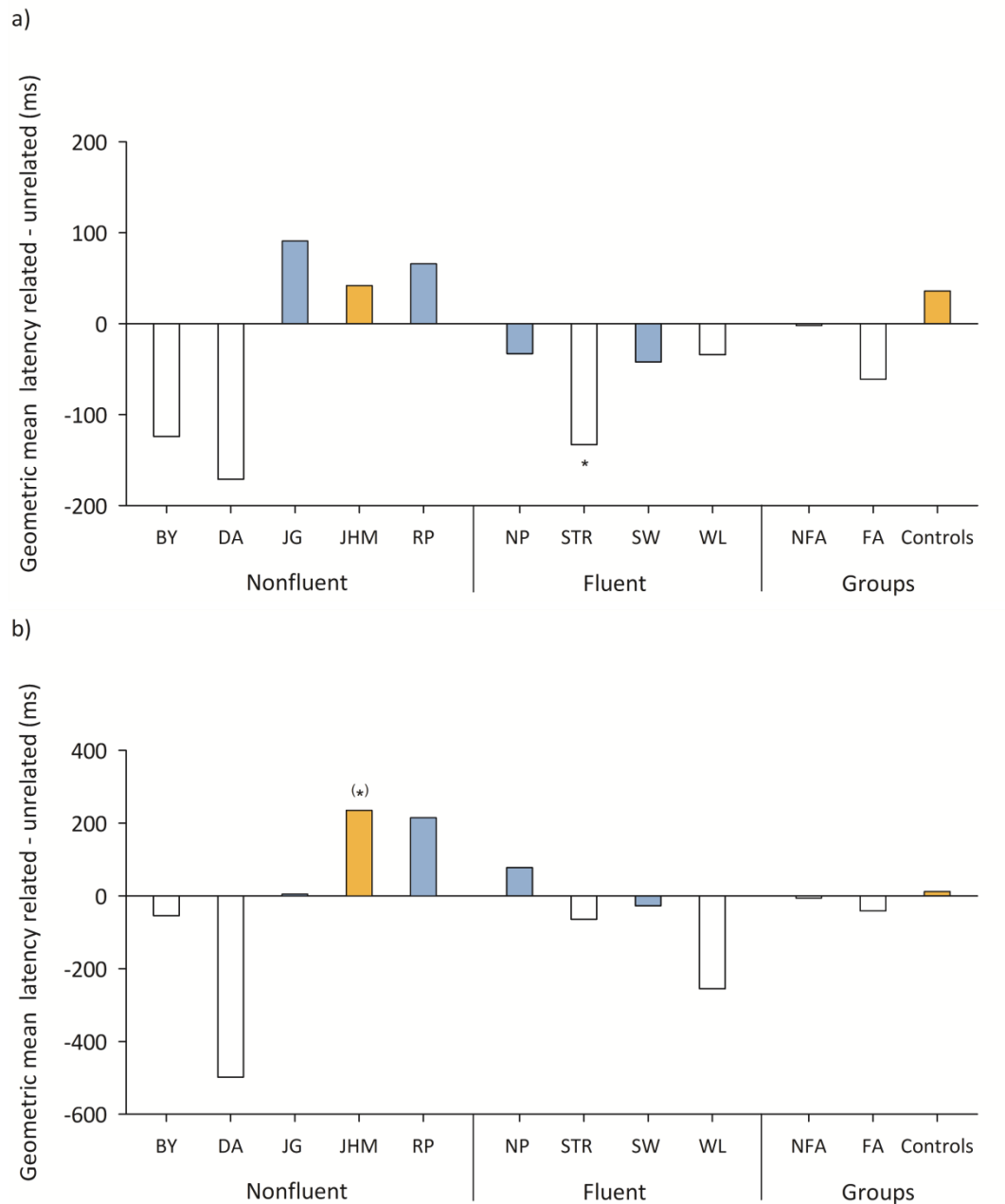
**Analysis of response latency.** Latency data were trimmed of outliers in the same way as for Experiment 1. When collapsed across all individuals, this led to a removal of an average of 2.4% of data for the aphasic groups, and 1.0% for the control group<sup>9</sup>. Appendix D reports geometric mean latencies for all latency measures and experimental conditions for each aphasic individual and the three participant groups. As in Experiment 1, results for subject-noun initiation times were very similar to results obtained from the analysis of sentence initiation times, and are therefore reported in the appendix (Appendix D reports full statistical details of the subject-noun initiation time analysis for aphasic individuals and the three participant groups).

**Sentence initiation times.** Panel a) of Figure 6.3 shows the difference in sentence initiation times for each group and each individual aphasic participant as a function of relatedness. The nonfluent participants, as a group, did not show a main effect of relatedness on sentence initiation times [ $p=.91$ ]. For the fluent participants as a group and the controls, relatedness had also no significant effect on sentence initiation times [ $p=.85$  and  $p=.42$ ].

When compared across groups, the nonfluent participants were slower to initiate the sentence than the control participants [ $F(1, 2081)=339.79, p<.0001$ ], but there was no significant group by relatedness interaction [ $p=.87$ ]. A direct comparison of the data for the nonfluent and the fluent groups revealed that there was no difference in their sentence initiation times [ $p=.44$ ], and no significant group by relatedness interaction [ $p=.49$ ]. Finally, the fluent participants showed slower sentence initiation times than the controls [ $F(1,1851)=126.30, p<.0001$ ], and in this case, there was a group by relatedness interaction, [ $F(1, 1851)=6.22, p<.05$ ]. However, this interaction fell out when latencies were transformed into z-scores [ $p=.06$ ], using the same procedures as outlined in Experiment 1.

---

<sup>9</sup> Due to technical problems and other disturbances, an additional 2.2% of data points (across all groups) were lost.



*Figure 6.3.* Noun relatedness effects on sentence initiation times (Panel a) and verb-object noun initiation times (Panel b). A positive value indicates a faster response time for sentences containing unrelated nouns. Blue bars represent whole sentence planners as classified in Experiment 1, and orange bars represent phrase-by-phrase planners who showed evidence of some advance planning. White bars represent phrase-by-phrase planners. NFA = nonfluent aphasic group; FA = fluent aphasic group. (\*) =  $p < .10$ , \* =  $p < .05$ .



At the individual level, fluent participant STR was significantly faster to initiate sentences when they contained meaning-related nouns. There were no main effects of relatedness for any of the other aphasic individuals (see Appendix D for full individual subject analyses). However, when examined on the descriptive level, it stands out that the nonfluent participants that were classified as whole sentence planners in Experiment 1 (JG and RP) showed consistently slower initiation times for sentences that contain meaning-related nouns (see Figure 6.3). The same pattern was found for nonfluent participant JHM and the controls, both of which were classified as phrase-by-phrase planners but indicated some degree of parallel subject and object noun retrieval. None of the fluent participants, not even those classified as whole sentence planners, showed this trend. However, these differences only existed on the descriptive level and were not backed up by inferential statistics.

**Object noun initiation times.** Panel b) of Figure 6.3 shows average latencies from the offset of the verb response to the onset of the object name, as a function of relatedness effects for each aphasic individual and the three participant groups. Results were similar to sentence initiation times. Again, no relatedness effects were found for any of the participant groups [nonfluents:  $p=.95$ ; fluents:  $p=.66$ ; controls:  $p=.55$ ].

When compared across groups, the nonfluent participants were slower to initiate the object noun than the controls [ $F(1, 2034)=2297.91, p<.0001$ ], but there was no significant group by relatedness interaction [ $p=.89$ ]. A direct comparison of the nonfluent and fluent groups revealed that there was no significant group effect [ $p=.15$ ], indicating that there was no difference in their verb-object noun initiation times, and there was no significant group by relatedness interaction [ $p=.52$ ]. Finally, when data of the fluent group was directly compared to that of the control group, there was a significant group effect [ $F(1, 1810)=64.23, p<.0001$ ],

indicating that the fluent participants were slower to produce the object noun than the controls, but again, there was no significant group by relatedness interaction [ $p=.52$ ].

Considering the individual level, no relatedness effects were found for any of the aphasic participants, however, there was a trend for nonfluent participant JHM, revealing that she took longer to produce the sentence object when the nouns in the sentence were related in meaning (see Appendix D for full individual subject analyses). Furthermore, as was found for sentence initiation and subject-noun initiation times, JG, RP, JHM and the controls showed consistently slower verb-object noun initiation times for sentences that contained meaning-related nouns. This time, also one of the fluent participants showed this trend: NP, who was classified as a whole sentence planner in Experiment 1.

**Comment.** Overall, the latency analyses yielded few statistically reliable results. However, when examined on the descriptive level, it stands out that the *nonfluent* participants that were classified as whole sentence planners in Experiment 1 (JG and RP) as well as the less incremental phrase-by-phrase planners (JHM and control participants) all showed consistently slower sentence initiation times for sentences containing meaning-related nouns. Nonfluent participants BY and DA, who were previously classified as phrase-by-phrase planners, did not show this pattern. This finding makes some intuitive sense, given that the semantic similarity between the object noun and its subject only has the opportunity to delay sentence initiation if the participant is planning the whole sentence significantly in advance. These suggestions are made very tentatively, however, as the described differences existed at the descriptive level only.

## Discussion

In this experiment, we manipulated the semantic relatedness of the subject and object nouns to be produced in SVO sentences. When we considered participants' accuracy in producing the key nouns themselves, irrespective of the relatedness condition, the results replicated those from Experiment 1: Again, the nonfluent participants were considerably more accurate at producing nouns in isolation than in the context of a sentence, and within the sentence itself they were more accurate at producing nouns in subject position than those in object position. Again, these effects were consistently observed only in individuals with nonfluent aphasia, while the performance of the control participants and the participants with fluent aphasia varied.

The primary hypothesis that individuals with nonfluent aphasia would be less accurate at producing sentences containing semantically related nouns than those containing unrelated nouns, was supported. Importantly, this effect was specific to individuals with nonfluent aphasia: The participants with fluent aphasia actually showed a trend in the opposite direction. Also, the control participants showed an effect in the same direction as the nonfluent participants, which is in line with previous research (e.g., Freedman et al., 2004). However, the relatedness effect was much more pronounced in the nonfluent group. This difference was not statistically reliable in a logistic regression analysis, but it did reach significance in an analysis of aggregated error scores, which is arguably a more appropriate method to use given the extremely low rate of errors in controls.<sup>10</sup>

---

<sup>10</sup> The logistic regression analysis may not have shown a significant group by relatedness interaction when comparing the controls and the nonfluent group because of low error rates in the control data. Further, both groups displayed a relatedness effect in the same direction, which makes it more difficult to find a statistically reliable interaction than when the effect across groups is diverging.

As mentioned previously, these findings can be incorporated into Martin and colleagues' framework: Inefficient binding of content elements to their respective "slots" in the semantic short-term memory buffer system could result in a failure to produce the sentence correctly. More specifically, the various elements of the utterances might compete more intensely for selection to a given position than they would if they were more efficiently bound to their appropriate positions, again making errors more likely to occur (see also Martin & Freedman, 2001; Martin & He, 2004; Martin et al., 1999). Although this framework deals only with planning within a single noun phrase, it could in principle be extended to account for planning over more complex syntactic structures.

Curiously, although semantic relatedness had a robust effect on our nonfluent participants' overall sentence accuracy, their sentence initiation times were not reliably influenced by this variable. Of course, an effect of semantic relatedness on early planning would only be expected if the speaker plans the object noun substantially in advance. In Experiment 1, only three of the five nonfluent participants showed any evidence of advance planning of this type: JG and RP were the only ones to show any suggestions of an object frequency effect on their sentence and/or subject initiation times, while JHM showed more marked effects of object frequency on verb-object noun initiation time when the subject noun was of high frequency, indicating a temporal overlap of subject and object noun retrieval. Interestingly, in this experiment, these three nonfluent participants also showed a trend towards slower initiation times for sentences containing semantically related nouns than for those containing unrelated ones. Taken together, these findings indicate that semantic relatedness slowed down initiation times only in instances where the speaker planned both subject and object nouns in advance (at least to some extent). Again, comparison between the nonfluent and the other participant groups revealed a similar pattern to that shown for the other effects: The control participants showed a trend in the same direction as the nonfluents,

except of a much smaller magnitude, and the fluents showed no evidence of any detrimental effect of relatedness, regardless of whether or not they were advance planners.

It should be emphasised that the semantic interference effects observed in the nonfluent participants involved different noun phrases. Previous studies of individuals with nonfluent aphasia have obtained semantic interference effects only when the related elements occurred within the same noun phrase (e.g., *nose and ear*; Freedman et al., 2004). Indeed, Martin and colleagues drew directly on such findings to support an incremental phrase-by-phrase model of sentence planning, in which speakers lexically encode only one single phrase at a time (see esp. Martin & Freedman, 2001; Martin & He, 2004). Two aspects of the current findings are relevant for reconciling these apparently inconsistent findings. The first concerns the directionality of the effects observed here: Evidence from our latency analyses indicates that much of the interference effect might reflect a perseveratory rather than an anticipatory interference of relatedness – that is, elements early in the sentence appeared to interfere with those later in the sentence to a greater extent than vice versa. If this is the case, then interference effects might plausibly extend beyond the “window” in which advance planning occurs. In fact, recent research has shown that semantic interference effects occur across phrases in normal speakers (Smith & Wheeldon, 2004). Second, it is very possible that the scope of planning varies according to current task demands. In our study, advance planning of the entire sentence was actively encouraged by removing the stimulus picture from view at the time of response initiation. Further, the noun phrases in our target sentences were extremely simple: They consisted only of a single noun plus determiner. These would be just the kinds of conditions which might engender advance planning across more than one phrase at a time.

These findings are also compatible with sentence production models which postulate that a robust structural representation of the target utterance helps to manage the naturally

competitive influence of one word on another when both are planned for the same utterance (Dell et al., 2008; Gordon & Dell, 2003; Stemberger, 1985). If the syntactic sequencing mechanism operates inefficiently, the content elements are freer to compete with one another for selection to a particular position in the sentence, particularly if they obey the same selection constraints. For example, in the DoL model (Gordon & Dell, 2003) several of the key characteristics of agrammatic speech were simulated by reducing the input activation of the syntactic sequencing mechanism. Further, a syntactically lesioned model was also successful to simulate a higher production accuracy for single words when compared to producing the same words in the context of a sentence. This pattern has been observed with nonfluent aphasic speech production in previous research (e.g., Schwartz & Hodgson, 2002; Williams & Canter, 1982) as well as in the current study. Gordon and Dell (2003) argue that the higher output accuracy of the DoL model for single word production reflects the missing aggravated syntagmatic interference that occurs during sentence production, when multiple words are planned for the same utterance but syntactic input activation is insufficient to activate the right word at the right time. In contrast to simulating agrammatic speech, the pattern of fluent but lexically empty speech was simulated by reducing the efficiency of conceptual-semantic activation of sentence elements. Therefore, individuals with fluent aphasia would be expected to show a better performance for sentences containing meaning-related nouns, as both nouns activate each other via shared semantic feature nodes and are so more likely to reach sufficient activation for production. This prediction was supported by results from the current experiment. Both conventional structure-based accounts of nonfluent aphasic sentence production and Kolk's framework (1995, 2006) fail to provide a plausible explanation for any of these findings.

Finally, the results from Experiment 2 provide further evidence against an account of nonfluent aphasic performance in terms of rapidly decaying conceptual representations.

Clearly, in order to obtain semantic relatedness effects, there must be substantial activation of both relevant conceptual representations in the first place. Further, most accounts require that the relevant lexical representations have also received substantial activation (Bloem & La Heij, 2003; Bloem, van den Boogaard, & La Heij, 2004; Levelt et al., 1999; Rahman & Melinger, 2009, 2011; Roelofs, 1992). Put simply, the semantic interference effects themselves provide further evidence that the difficulty in nonfluent aphasia is not due to abnormally fast decay of the object's conceptual representation, but rather to effects occurring during linguistic specification of the message.

### **Lesion Analysis**

The aim of the current lesion analysis was to investigate the relationship between the behavioural phenomena observed in Experiments 1 and 2 and lesion localisation. In Experiment 1 and 2, we found that nonfluent participants were particularly susceptible to the frequency of elements appearing early in the sentence (Experiment 1), and to the semantic relationship of the constituent nouns (Experiment 2); fluent participants did not show these effects. We argued that the cognitive impairments underlying these findings in nonfluent aphasia arise from weak structure building processes. This results in an abnormal reliance on lexical elements to generate a sentence, and also gives rise to difficulties managing competition between lexical candidates planned for the same utterance.

As mentioned previously, lesions in nonfluent aphasia generally encompass the anterior left hemisphere language regions, including most commonly the posterior portion of the LIFG (BA44/Broca's area) and the insula (Goodglass et al., 2001). Damage to the insula has been specifically linked to articulatory problems (Dronkers, 1996). The role of the LIFG is less well understood, but some studies have demonstrated a relationship between damage to this region and poor comprehension in sentence-picture matching tasks for syntactically reversible sentences (BA 45/47; Baldo & Dronkers, 2007) or those which required the use of syntactic structure to interpret their meanings (BA 47; Dronkers, Wilkins, Van Valin Jr., Redfern, & Jaeger, 2004).

The exact relationship between the LIFG and sentence production difficulties is still unclear, as few studies have explored this question. Nevertheless, there is considerable evidence that the LIFG in general, and BA 44 in particular, is crucial for selecting the appropriate lexical element among other highly activated competitors (e.g., Schnur et al., 2005; Schnur et al., 2009), a proposal that directly relates to the finding that nonfluent



participants were highly susceptible to the manipulation of semantic relatedness of nouns in Experiment 2.

Further insights come from recent research that has suggested that damage to BA 44/6 (the interface of Broca's area to the left premotor cortex) is associated with difficulty sequencing lexical elements in multiword utterances (Thothathiri, Gagliardi, & Schwartz, 2012; Thothathiri et al., 2010). Thothathiri and colleagues (2010) found that two individuals with damage to BA 44/6 performed poorly in a two-picture naming task when asked to sequence lexical elements into a conjoined noun phrase (e.g., *the eye and the pencil*). This difficulty was even more exaggerated when it was not possible to rely on lexical variables to order the sentence elements (the two individuals made more errors when the phrase began with an inanimate noun, such as *football*, compared to when it began with an animate noun, such as *monkey*, which violates the usual transitive pattern in English of placing an animate agent before an inanimate patient, e.g., *the monkey kicked the football*). Taken together, a difficulty sequencing lexical elements may be genuinely associated with difficulty in sentence structure generation, which we propose to be the primary cause of sentence production failure in nonfluent aphasia.

In order to systematically investigate whether the performance patterns observed in nonfluent participants in Experiments 1 and 2 were consistently associated with damage to particular cortical regions, we analysed structural MRI scans of all participants who were willing to undergo the procedure (all except participant RP). We then created lesion overlay maps and conducted a simple voxel-based lesion symptom mapping analysis (VLSM). Two main limitations to this analysis need to be noted in advance: First, potential findings using VLSM need to be interpreted with caution, because of the small sample size. Second, individuals with aphasia were initially selected based on a relatively good naming performance, because we wanted to investigate sentence production difficulties independent

of naming difficulties. This selection criterion undoubtedly affected the distribution of damage that was seen in the analysis. Nevertheless, we were confident that the lesion analysis would provide valuable information about brain regions that corresponded to the performance deficits observed in the preceding behavioural experiments.

Based on previous research, we tentatively predicted that damage to the LIFG (and in particular to the junction between BA 44 and BA 6) would be positively associated with performance patterns observed in the nonfluent participants in Experiments 1 and 2. However, due to the scarcity of lesion studies investigating neuroanatomical correlates of components specifically involved in sentence production difficulties in nonfluent aphasia, we adopted a mainly exploratory approach to the lesion analysis.

### **Method**

The Capital & Coast District Health Board of Wellington, New Zealand, gave ethical approval for this research in 2013.

#### **Participants**

The participant group in this study consisted of all the individuals with aphasia who had taken part in Experiment 1 and 2, except for nonfluent participant RP. He had experienced anxiety brought on by confinement in the MRI scanner during previous scans and therefore decided not to participate.

#### **Imaging Methods**

**Image acquisition.** All participants underwent a magnetic resonance imaging scan specifically for this study, within one year after being tested in both experiments. Whole brain T1-weighted as well as T2-weighted structural scans with inversion recovery (FLAIR) were collected at 1.5 Tesla (T1 3D FFE: TR = 25 ms, TE = 4.6 ms, FOV = 252mm x 238mm, slice thickness = 1mm; Sagittal 3D FLAIR: TR = 4800ms, TE = 329ms, TI = 1660ms, FOV = 252mm x 250mm, slice thickness = 1mm). While T1-weighted images provide high

structural definition, T2-weighted FLAIR scans are particularly useful for identifying older infarctions and scarred tissue.

**Image preparation.** For each individual, the lesions were manually drawn onto the participant's own T1-weighted structural image using MRICron (Rorden, Karnath, & Bonilha, 2007; <http://www.mccauslandcenter.sc.edu/mricro/mricron/index.html>), while consulting the T2-weighted FLAIR image as additional guidance. Subsequently, the scans and the lesions were spatially normalised using SPM8 (Ashburner et al., 2008; see also <http://www.fil.ion.ucl.ac.uk/spm/>), implemented in MatLab 8.0 (The Mathworks Inc, Natick, MA). The lesions were then overlaid onto a standard template based on healthy elderly individuals (with a mean age of 65 years; Rorden et al., 2012) to roughly account for the age group of the participants.

The exact locations of lesions were determined using the SPM Anatomy Toolbox (Eickhoff et al., 2005). The SPM Anatomy Toolbox is based on three-dimensional cytoarchitectonic probabilistic maps that were created analysing ten human brains post-mortem (e.g., Amunts et al., 1999). However, it should be noted that the SPM Anatomy Toolbox has not yet been completed because not all cortical regions have been mapped yet. Therefore, damage that involved brain regions that were not accounted for so far in the SPM Anatomy Toolbox was determined using the Brodmann Atlas as implemented in MRICron (Rorden et al., 2007).

## **Image Analysis**

**Lesion overlap analysis.** First, to identify brain regions that were commonly damaged in individuals within each aphasia group, MRICron (Rorden et al., 2007) was used to create lesion overlay maps for both the nonfluent and fluent groups. It is important to note that the regions of maximum overlap may not necessarily reflect the functional roles of the relevant brain areas, but rather their susceptibility to vascular damage. Nevertheless, this

method is useful to investigate the distribution of brain damage within a group of individuals suffering from the same pathology, which was the main goal here.

**Creation of behavioural measures for VLSM analysis.** In order to examine the relationships between lesion location and our key behavioural phenomena, we first created a continuous measure of each behaviour. A measure of the *subject noun frequency effect* (Experiment 1) was calculated by subtracting the overall percentage of errors in sentences commencing with a low frequency subject from the percentage of errors in sentences commencing with a high frequency subject. The algebraic sign of the score indicated the direction of the subject frequency effect (*negative* scores indicated higher error rates on sentences with a *low* frequency subject, *positive* scores indicated higher error rates on sentences with a *high* frequency subject), and the absolute value of the score indicated the magnitude of the subject noun frequency effect. A measure for the *noun relatedness effect* (Experiment 2) was calculated by subtracting the percentage of errors on the semantically related sentences from the percentage of errors on the unrelated sentences. Again, the direction of the effect was indicated by the algebraic sign of the score (*negative* scores indicated higher error rates for sentences containing meaning-related nouns, *positive* scores indicated higher error rates for sentences containing meaning-*unrelated* nouns), and the absolute value of the score reflected the magnitude of the noun relatedness effect.

**VLSM analysis.** Using the behavioural measures described above, we then performed a VLSM analysis to investigate whether participants' scores on each measure were predicted by damage to particular brain voxel(s). VLSM analyses the strength of contribution of each voxel (the smallest element that can be defined in a three-dimensional space) to a specific behaviour of interest, and identifies voxels that are significantly associated with the measured behaviour (Baldo, Wilson, & Dronkers, 2012). For each voxel, a statistical test determines whether damage to this particular voxel is associated with participants' performance in the

behavioural data (Bates et al., 2003). The VLSM method has a significant advantage over simple overlap analysis, because each voxel must *reliably predict performance* on the key measure in order to be considered implicated. This reduces the possibility that the results of the analysis will be skewed by differences in the susceptibility of different brain regions to stroke. Also, since the behavioural measures are continuous rather than a binary grouping, inter-individual differences in the magnitude of the behaviour of interest will be maintained, and statistical power maximised (Baldo et al., 2012). In the current VLSM analysis, we used the non-parametric Brunner-Munzel (BM) rank order test (Rorden et al., 2007) provided in NPM (implemented in MRICron; <http://www.mccauslandcenter.sc.edu/mricron/npm/>).

Further, the analysis was set to discount any voxels damaged only in a single participant.

When using VLSM, there is a general risk of increased Type 1 errors, since this method will conduct as many statistical tests as there are voxels in the image. Several studies have recommended the use of False Discovery Rate (FDR) to correct for multiple comparisons over the use of the very conservative Bonferroni correction (e.g., Rorden & Karnath, 2004), because it dramatically increases the probability for Type II errors (Kimberg, Coslett, & Schwartz, 2007). However, FDR does not control for familywise error rates, so it may be advisable to consider both types of corrections. Also, problems in computing accurate  $Z$  scores and  $p$  values for small samples have been reported for the Brunner Munzel rank order test (Medina, Kimberg, Chatterjee, & Coslett, 2010). In order to account for this, we followed the recommendations of Medina and colleagues (2010) and used permutation-generated test scores for the Brunner-Munzel test (and a permutation threshold of 1000). Permutation testing involves the random assignment of each participant's score in the behavioural measures across each individual voxel. Critical values are calculated for each permutation, "thereby creating a simulation of how often extreme observations would be generated under the null hypothesis of no association between lesions site and behavioural

score” (Baldo et al., 2012, p.7). These iterations ensure that accurate Z scores are created, even when sample sizes are small and the distribution of data is skewed (Kimberg et al., 2007).

One danger in performing VLSM analyses across a sample as small as this one is lack of statistical power. Therefore, we also present a statistical power map indicating the probability of detecting damage to voxels in different brain regions (the map corresponds to Rudrauf et al.’s (2008) “effective coverage map”, which takes the spatial distribution of lesions in the sample into account, however the map presented here is based on the Wilcoxon-Mann-Whitney probability due to the continuous nature of the behavioural data).

Finally, for both the overlap images and the VLSM, respectively, the resulting lesion maps were examined using the SPM Anatomy Toolbox (Eickhoff et al., 2005, procedure outlined above)<sup>11</sup> and were further explored using the Brodmann Atlas as implemented in MRICron (Rorden et al., 2007).

## Results

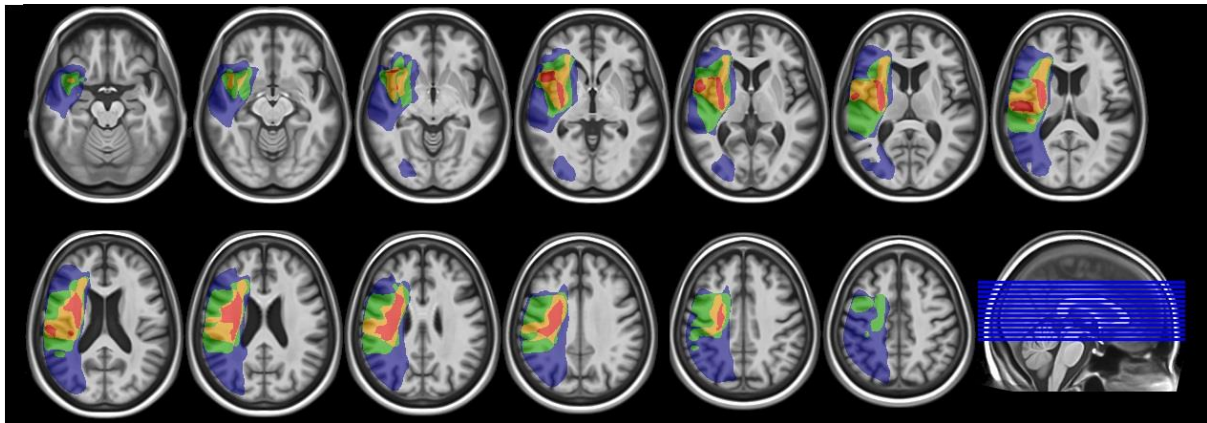
### Lesion Overlap Analysis

Figure 7.1 displays lesion overlay maps for individuals with nonfluent aphasia, and Figure 7.2 illustrates overlay maps for fluent aphasic individuals. Except for fluent participant STR, all participants in this analysis had lesions exclusively to the left hemisphere. When comparing overlay maps of nonfluent and fluent individuals, lesions associated with nonfluent aphasia were much more anterior than lesions associated with fluent aphasia, which

---

<sup>11</sup> Probabilistic maps (and publication describing these) relevant to the current analysis included: Broca’s area (BA 44, BA 45; Amunts et al., 1999); Motor cortex (BA 4, BA 6; Geyer, 2003; Geyer et al., 1996); Somatosensory cortex (BA 1, BA 2, BA 3; Geyer, Schleicher, & Zilles, 1999; Geyer, Schormann, Mohlberg, & Zilles, 2000; Grefkes, Geyer, Schormann, Roland, & Zilles, 2001); Parietal Operculum (Eickhoff, Amunts, Mohlberg, & Zilles, 2006; Eickhoff, Schleicher, Zilles, & Amunts, 2006); Parietal cortex (Caspers et al., 2006; Scheperjans, Eickhoff, et al., 2008; Scheperjans, Hermann, et al., 2008); and Insula (Kurth et al., 2010).

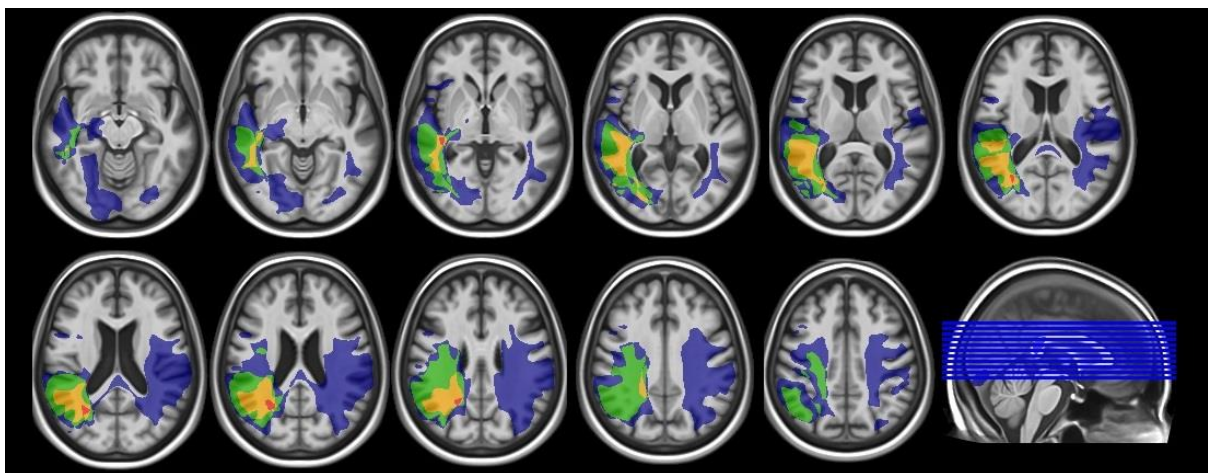
showed more posterior damage. For the individuals with nonfluent aphasia (Figure 7.1), the region of maximum (100%) overlap extended from the LIFG (including BA 44) and the left insula to the left parietal operculum, and a small portion of the left inferior parietal cortex. When also considering regions of 75% lesion overlap, the lesion map also included parts of the left premotor cortex (BA 6). Further analysis using the Brodmann Atlas in MRIcron revealed additional damage to the left posterior pars orbitalis in the LIFG (BA 47), and the left anterior tip of the auditory cortex (BA 42).



*Figure 7.1.* Lesion overlay maps for individuals with nonfluent aphasia showing axial slices on a standard template (Rorden et al., 2012) at MNI Z coordinates = -12, -7, -2, 3, 8, 13, 18, 23, 28, 33, 38. Red = overlap between all individuals; Green = overlap between three individuals; Yellow = overlap between two individuals; Blue = no overlap, lesion is specific to only one individual.

For the fluent aphasia group (Figure 7.2), the overlap analysis revealed only a small region of maximum overlap. This region could not be explored using the SPM Anatomy Toolbox (Eickhoff et al., 2005) because the majority of temporal areas has not yet been described with probabilistic maps. An informal analysis of the maximum overlap map using the Brodmann Atlas in MRIcron identified small portions of the left fusiform gyrus (BA 37),

the left inferior temporal gyrus (BA 20), and curiously, secondary and associative left visual cortex areas (V2 and V3). When considering regions of 75% lesion overlap, damage further included parts of the left superior temporal gyrus (BA 22), the left middle temporal gyrus (BA 21), the left angular gyrus (BA 39) as well as small parts of the left auditory cortex (BA 41 and BA 42).



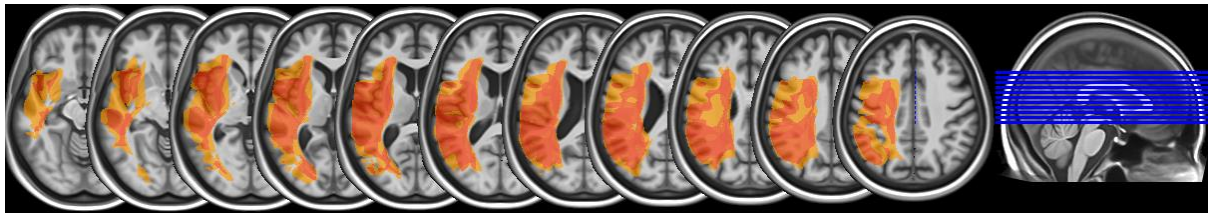
*Figure 7.2.* Lesion overlay maps for individuals with fluent aphasia showing axial slices on a standard template (Rorden et al., 2012) at MNI Z coordinates = -7, -2, 3, 8, 13, 18, 23, 28, 33. Red = overlap between all individuals; Green = overlap between three individuals; Yellow = overlap between two individuals; Blue = no overlap, lesion is specific to only one individual.

### Voxel-Based Lesion-Symptom Mapping

**Power map.** Figure 7.3 shows the power map for the VLSM analyses. Note that only one power map is reported, since both VLSM analyses were performed on the same set of MRI scans, and therefore the probability to detect damage to voxels across different brain regions was the same. As can be seen from Figure 7.3, statistical power was generally low across different brain regions and not sufficient to reliably detect a significant relationship between lesion location and *behavioural measures* when correcting for multiple comparisons.



Power was particularly low for very anterior, posterior, inferior and superior regions, as well as the whole right hemisphere, due to the lesion distribution in the sample. Therefore, non-significant results should be interpreted with caution. Further, significant results may evolve due to the idiosyncratic nature of the lesions submitted to the analysis.



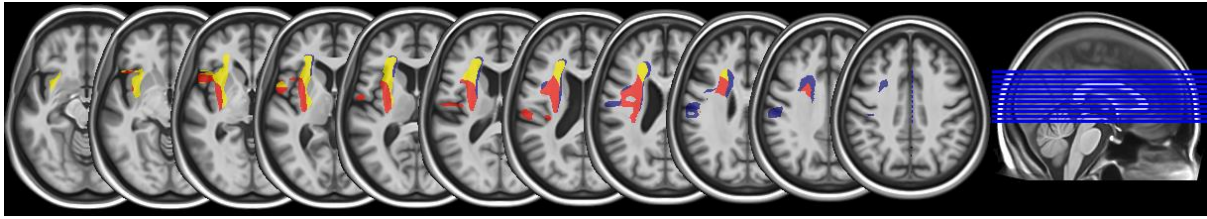
*Figure 7.3.* Representative axial slices showing the distribution of statistical power on a standard template (Rorden et al., 2012) at MNI Z coordinates: -7, -2, 3, 8, 13, 18, 23, 28, 33, 38, 43. The colour scheme indicates statistical power to find effects significant to different  $p$  values and types of multiple corrections. Red = sufficient power to detect effects at an uncorrected  $p < .01$  level. Orange = sufficient power to detect effects at an uncorrected  $p < .05$  level. There was not sufficient power in any regions to detect effects when accounting for multiple comparisons (using FDR or Bonferroni correction).

**Subject noun frequency effect.** Figure 7.4 shows the VLSM map of the voxels significantly associated with marked subject frequency effects. The map encompasses a small portion of the LIFG (BA 44 and BA 45), the insula as well as the inferior parietal lobe and the parietal operculum. Further analysis with the Brodmann Atlas in MRICron revealed additional damage to the posterior pars orbitalis in the LIFG (BA 47) and the anterior tip of the auditory cortex (BA 42).



*Figure 7.4.* Axial slices from VLSM maps computed for total subject noun frequency effect displayed on a standard template (Rorden et al., 2012) at MNI X coordinates = -7, -2, 3, 8, 13, 18, 23, 28, 33, 38, 43. The map shows regions that are associated with poorer performance in sentences containing a low frequency subject. The colour scheme indicates BM Z-scores: Red = regions significant at  $p < .05$  of the permutation generated threshold ( $Z = 3.89$ ); Yellow = regions significant at the uncorrected  $p < .01$  level; Blue = regions significant at the uncorrected  $p < .05$  level.

**Noun relatedness effect.** The VLSM map for the noun relatedness effect is shown in Figure 7.5. The map was almost identical to that obtained for the subject noun frequency effect: Left hemisphere regions affected included the LIFG (BA 44 and BA 45) and the insula, as well as the left inferior parietal lobe and the parietal operculum. When further explored using the Brodmann Atlas in MRICron, additional damage was found in the posterior pars orbitalis in the LIFG (BA 47) and the anterior part of the auditory cortex (BA 42).



*Figure 7.5.* Axial slices from VLSM maps computed for total noun relatedness effect displayed on a standard template (Rorden et al., 2012) at MNI X coordinates = -12, -7, -2, 3, 8, 13, 18, 23, 28, 33, 38. The map shows regions that are associated with poorer performance in sentences containing meaning-related nouns. The colour scheme indicates BM Z-scores: Red = regions significant at  $p < .05$  of the permutation generated threshold ( $Z = 3.89$ ); Yellow = regions significant at the uncorrected  $p < .01$  level; Blue = regions significant at the uncorrected  $p < .05$  level.

## Discussion

The lesion overlay maps for the two different aphasia groups were broadly in line with previous literature: Lesions in the nonfluent group generally encompassed the anterior left hemisphere, while those in the fluent group encompassed mainly posterior regions of the left hemisphere (e.g., Goodglass et al., 2001; Goodglass, 1993). Importantly, regions of maximum overlap in the nonfluent participants were entirely spared in the fluent group, indicating that the group-specific performance observed in Experiment 1 and Experiment 2 may also be reflected in selective brain damage. However, as noted above, lesion overlap analysis does not allow the inference of structure-function relationships, because regions may differ in their susceptibility to vascular damage.

The VLSM maps for the two key behavioural measures were extremely similar. This is not surprising since the nonfluent individuals showed a distinctive performance pattern across all behavioural measures as well as a distinctive lesion distribution. Therefore, the next section will provide a combined discussion of lesions associated with poor performance on

both low subject frequency sentences and on sentences containing meaning-related nouns. However, it is important to keep in mind that the role of other anatomical structures in the behavioural deficits observed cannot be reliably excluded, because null findings may be due to insufficient power to detect effects in these brain regions.

In line with the hypothesis, performance on both our key measures was associated with lesions to regions in the LIFG, including BA 44, BA 45, and BA 47, although the critical area of damage to each region was generally small. The involvement of the LIFG is particularly noteworthy, given recent findings that implicate this region in exaggerated semantic blocking effects in continuous naming tasks, a phenomenon that may reflect an inability to regulate lexical competition (e.g., Schnur et al., 2005; Schnur et al., 2009).

We further hypothesised that damage to the interface of BA44/6 may be associated with both of the behavioural measures, because this area has been linked to difficulties selecting words for specific positions in an utterance (e.g., Thothathiri et al., 2010). This could in principle reflect weak structure building processes. However, no significant association between BA 44/6 and performance on any of our behavioural measures was found. Nevertheless, there was some indication of possible involvement of this region in the maps obtained using uncorrected *p* values, and given the low power in the current study, we cannot rule out a possible involvement of this area in the experimental manipulations. With a potentially bigger sample size, the association with this region and the subject noun frequency and noun relatedness effects might have reached significance after correcting for multiple comparisons.

In addition to the LIFG, slight damage to other brain regions was also associated with the key behavioural measures. First, lesions to the left insula were linked to both the subject noun frequency effect and the noun relatedness effect. As mentioned previously, damage to the insula is commonly observed in nonfluent aphasia and seems to play a crucial role in the

articulatory difficulties that often occur in this disorder (Dronkers, 1996). Nevertheless, its involvement in the higher-level phenomena measured here is somewhat surprising. Second, lesions to portions of left inferior parietal lobe structures (the parietal operculum and the anterior supramarginal gyrus) also affected performance in the key measures, although the critical region implicated was relatively small. Again, these findings were surprising. Previous lesion studies have linked this region with poor phonological short-term memory. The supramarginal gyrus has been associated with the production of phonological errors in naming (Schwartz, Faseyitan, Kim, & Coslett, 2012) and auditory-verbal short-term memory, as measured by span tasks (Koenigs et al., 2011). While these processes are undisputedly involved in many aspects of language production, it remains unclear why they were significantly associated with the behavioural phenomena under investigation here. It might be argued that impairments to the verbal short-term memory in the nonfluent participants may have affected performance in Experiment 1 and 2. However, this explanation seems unlikely, because there were no differences in verbal short-term memory spans between the fluent and the nonfluent participants. Further, there were no differences in any of the sentence latency measures between the two aphasia groups that could have affected short-term memory maintenance demands of lexical elements. Consequently, we can rule out the possibility that sentence elements may have decayed before it was time to produce them due to longer speech initiation times in both of the aphasia groups. Another, perhaps more persuasive, explanation is that these regions are implicated purely due to their proximity to other, more critical regions. Future research with a bigger sample size and a broader lesion distribution is needed to confirm the reliability of these findings.

## General Discussion

Most studies of sentence production in nonfluent aphasia have examined accuracy across different types of sentence structures, while minimising the influence of lexical content. In this research, we have done the opposite: We have systematically examined the influence of lexical content on sentence production while keeping syntactic structure constant. In Experiment 1, the effect of noun frequency on SVO sentence production was investigated. It was found that in participants with nonfluent aphasia, the frequency of early noun elements in the sentence had a marked downstream effect on overall sentence production accuracy. This effect was observed even though the nouns themselves were within every participant's range of vocabulary and there were few errors involving these early nouns themselves. This effect was not observed in the participants with fluent aphasia. Control participants showed a similar, however less marked effect of lexical frequency on sentence production accuracy when compared to the nonfluent aphasic individuals. Further, the analysis of sentence latency measures identified two different types of sentence planning strategies: an incremental strategy, and an advance planning strategy. While controls, as a group, employed the incremental strategy, there was considerably individual variability in the strategies adopted by the various aphasic participants. This variability cut across the fluent-nonfluent distinction.

In Experiment 2, the same aphasic and control participants performed a sentence production task using a similar paradigm to that in Experiment 1, but this time the semantic relationship between the subject and object nouns was manipulated. The main finding was that the nonfluent individuals were less accurate at producing the SVO sentences when the subject and the object noun elements were semantically related than when they were unrelated. The fluent participants actually showed a trend in the *opposite* direction. As in Experiment 1, control participants showed a similar, but attenuated effect than the nonfluent

participants. Unexpectedly, the semantic manipulation had no reliable effect on sentence latency times; however, some nonfluent individuals showed a non-significant trend towards longer sentence initiation times when the sentence nouns were semantically related. These were the same individuals who showed evidence of (at least some) advance sentence planning in Experiment 1. Importantly, this effect was not found for any of the fluent participants, not even for the ones who also seemed to plan the sentence in advance.

Importantly, the experiments not only showed that lexical variables had a powerful effect on sentence production accuracy in nonfluent aphasia, but that, conversely, utterance structure also strongly influenced lexical retrieval success in these individuals. The critical finding in this respect was that the participants with nonfluent aphasia were very accurate at producing the key nouns in isolation, but their accuracy dropped substantially when the same nouns had to be produced in the context of a sentence, especially when they appeared in object position. This pattern was again unique to the individuals with nonfluent aphasia. Taken together, the findings from Experiment 1 and 2 suggest that a full understanding of sentence production difficulties in nonfluent aphasia can only be achieved if we consider the way in which lexical retrieval interacts with structural aspects of sentence planning.

Finally, voxel-based lesion-symptom mapping was used to explore the lesion correlates of the two major behavioural phenomena we observed in individuals with nonfluent aphasia (that is, the powerful influence of subject noun frequency on downstream sentence production accuracy, and the inhibitory influence of noun semantic relatedness). The resulting lesion maps were very similar for both behavioural phenomena: Both encompassed portions of the LIFG, a region that has previously been associated with lexical selection when competition is high (e.g., Schnur et al., 2005). However, the lesion maps also identified some regions that have not previously been associated with processes manipulated in the behavioural experiments, including the left inferior parietal cortex and the insula. Yet, it is

difficult to draw confident conclusions from the lesion analysis due to low statistical power and the small size of the participant pool.

In the following section, we will discuss the implications of our findings for existing theories of nonfluent aphasic sentence production and will provide an alternative theoretical framework that considers the interaction of lexical and structural aspects in nonfluent aphasic speech. Further, our findings have implications for the treatment of sentence production difficulties in nonfluent aphasia and provide important insights into how normal sentence production may operate, which will be discussed in a subsequent section.

### **Implications for Theories of Nonfluent Aphasic Sentence Production**

The findings from Experiment 1 and 2 clearly point out that lexical variables have an effect on sentence production accuracy in nonfluent aphasia, and that it is crucial to consider the way in which lexical retrieval and more structural aspects of sentence planning interact. As mentioned previously, conventional structure-based accounts attribute sentence production breakdown in nonfluent aphasia solely to difficulties with generating certain types of structural relations. Consequently, within these accounts sentence production success depends on the utterance syntactic structure only; the lexical content of the utterance is not considered. Even frameworks that pay attention to the dynamics of the sentence generation process, such as that of Kolk (1995, 2006), have difficulty to account for these findings. Kolk's framework attributes sentence production difficulties to a lack of synchronicity between the peak activation of lexical elements and that of their corresponding syntactic slots, as a consequence of slow retrieval and/or fast decay of either type of information. However, we found no differences in the speed of lexical retrieval between the two aphasia groups, neither in the single picture naming nor the overall sentence production task. Furthermore, delayed activation of syntactic slots in the sentence frame seems an unlikely explanation for the effects observed in nonfluent aphasia, since such a delay would have led



to greater synchronicity between syntactic slots and lexical elements for sentences containing low frequency nouns, which was the opposite of what we found. Furthermore, Kolk's framework does not address any processes that may account for the semantic relatedness effect observed in Experiment 2. Finally, the distinctive nonfluent aphasic performance patterns found across the two experiments make it seem implausible that the same deficit underlies both nonfluent and fluent aphasia, as proposed by Kolk. Also, the use of compensatory strategies employed by individuals with nonfluent aphasia to improve communication is unlikely to account for any of the observed effects, since syntactic structure in the sentence production tasks was held consistent, and only lexical content was manipulated.

We propose an explanation for our findings within a framework which allows for significant interplay between the processes responsible for the generation of structure and content, respectively (e.g., Bock, 1982; Chang et al., 2006; Stemmer, 1985). Within this type of framework, the sentence planning difficulty in nonfluent aphasia would be characterised as an inability to generate a robust, stable structural representation of the sentence being planned. In this situation, the bidirectional interplay between structure generation and lexical element retrieval processes becomes particularly crucial. Specifically, the construction of the sentence plan would rely disproportionately heavily on activation input from the constituent lexical elements. Sentence production success in nonfluent aphasia therefore depends particularly heavily on the ability to successfully and efficiently retrieve content words. We call this the *Content Drives Structure (COST)* hypothesis. The COST hypothesis offers a plausible explanation as to why sentence production accuracy in nonfluent aphasia is more susceptible to manipulations of lexical content than it is in fluent aphasia. Unlike those with nonfluent aphasia, individuals with fluent aphasia can rely on robust structure building processes to guide sentence production, so any inefficiency in producing a

particular lexical content element has only a localised effect, and does not impact on the remainder of the sentence. Furthermore, the COST hypothesis is compatible with the context-sensitivity of the nonfluent participants' noun production: They performed better when producing the key nouns in the pre-test than in the context of a sentence, and within the sentence context, they were more accurate to produce the nouns in an early than in a late position (see Figure 5.6 and 6.1). Both findings can be accommodated within the COST framework: Naming nouns in the pre-test does not rely on the successful generation of a structural plan, nor does it involve resolving syntagmatic interference from other sentence elements. Further, an inability to generate a robust structural plan would impact more heavily on constituents appearing late in the sentence, because the structure building process may not have been completed before their representation has decayed.

As discussed earlier, a complete account of nonfluent aphasic sentence production also needs to be able to explain the observed variability in performance between and within individuals. In the COST framework, the phenomenon of between-participant variability can be explained in terms of the degree of impairment to the structure generation process and/or how this interacts with other constraints imposed by the aphasic impairment (for example, slowed motor-articulatory planning). The framework also offers a means for explaining within-participant variability, since success on any given utterance will depend upon the degree to which lexical content information can be used to help support sentence planning, and may be further modulated by the degree of contextual support available (for example, a picture scene), and the specific time demands of the current conversational context.

Furthermore, the COST hypothesis can account for why the production of single words is relatively well preserved in individuals with nonfluent aphasia, since single word production does not require any coordination of syntactic and lexical elements, and lexical competition is reduced to a minimum. To summarise, our theoretical framework attributes a central role to

structure planning in nonfluent aphasia, but unlike more traditional accounts, it focusses in particular on the interplay between structural aspects of planning and the retrieval of content.

The semantic short-term memory model of Martin and colleagues offers a useful framework for elaboration of this hypothesis (e.g., Freedman et al., 2004; Martin et al., 1999; Martin et al., 2004). These researchers suggest that in nonfluent aphasia, there may be weak or inefficient binding of lexical elements to their structural positions within a short-term memory representation of the dominant phrase. This weak structure binding leads to increased production failures, particularly on phrases containing multiple words with similar selection constraints, where structure binding is crucial to resolve order. Importantly, the likelihood of such production failures depends on the lexical content elements themselves. For example, a strongly activated lexical element in turn strongly activates its corresponding structural representation in semantic short-term memory, thereby mitigating the effects of the impairment. Moreover, if elements are only weakly bound to their positions in semantic short-term memory, then they will be freer to compete with one another for selection to a given position (provided, of course, they have similar selection constraints, so are both plausible candidates for a particular position). This competition might be particularly intense if the words are also semantically related, as they were in Experiment 2. However, Martin and colleagues' framework deals only with planning within a single phrase. In this study, we were able to demonstrate semantic interference effects that span more than a single phrase, so, to account for our results, we need a framework that considers sentence planning more broadly.

One model of aphasic sentence production that is compatible with the COST framework is the DoL model (Gordon & Dell, 2003; see also Dell et al., 2008). In this model, activation of different lexical elements depends on combined syntactic-sequential and conceptual-semantic input. The lexical elements required for an utterance receive activation

from their corresponding conceptual-semantic input nodes, and also from the syntactic-sequential nodes corresponding to the element's syntactic class. Implemented through a learning algorithm, the connection strength between lexical elements and their corresponding conceptual-semantic and syntactic-sequential input nodes was found to differ depending on whether the words were primarily represented by semantic features or whether they primarily served a syntactic function: Nouns and semantically rich verbs were mainly activated by conceptual-semantic input nodes, whereas determiners and semantically impoverished verbs were primarily activated by syntactic-sequential input nodes (Gordon & Dell, 2003).

Accordingly, by lesioning the syntactic-sequential nodes in the DoL model, Gordon and Dell (2003) were able to simulate the consequences of selective damage to the structure building process. The resultant model, where lexical elements were primarily reliant on top-down activation from semantic-conceptual input nodes, produced an agrammatic pattern of speech, characterised by the selective omission of determiners and semantically impoverished verbs. Importantly, when the syntactic-sequential nodes were lesioned, enhanced lexical competition also occurred, because all the planned sentence elements became concurrently activated, and no structural representation was in place to enhance activation of the right word for the right position. The DoL model only simulated subject-verb sentences, but presumably this effect would be considerably intensified for sentence elements that share the same word class (e.g., subject and object nouns), and particularly so if they also share semantic features. This is exactly what we found in Experiment 2 in the participants with nonfluent aphasia.

A recent model that may offer a more elegant account of these phenomena is the Dual-path model of Chang and colleagues (2006). In this model, the retrieval of lexical content is controlled primarily by the *meaning system*, and the encoding of syntactic structure is controlled primarily by the *sequencing system*. Although these two systems have no direct

connections, both are capable of interacting via a common set of *thematic role units*.

Specifically, the thematic role units simultaneously receive activation from two sources: 1)

The sequencing system, which keeps track of the sequencing constraints that currently apply given the recent production history, activates those thematic units that are appropriate for selection to the next word position; and 2) the meaning system, which transmits activation from currently activated concept units to their corresponding thematic units, according to their thematic role in the sentence being planned. With regards to nonfluent aphasia, the key impairment could in principle involve either the sequencing system or the process of thematic role assignment itself. An impairment to the sequencing system would in principle result in an overreliance on the meaning system to drive element selection, which may possibly results in the systematic omission of function words. Also, without the sequencing system to manage the sequential processes of element selection, utterance production will rely more heavily on the meaning system, thereby giving rise to powerful lexical content effects. The influence of lexical content will be particularly powerful at the time of initiation of the sentence, because at this stage, selection of the initial thematic unit is based primarily on the relative levels of activation of the various candidates from the meaning system. In addition, if the sequencing system is ineffective at activating each lexical element in the appropriate sequence, lexical elements planned for the same utterance may compete more intensely with one another.

### **Implications for the Treatment of Sentence Production Difficulties in Nonfluent Aphasia**

A consideration of the way in which structural planning and content retrieval interact in nonfluent aphasia is important not just from a theoretical point of view, but also has practical consequences. For example, treatment devices that offer lexical support during sentence production may be particularly beneficial for individuals with nonfluent aphasia. One such treatment device is a therapy software from Linebarger and colleagues, the

*SentenceShaper*®, which specifically supports the retrieval and/or maintenance of key lexical elements during sentence planning (Linebarger et al., 2007; Linebarger et al., 2000). The *SentenceShaper*® allows the patients to record, replay, and reorder sentence elements that they wish to produce, thus enabling them to make lexical content more readily available before constructing a sentence. This has been found to result in an increase in longer and well-formed sentences in many cases with nonfluent aphasia – even though no syntactic assistance is provided.

These findings suggest that word-retrieval treatments that focus on making lexical content available directly by training specific words may not only improve lexical retrieval itself, but may also positively impact on overall sentence production performance in nonfluent aphasia. In fact, results from a recent therapy study support this idea: After word-retrieval training, an individual with nonfluent aphasia showed higher overall sentence production accuracy for sentences containing trained nouns and verbs, including the use of correct inflections and determiners (Raymer & Kohen, 2006).

The finding from treatment studies are in line with the COST framework, which predicts that activation from lexical elements will support a more elaborated construction of the syntactic frame and improve overall sentence production performance. Taken together, the findings from the current study suggest that a greater attention to lexical support in treatment of nonfluent aphasia can result in significant gains.

### **Implications for Theories of Normal Sentence Production**

Research into disorders that selectively impair different aspects of the speech planning process, such as the ones studied here, can provide important insights into how normal sentence production operates. As discussed earlier, models of sentence production differ to the extent to which they assume an interaction between structure and content retrieval processes during speech planning. If the COST hypothesis is correct, then this would imply

that effective content retrieval processes can support syntactic structure building and conversely, that effective syntactic structure generation can reduce levels of syntagmatic competition between lexical elements planned for the same utterance. A theoretical model capable of accommodating these phenomena would therefore need to allow for some considerable interaction between structure and content encoding processes.

Examples of early models that allow for such interaction include those of Bock (1982) and Stemberger (1985). Within Bock's (1982) framework, currently active lexical content has the capacity to influence the overall syntactic plan, because the lexical arena interacts with the syntactic arena. Within this framework, a stable activation of lexical content corresponds to a stable activation of syntactic structure. This feature of the model could help explain the subject frequency effect observed in nonfluent aphasic participants in Experiment 1: If activation of elements in the structural arena is poor, sentence planning processes would benefit particularly from the robust activation and ready availability of lexical elements, particularly if those elements appear early in the sentence. This may help speed up the syntactic encoding process, enabling late elements to be integrated into the syntactic frame before they decay from verbal working memory. Conversely, if those early lexical elements do not receive robust and rapid activation, the syntactic encoding process may not be completed before subsequent elements have decayed. However, Bock's account does not provide enough detail regarding the principles of spreading activation in and between the different arenas in order to explain the semantic relatedness effect observed in the nonfluent participants in Experiment 2. Also, the model does not specifically allow for feedback from the syntactic arena to the lexical arena, a principle useful to help reduce syntagmatic interference during sentence planning, as will be discussed in the following.

Unlike Bock, Stemberger (1985) proposes that flow of information between lexical and syntactic processing during sentence planning is *bi-directional*. In other words, each slot

in the syntactic plan for the sentence is capable of activating its corresponding lexical elements, and in turn, each lexical element is capable of activating its corresponding syntactic slot. This means that currently active lexical content has the capacity to drive structure building processes, as in Bock's model. However, the assumption of bi-directional flow of information allows for syntactic processes to influence lexical processing as well: Stemberger postulates that slots in the phrase frame (e.g., noun phrase) can activate their corresponding lexical elements (e.g., nouns) and inhibit non-corresponding ones (e.g., verbs), thereby reducing syntagmatic interference between the sentence elements. If, however, the phrase frames are poorly activated from higher message levels, lexical elements that fit current structural requirements will not receive a sufficient additional increase of activation, whilst unfitting ones will not be sufficiently inhibited. As a consequence, lexical elements are freer to compete with each other for production, aggravating lexical selection and leading to poorer sentence production performance. This account may go some way towards explaining why individuals with nonfluent aphasia are particularly susceptible to syntagmatic competition. However, since the primary structural units in this model are syntactic units (noun phrases, verb phrases and so forth), it has a more difficult time managing competition between words from the same syntactic class.

Another, more recent example of a model that allows for interaction between content and structure generation is the DP model of Chang and colleagues (2006). Although this model does not allow for direct interaction between lexical and syntactic units, it does include a mechanism whereby content and structure generation processes can interact during sentence planning: Both are capable of interacting via a common set of *thematic role units*. Because the primary units of interaction in this model are the thematic role units, rather than the syntactic category units, this model cannot only manage competition between words of different syntactic classes, but also that between words of the same syntactic class that are



destined to perform different thematic roles. In addition, processes occurring in the meaning system have a direct influence upon which elements are selected at the beginning of the sentence, since the selection of the first thematic unit is based primarily on the level of activation of the various candidate units at the time of sentence initiation. For example, if the patient unit is more activated than the agent unit due to a higher prominence of the concept it represents, it will be linked to an early position in the sentence and signal for a passive construction.

The findings from Experiments 1 and 2 also have implications for our understanding of the scope of planning in sentence production. Overall, findings from the latency analysis are consistent with previous research that has suggested a phrasal scope of sentence planning in healthy speakers (e.g., F. Ferreira, 1991; Smith & Wheeldon, 1999; Wheeldon et al., 2013). However, because each phrase contained only one content word, one could argue that the scope of planning might even be narrower, perhaps only consisting of one content word at a time (see Griffin & Bock, 2000, for a consistent proposal). However, control participants were slower to produce a low frequency object noun when the subject noun was of high frequency, indicating a certain degree of overlap between retrieving the subject and the object noun.

Results from the individual analysis of aphasic latency measures were less clear. While some aphasic participants showed a similar phrase-by-phrase planning strategy than the controls, others showed evidence of more extensive advance planning. Since there was no systematic relationship between sentence planning style and broad aphasia subtype, the nature of these differences is unclear. One possible explanation is that these findings simply reflect different individual preferences for how far to plan an utterance in advance. An alternative explanation is that the experimental design of the sentence production task had an impact on the planning strategy used by the aphasic individuals. Recent research has shown

that the degree of advance planning may depend on current processing demands, in the form of task complexity and time constraints, favouring an adaptable scope of planning (F. Ferreira & Swets, 2002). In the current study, some aphasic individuals may have adopted a more incremental planning strategy in favour of producing the sentences as quickly as possible (like the control participants), whilst others may have carefully planned the sentence in advance in favour of reducing processing load.

### **Limitations and Suggestions for Future Research**

One important limitation to our study was that syntactic structure was kept constant to very simple SVO sentences, which were chosen to ensure that all the participants with aphasia were likely to produce a substantial number of the sentences correctly. Since the syntactic structure of the target sentences was not varied, we cannot say whether the lexical effects observed here would extend to the production of other types of syntactic structures, and we cannot say how these lexical effects might interact with more elaborated structural aspects of the sentence. Indeed, since content effects in this study have been shown to play a role in very simple canonical sentences containing words that are in the participant's vocabulary, this effect might be expected to play an even more important role in more complex and more resource demanding utterances. This, in fact, is an avenue for future research. However, designing such future studies will be challenging because of the difficulty eliciting specific sentence structures, especially if one wants to avoid the use of linguistic and other types of prompts (for example, arrows to indicate which element should be produced first). If the task becomes too difficult and participants produce many errors, this could easily result in floor effects, and a very sensitive scoring procedure would be needed. Other elicitation techniques that do not remove the picture at speech onset or use arrows to point to the agent of the sentence may be used; however, patients might then rely on using a simple

labelling strategy which may not resemble the demands of naturally occurring speech. Thus, the design of the current study clearly provides some advantages.

Another important limitation to the current study is that the sample size of participants with aphasia was too small to draw confident conclusions from the lesion analysis. Having a larger sample for both the nonfluent and the fluent group would enable making more valid inferences from the brain region that were associated with performance in Experiment 1 and 2, because it would result in a higher statistical power of the administered statistical tests and more variation in the lesion locations. In theory, including a sample that is not selected based on good naming abilities and thus does not restrict the distribution of brain lesions would further strengthen the analysis. However, testing aphasic individuals with severe naming deficits would obscure the interpretation of latency and accuracy results in the sentence production tasks, as sentence production difficulties would most likely be confounded with lexical impairments.

Overall, it would be desirable to test the specific assumptions made in the COST framework more closely in future studies. For example, it would be interesting to investigate the effect of lexical availability on sentence production using other experimental paradigms, such as lexical priming. A lexical priming experiment could manipulate the availability of lexical elements directly by presenting an auditory prime before the onset of a picture that needs to be described in a sentence. The primes could be identical to the subject or the object of the sentence to be produced, and could also involve neutral words. Within the COST framework, priming the subject of a sentence would be expected to support structure building processes and improve overall sentence production accuracy in nonfluent aphasic participants, while it may have more local effects in participants with fluent aphasia.

Finally, an important issue that has not been addressed in the current study concerns whether the structure generation processes hypothesised to be impaired in nonfluent aphasia

are specialised for the production of grammatical sentence frames, or whether they perform a more general function in modulating levels of lexical activation. Several recent studies have observed semantic interference effects in nonfluent aphasia even on tasks that involve the production of only one word at a time, such as in blocked cycling tasks (e.g., Schnur et al., 2006; Schwartz & Hodgson, 2002; Scott & Wilshire, 2010; Wilshire & McCarthy, 2002). In consideration of such findings, some researchers have proposed the existence of a more general mechanism that modulates lexical activation in language production in response to specific task demands, and it is this mechanism that is impaired in nonfluent aphasia (e.g., Dell et al., 2008; Oppenheim, Dell, & Schwartz, 2010; Scott & Wilshire, 2010; Wilshire & McCarthy, 2002). This proposed mechanism seems to operate in a very similar way to the structure building process proposed here, which reduces competition between words in an utterance by boosting the ones that fit current syntactic constraints while inhibiting the rest. Though being involved in eliminating different types of lexical competition (paradigmatic vs. syntagmatic), both mechanisms enforce the inhibition of non-target lexical items and the activation of target lexical items. One possible explanation may be that both of these phenomena arise from the same source, an overall (anterior) mechanism that is responsible for general goal direction and inhibitory control in verbal tasks. In order to explore this idea further, it would be interesting to examine the performance of individuals with nonfluent aphasia on tasks that manipulate the degree of lexical competition at the single word and at the sentence level.

In fact, one recent case study speaks directly to this issue: Nonfluent participant JHM, who showed semantic interference effects in Experiment 2, has also shown impairments in a variety of other tasks manipulating lexical competition (Scott & Wilshire, 2010). For example, JHM was less accurate and slower when she had to name semantically related pictures compared to unrelated pictures in a blocked cycling task. These findings indicate a

potential relationship between the mechanisms that operate to resolve paradigmatic interference at the single word level and syntagmatic interference at the sentence level, supporting the idea of one general process responsible for minimising competition throughout the language system, which may be compromised in nonfluent aphasia. Furthermore, it has been suggested that this general process is localised in the LIFG (particularly BA 44; e.g., Novick et al., 2009; Schnur et al., 2009; Thompson-Schill et al., 1998). In fact, the analysis of JHM's MRI scan revealed that she had damage to this particular brain area. However, this issue requires further investigation and the converging findings from single word and sentence production tasks need to be confirmed in a larger number of participants before stronger conclusions can be drawn.

## **Conclusions**

In two experiments, the effect of lexical content on sentence production performance in nonfluent aphasic, fluent aphasic and control participants was investigated. Traditionally, most studies of sentence production in nonfluent aphasia have examined accuracy across different types of sentence structures, while paying little attention to lexical content. However, the results of the current research demonstrate that lexical content has a powerful influence on sentence production success in nonfluent aphasia, even when the key words are well within the participants' vocabulary (they are produced quite accurately in isolation). We observed marked effects of lexical frequency (Experiment 1) and semantic relatedness (Experiment 2) of sentence nouns on overall sentence production accuracy only in the participants with nonfluent aphasia. Furthermore, the analysis of lesion sites and performance accuracy in the aphasic participants across the two experiments indicates a potential role of the LIFG in structure building processes and in resolving syntagmatic interference.

Overall, the findings are consistent with the COST hypothesis, proposing that structure building processes in nonfluent aphasia are compromised and that sentence production

success depends heavily on the ability to successfully and efficiently retrieve content words. In fact, the reported findings demonstrate that difficulties with structural planning should not be considered in isolation, and point out the importance to consider the interaction of content and structure in future research into sentence production in nonfluent aphasia. Furthermore, the findings highlight the need for greater attention to processing aids in the treatment of individuals with nonfluent aphasia.

## List of Appendix Tables

Table A1 .....	177
<i>Target nouns used in Experiment 1 and corresponding frequency ratings</i>	
Table A2 .....	178
<i>Target pictures used in Experiment 1 with corresponding experimental condition</i>	
Table A3 .....	179
<i>Distracter pictures used in Experiment 1 and 2 with corresponding sentence types</i>	
Table B1 .....	181
<i>Geometric mean pre-test naming latencies for each individual aphasic participant and the three participant groups as a function of frequency condition (Experiment 1)</i>	
Table B2 .....	182
<i>Naming latency in the pre-test as a function of frequency for participants with aphasia (Experiment 1)</i>	
Table B3 .....	183
<i>Individual error analyses in the sentence production task for participants with aphasia (Experiment 1)</i>	
Table B4 .....	184
<i>Supplementary details for the group error analyses in the sentence production task for each of the three participant groups and across groups (Experiment 1)</i>	
Table B5 .....	185
<i>Geometric means of participants with aphasia and the three participant groups for the different latency measures in the different conditions (Experiment 1)</i>	
Table B6 .....	187
<i>Subject and object frequency effects on sentence latency measures in participants with aphasia and supplementary details for the three participant groups (Experiment 1)</i>	
Table C1 .....	188
<i>Target nouns used in Experiment 2 and corresponding frequency ratings</i>	
Table C2 .....	189
<i>Target pictures used in Experiment 2 with corresponding experimental condition</i>	

Table D1 .....	190
<i>Pre-test naming errors for aphasic participants and the three participant groups in the different frequency conditions (Experiment 2)</i>	
Table D2 .....	191
<i>Geometric mean pre-test naming latencies for each individual aphasic participant and the three participant groups as a function of frequency condition (Experiment 2)</i>	
Table D3 .....	192
<i>Naming latency in the pre-test as a function of frequency for participants with aphasia (Experiment 2)</i>	
Table D4 .....	193
<i>Individual error analyses in the sentence production task for participants with aphasia (Experiment 2)</i>	
Table D5 .....	194
<i>Supplementary details for the group error analyses in the sentence production task for each of the three participant groups and across groups (Experiment 2)</i>	
Table D6 .....	195
<i>Geometric means of participants with aphasia and the three participant groups for the different latency measures in the different conditions (Experiment 2)</i>	
Table D7 .....	197
<i>Noun relatedness and frequency effects on sentence latency measures for participants with aphasia and supplementary details for the three participant groups (Experiment 2)</i>	



## Appendix A: Stimulus Material for Experiment 1

Table A1

*Target nouns and corresponding frequency ratings*

Noun	Frequency rating
High frequency	( $\geq 30$ occ. per million)
Bear	57.41
Cat	66.33
Dog	192.84
Ghost	36.59
Horse	92.88
King	129.52
Nurse	44.98
Pig	39.14
Queen	54.69
Low frequency	(< 30 occ. per million)
Clown	15.82
Cow	25.51
Fox	21.61
Goat	10.53
Maid	22.82
Nun	6.96
Sheep	22.35
Snake	13.43
Witch	27.65

*Note.* occ = occurrence; Frequency ratings according to the Subtitle Analysis Project (Brysbaert & New, 2009).

Table A2

*Target pictures with corresponding experimental condition*

high frequency subject / high frequency object	high frequency subject / low frequency object	low frequency subject / high frequency object	low frequency subject / low frequency object
The dog is licking the cat	The cat is chasing the sheep	The sheep is shooting the pig	The fox is shooting the cow
The pig is lifting the cat	The pig is shooting the sheep	The fox is chasing the dog	The snake is biting the fox
The pig is biting the bear	The nurse is kissing the witch	The sheep is chasing the cat	The fox is chasing the goat
The ghost is shooting the king	The horse is licking the fox	The fox is kicking the bear	The cow is shooting the fox
The queen is pushing the nurse	The bear is kicking the fox	The cow is kicking the horse	The goat is licking the cow
The ghost is washing the queen	The nurse is shooting the clown	The snake is licking the pig	The fox is biting the snake
The cat is licking the dog	The king is pushing the maid	The snake is kissing the bear	The witch is kissing the nun
The king is shooting the ghost	The dog is chasing the fox	The witch is washing the nurse	The goat is chasing the fox
The horse is chasing the bear	The bear is kissing the snake	The maid is hitting the ghost	The sheep is kicking the cow
The cat is lifting the pig	The ghost is hitting the maid	The witch is kissing the nurse	The nun is kissing the witch
The bear is chasing the horse	The pig is licking the snake	The maid is pushing the king	The cow is kicking the sheep
The horse is kicking the pig	The nurse is washing the witch	The clown is shooting the nurse	The maid is lifting the nun
The queen is washing the ghost	The ghost is pushing the clown	The clown is pushing the ghost	The sheep is brushing the goat
The nurse is pushing the queen	The bear is lifting the goat	The goat is lifting the bear	The nun is lifting the maid
The pig is kicking the horse	The dog is biting the goat	The fox is licking the horse	The cow is licking the goat
The bear is biting the pig	The horse is kicking the cow	The goat is biting the dog	The goat is brushing the sheep
			The clown is washing the maid
			The maid is washing the clown

Table A3

*Distracter pictures with corresponding sentence types*

Sentence Type	Distracter Picture
Direct object, Prepositional phrase	<p>The librarian is giving a book to the boy</p> <p>The girl is handing a paint brush to the man</p> <p>The father is reading a story to his son in bed*</p> <p>The girl is throwing a hammer to the boy</p> <p>The man is giving a present to the girl</p> <p>The mother is bringing a cake to her son</p> <p>The children are showing a picture to the teacher</p>
Double object, Possessive phrase	<p>The woman is bathing her foot in a bowl</p>
Double object, Prepositional phrase	<p>The girl has cut her hand with a broken bottle</p> <p>The girl is drawing a smiley face on the blackboard</p> <p>The woman is taking the baking out of the oven</p>
Intransitives	<p>The boy is skating</p> <p>The boy is cooking</p> <p>The girl is reading</p> <p>The girl is cooking</p> <p>The man is shaving</p> <p>The man is sleeping</p> <p>Two men are running</p> <p>The girl is crying</p> <p>The woman is ironing</p> <p>The woman is sowing</p> <p>The woman is singing</p> <p>The woman is swimming</p> <p>The women are raking</p>
Passives	<p>The man is robbed by a robber</p> <p>The boy is getting hit by a baseball</p> <p>The church is struck by lightening</p> <p>The fly has been swatted by a fly swat</p> <p>The man is getting hit by a rock</p> <p>The window has been broken by a stone</p> <p>The fire-fighter is getting splashed by the hydrant</p>

*Table A3 continued.*

Sentence Type	Distracter Picture
Prepositional phrase	<p>The car is on top of the garage  The boy is sliding down a slide  The girl is skipping with a rope  The car is towing a boat  Two babies are playing with balls  The car is running over a ball  The truck crashes into a tree  The children sing the national anthem around the flag  The girl is climbing over the fence  The girl is looking into the cookie jar  The man is sitting on a chair</p>
Single object	<p>The boy is watching TV  The girl is eating a lollypop  Two girls are riding tricycles  The man is opening a door  The man is painting a wall  The artist is painting a woman  The man is sharpening an axe  The man is throwing a rugby ball  The man is watching TV  The man is writing a letter  The men are shaking hands  The woman is carrying a box  The woman is eating breakfast*  The woman is hanging up some clothes  The woman is sweeping leaves  The woman is cleaning the dishes  The boy is frying a meatball</p>
Single object, Conjoined Subject phrase	<p>The boy and the girl are digging a hole  The boy and the girl are picking flowers  The father and the son are reading the newspaper</p>
Single object, Possessive phrase	<p>The girls are brushing their hair  The women is brushing her hair  The woman is spilling her drink</p>

*Note.* \* Pictures not used in Experiment 2.

## Appendix B: Supplementary Data and Analyses for Experiment 1

Table B1

*Geometric mean pre-test naming latencies for each individual aphasic participant and the three participant groups as a function of frequency condition*

	Geometric mean naming latency in milliseconds	
	High Frequency	Low Frequency
Nonfluent		
BY	1097 (783-1629)	1588 (919-2744)
DA	1468 (1051-2047)	1644 (1085-2490)
JG	1325 (908-1932)	2252 (957-5300)
JHM	986 (804-1210)	1140 (969-1342)
RP	1076 (865-1337)	1526 (995-2340)
Group mean	1182 (883-1581)	1628 (998-2655)
Fluent		
NP	1326 (961-1830)	1517 (958-2401)
STR	1461 (1074-1987)	1567 (1203-1980)
SW	1438 (1109-1866)	1692 (1211-2362)
WL	1425 (978-2098)	1901 (1050-3443)
Group mean	1413 (348-3821)	1669 (638-4051)
Controls		
Group mean	744 (565-980)	818 (650-1031)

*Note.* Figures in parentheses represent values one standard deviation above and below the mean, respectively (calculated from logged reaction time values, then presented in antilogged form).

Table B2

*Naming latency in the pre-test as a function of frequency for participants with aphasia*

Participant	<i>df</i>	<i>F</i>	<i>p</i>
Nonfluent			
BY	1, 32	5.33	<b>&lt;.05</b>
DA	1, 33	1.14	.29
JG	1, 29	3.97	<.10
JHM	1, 33	4.36	<b>&lt;.05</b>
RP	1, 28	10.28	<b>&lt;.01</b>
Fluent			
NP	1, 30	1.12	.29
STR	1, 35	0.27	.61
SW	1, 26	3.12	<.10
WL	1, 23	2.04	.17

*Note.* Effects significant at the  $p < .05$  level are in bold font.

Table B3

*Individual error analyses in the sentence production task for participants with aphasia*

Participant	Errors in Context			Sentence Errors							Sentence Errors excluding Subject Noun Errors						
	Error Position			Subject Freq			Object Freq		SF*OF		Subject Freq			Object Freq		SF*OF	
	<i>df</i>	$\chi^2$	<i>p</i>	<i>df</i>	$\chi^2$	<i>p</i>	$\chi^2$	<i>p</i>	$\chi^2$	<i>p</i>	<i>df</i>	$\chi^2$	<i>p</i>	$\chi^2$	<i>p</i>	$\chi^2$	<i>p</i>
Nonfluent																	
BY	1	3.91	<.05	1	2.82	<.10	2.04	.15	0.08	.78	1	0.99	.32	2.33	.13	0.01	.93
DA*	1	10.99	<.001	1	2.73	<.10	0.44	.51	0.38	.54	1	2.73	<.10	0.44	.51	0.38	.54
JG	1	12.96	<.001	1	4.12	<.05	1.38	.24	0.25	.62	1	2.98	<.10	0.05	.82	0.47	.49
JHM	1	18.26	<.0001	1	5.40	<.05	0.03	.86	0.46	.50	1	5.37	<.05	0.09	.77	0.45	.50
RP	1	3.86	<.05	1	1.86	.17	4.06	<.05	0.23	.63	1	2.38	.12	3.15	<.10	0.61	.43
Fluent																	
NP	1	0.28	.60	1	3.56	<.10	5.72	<.05	4.35	<.05	1	2.24	.13	1.67	.20	3.76	<.10
STR	1	7.44	<.01	1	0.31	.58	0.39	.53	1.88	.17	1	1.82	.18	1.83	.18	1.36	.24
SW	1	1.00	.32	1	7.25	<.01	0.12	.72	1.23	.27	1	3.39	<.10	0.74	.39	0.82	.37
WL	1	2.04	.15	1	2.92	<.10	1.98	.18	2.36	.12	1	3.40	<.10	0.65	.42	1.32	.25

*Note.* Effects significant at the  $p < .05$  level are in bold font. Freq = Frequency. SF = Subject Frequency. OF = Object Frequency.

\* DA made no errors on nouns in the subject position, therefore results from the analysis of sentence errors excluding subject noun errors are identical to results from the sentence error analysis.

Table B4

*Supplementary details for the group error analyses in the sentence production task for each of the three participant groups and across groups*

Groups	Sentence Errors							Sentence Errors excluding Subject Noun Errors										
	Object Freq			SF*OF		Group*OF		Subject Freq			Object Freq		SF*OF		Group*SF		Group*OF	
	<i>df</i>	$\chi^2$	<i>p</i>	$\chi^2$	<i>p</i>	$\chi^2$	<i>p</i>	<i>df</i>	$\chi^2$	<i>p</i>	$\chi^2$	<i>p</i>	$\chi^2$	<i>p</i>	$\chi^2$	<i>p</i>	$\chi^2$	<i>p</i>
Nonfluent	1	7.05	<.01	0.03	.87			1	7.43	<.01	0.00	.97	1.01	.31				
Fluent	1	1.21	.27	0.11	.74			1	0.87	.35	1.15	.28	0.06	.80				
Controls	1	8.15	<.01	1.14	.29			1	10.38	<.01	8.21	<.01	2.94	<.10				
Nonfluent vs. Controls	1					5.04	<.05	1							0.61	.44	2.03	.15
Nonfluent vs. Fluent	1					0.90	.34	1							6.23	<.05	0.85	.36
Fluent vs. Controls	1					8.69	<.01	1							5.31	<.05	3.48	<.10

*Note.* Effects significant at the  $p < .05$  level are in bold font. Freq = Frequency. SF = Subject Frequency. OF = Object Frequency.



Table B5

*Geometric means of participants with aphasia and the three participant groups for the different latency measures in the different conditions*

Participant	Sentence Initiation Time				Subject-Noun Initiation Time				Verb-object Noun Initiation Time			
	Subject Frequency		Object Frequency		Subject Frequency		Object Frequency		Subject Frequency		Object Frequency	
	high	low	high	low	high	low	high	low	high	low	high	low
Nonfluent												
BY	3354 (2013-4631)	3596 (2275-5683)	3209 (2040-5049)	3430 (2221- 5296)	3602 (2243-5784)	4082 (2458-6778)	3762 (2244-6308)	3915 (2248-6262)	851 (341-2122)	1023 (339-3081)	787 (301-2057)	1106 (390-3136)
DA	4478 (2675-7498)	5068 (3087-8322)	4961 (2887-8523)	4573 (2849-7341)	5256 (3168-8722)	5487 (3390-8883)	5543 (3228-29516)	5200 (3348-8076)	2075 (1030-4177)	2421 (1218-4813)	2409 (1179-49223)	2096 (1070-4108)
JG	3624 (2298-5714)	4038 (2594-6287)	3659 (2370-5649)	3996 (2504-6379)	4100 (2624-6404)	4340 (2785-6765)	3986 (2645-6005)	4477 (2785-7198)	885 (389-2312)	732 (299-1792)	780 (275-2141)	855 (376-1948)
JHM	2149 (1436-3217)	2514 (1651-3828)	2431 (1639-3608)	2242 (1449-3469)	2643 (1710-4084)	2985 (1971-4520)	2878 (1931-4290)	2756 (1753-4334)	1230 (572-2645)	1260 (621-2554)	1115 (541-2300)	1374 (659-2865)
RP	4907 (3060-7867)	6500 (4088-10337)	5364 (3235-8893)	6016 (3783-9567)	5319 (3375-8387)	7692 (4405-10473)	5604 (3503-8964)	6472 (4169-10048)	1125 (486-2606)	1236 (505-3028)	1168 (484-2819)	1196 (506-2830)
Group mean	3496 (2517-4858)	4131 (2887-5911)	3768 (2726-5208)	3850 (2670-5551)	4051 (3029-5419)	4560 (3340-6224)	4221 (3184-5596)	4388 (3186-6042)	1167 (816-1667)	1231 (795-1904)	1136 (716-1804)	1267 (910-1762)

Table B5 continued.

Participant	Sentence Initiation Time				Subject-Noun Initiation Time				Verb-object Noun Initiation Time			
	Subject Frequency		Object Frequency		Subject Frequency		Object Frequency		Subject Frequency		Object Frequency	
	high	low	high	high	low	high	high	low	high	high	low	high
Fluent												
NP	3863 (2954-5051)	4250 (3099-5828)	3820 (2947-4952)	4282 (3115-5886)	4177 (3250-5368)	4660 (3370-6431)	4204 (3221-5486)	4617 (3384-6299)	705 (295-1687)	606 (260-1410)	592 (254-1381)	720 (303-1709)
STR	2536 (1850-3475)	2805 (2081-3781)	2717 (2011-3671)	2626 (1907-3615)	2965 (2127-4132)	3320 (2347-4697)	3088 (2273-4193)	3204 (2192-4683)	484 (279-840)	533 (249-1140)	509 (251-1033)	509 (271-958)
SW	2620 (1860-3690)	2971 (2207-4000)	2601 (1841-3674)	2957 (2194-3986)	2803 (2013-3903)	3101 (2351-4090)	2737 (1996-3752)	3132 (2345-4183)	553 (277-1103)	643 (297-1391)	566 (267-1203)	623 (303-1284)
WL	4741 (2134-10531)	5053 (2272-11237)	5004 (2242-11171)	4812 (2172-10664)	6141 (3091-12205)	6593 (3295-13194)	6644 (3377-13073)	6141 (3443-12393)	1787 (728-4384)	2631 (931-7433)	2262 (776-6591)	2140 (850-5385)
Group mean	3321 (2449-4504)	3658 (2755-4855)	3409 (2505-4640)	3557 (2662-4752)	3821 (2660-5489)	4217 (2981-5967)	3920 (2639-5821)	4107 (2977-5666)	762 (423-1374)	860 (406-1820)	789 (389-1597)	836 (440-1589)
Controls												
Group mean	1198 (896-1455)	1290 (1041-1597)	1248 (1015-1533)	1239 (1005-1528)	1353 (1100-1664)	1437 (1141-1809)	1395 (1120-1736)	1394 (1115-1743)	191 (128-285)	196 (129-297)	186 (125-276)	202 (133-306)

*Note.* Figures in parentheses represent values one standard deviation above and below the mean, respectively (calculated from logged reaction time values, then presented in antilogged form).

Table B6

*Subject and object frequency effects on sentence latency measures in participants with aphasia and supplementary details for the three participant groups*

Participant	Sentence Initiation Times								Subject-Noun Initiation Times								Verb-object Noun Initiation Times								
	Subj. Freq				Obj. Freq				SF*OF				Subj. Freq				Obj. Freq				SF*OF				
	<i>df</i>	<i>F</i>	<i>p</i>		<i>F</i>	<i>p</i>			<i>df</i>	<i>F</i>	<i>p</i>		<i>F</i>	<i>p</i>			<i>df</i>	<i>F</i>	<i>p</i>		<i>F</i>	<i>p</i>			
Nonfluent																									
BY	1, 84	5.14	<.05		0.62	.44		1.28	.26	1, 84	2.01	.16	0.10	.75	0.84	.36	1, 78	0.87	.35	3.22	<.10	0.40	.53		
DA	1, 94	9.16	<.01		0.00	.98		0.05	.83	1, 92	1.85	.17	0.11	.75	1.28	.26	1,100	1.61	.21	0.95	.33	0.35	.55		
JG	1, 72	1.86	.18		0.88	.35		0.94	.33	1, 72	0.57	.45	1.95	.17	0.44	.51	1, 72	1.42	.24	0.34	.56	0.18	.68		
JHM	1,103	2.78	<.10		0.74	.40		0.35	.55	1,100	1.31	.25	3.94	<.05	0.17	.68	1, 82	0.02	.88	5.03	<.05	4.39	<.05		
RP	1, 70	9.30	<.01		2.06	.16		0.52	.47	1, 67	7.92	<.01	3.38	<.10	0.28	.60	1, 71	0.35	.56	0.04	.83	0.20	.65		
Group										1, 667	8.21	<.01	0.67	.44	0.63	.43									
Fluent																									
NP	1, 66	2.88	<.10		3.96	<.10		0.02	.90	1, 66	3.96	<.10	2.63	.11	0.06	.81	1, 64	1.47	.23	1.49	.23	0.11	.74		
STR	1, 76	3.67	<.10		0.32	.58		0.01	.91	1, 76	3.80	<.10	0.33	.57	0.12	.73	1, 70	0.59	.44	0.00	.96	2.41	.12		
SW	1, 48	4.45	<.05		4.16	<.05		0.86	.36	1, 48	2.98	<.10	4.71	<.05	0.67	.42	1, 46	1.12	.29	0.45	.51	0.00	.98		
WL	1, 71	2.37	.13		0.07	.80		0.33	.57	1, 71	1.47	.23	0.35	.56	2.39	.13	1, 74	4.96	<.05	0.04	.84	2.44	.12		
Group										1, 440	8.60	<.01	1.05	.31	1.26	.26									
Controls										1, 1337	6.43	<.05	0.00	.95	0.41	.52									

*Note.* Effects significant at the  $p < .05$  level are in bold font. Subj. Freq/SF=Subject Frequency. Obj. Freq/OF=Object Frequency. Values for group analyses that are not presented here are reported in the main text.

## Appendix C: Stimulus Material for Experiment 2

Table C1

*Target nouns and corresponding frequency ratings*

Noun	Frequency rating
High frequency	( $\geq 30$ occ. per million)
Bear	57.41
Cat	66.33
Ghost	36.59
Horse	92.88
King	129.52
Nurse	44.98
Pig	39.14
Queen	54.69
Low frequency	(< 30 occ. per million)
Clown	15.82
Cow	25.51
Goat	10.53
Maid	22.82
Nun	6.96
Sheep	22.35
Skunk	3.25
Witch	27.65

*Note.* occ = occurrence; Frequency ratings according to the Subtitle Analysis Project (Brysbaert & New, 2009).

Table C2

*Target pictures with corresponding experimental condition*

Related high frequency	Related low frequency	Unrelated high frequency	Unrelated low frequency
The pig is lifting the cat	The goat is licking the cow	The ghost is lifting the horse	The sheep is kicking the witch
The pig is biting the bear	The witch is kissing the nun	The pig is kicking the nurse	The nun is lifting the sheep
The ghost is shooting the king	The sheep is kicking the cow	The king is kicking the cat	The goat is kicking the clown
The queen is pushing the nurse	The nun is kissing the witch	The cat is kicking the king	The clown is kicking the goat
The ghost is washing the queen	The cow is kicking the sheep	The cat is tickling the ghost	The sheep is lifting the nun
The king is shooting the ghost	The maid is lifting the nun	The pig is shooting the queen	The witch is pushing the goat
The horse is chasing the bear	The sheep is brushing the goat	The cat is washing the queen	The witch is kicking the sheep
The cat is lifting the pig	The nun is lifting the maid	The ghost is chasing the pig	The nun is washing the cow
The bear is chasing the horse	The cow is licking the goat	The queen is shooting the pig	The witch is chasing the cow
The horse is kicking the pig	The goat is brushing the sheep	The nurse is kicking the pig	The skunk is tickling the clown
The queen is washing the ghost	The clown is washing the maid	The queen is washing the cat	The nun is tickling the skunk
The nurse is pushing the queen	The maid is washing the clown	The ghost is tickling the cat	The clown is tickling the skunk
The pig is kicking the horse	The sheep is washing the skunk	The king is washing the bear	The cow is chasing the witch
The bear is biting the pig	The skunk is washing the sheep	The bear is washing the king	The skunk is tickling the nun
		The pig is chasing the ghost	The goat is pushing the witch
		The nurse is pushing the bear	The cow is washing the nun
		The horse is lifting the ghost	The maid is chasing the sheep
		The bear is pushing the nurse	The sheep is chasing the maid

## Appendix D: Supplementary Data and Analyses for Experiment 2

Table D1

*Pre-test naming errors for aphasic participants and the three participant groups in the different frequency conditions, collapsed across all three sessions*

		Total errors across all three sessions		Items consistently failed	
	Participant	high	low	high	low
Nonfluent					
	BY	0	1	0	0
	DA	0	1	0	0
	JG	2	2	0	0
	JHM	0	0	0	0
	RP	0	2	0	0
Total					
Fluent					
	NP	1	1	0	0
	STR	1	1	0	0
	SW	2	2	0	0
	WL	0	3	0	0
Total					
Controls					
Total					
		0	2	0	0

Table D2

*Geometric mean pre-test naming latencies for each individual aphasic participant and the three participant groups as a function of frequency condition*

	Geometric mean naming latency in milliseconds	
	High Frequency	Low Frequency
Nonfluent		
BY	1029 (749-1415)	1420 (1033-1953)
DA	1086 (872-1354)	1260 (896-1773)
JG	1576 (778-3191)	1503 (848-2665)
JHM	883 (752-1038)	1059 (878-1277)
RP	1067 (781-1459)	1224 (1044-1435)
Group mean	1109 (796-1545)	1311 (937-1766)
Fluent		
NP	971 (825-1142)	1254 (948-1658)
STR	1458 (1117-1901)	1619 (1203-2178)
SW	1387 (1182-1626)	1656 (1246-2202)
WL	1199 (865-1664)	1487 (939-2357)
Group mean	1251 (983-1592)	1535 (1134-2077)
Controls		
Group mean	765 (614-916)	829 (689-967)

*Note.* Figures in parentheses represent values one standard deviation above and below the mean, respectively (calculated from logged reaction time values, then presented in antilogged form).

Table D3

*Naming latency in the pre-test as a function of frequency for participants with aphasia*

Participant	<i>df</i>	<i>F</i>	<i>p</i>
Nonfluent			
BY	1, 28	5.36	<b>&lt;.05</b>
DA	1, 28	3.47	<.10
JG	1, 26	0.04	.84
JHM	1, 30	7.24	<b>&lt;.05</b>
RP	1, 29	2.68	.11
Fluent			
NP	1, 29	6.81	<b>&lt;.05</b>
STR	1, 30	2.68	.11
SW	1, 23	3.31	<.10
WL	1, 27	2.30	.13

*Note.* Boldfaced effects are significant at the  $p < .05$  level.



Table D4

*Individual error analyses in the sentence production task for participants with aphasia*

Participant	Errors in Context			Sentence Errors						Combined Noun Errors							
	Error Position			Relatedness		Frequency		Rel*Freq		Relatedness		Frequency		Rel*Freq			
	<i>df</i>	$\chi^2$	<i>p</i>	<i>df</i>	$\chi^2$	<i>p</i>	$\chi^2$	<i>p</i>	$\chi^2$	<i>p</i>	<i>df</i>	$\chi^2$	<i>p</i>	$\chi^2$	<i>p</i>	$\chi^2$	<i>p</i>
Nonfluent																	
BY	1	8.89	<.01	1	0.55	.46	3.65	<.10	0.58	.45	1	0.85	.36	3.03	<.10	0.07	.79
DA	1	0.46	.50	1	0.49	.49	0.50	.48	0.46	.50	1	0.17	.68	0.24	.63	0.91	.34
JG	1	3.30	<.10	1	3.66	<.10	1.69	.19	3.51	<.10	1	1.50	.22	4.98	<.05	0.00	.96
JHM	1	3.76	<.10	1	1.00	.32	0.06	.80	0.19	.66	1	0.01	.91	0.64	.42	0.00	.96
RP	1	2.03	.15	1	1.46	.23	1.35	.25	0.62	.43	1	0.42	.52	2.53	.11	0.94	.33
Fluent																	
NP	1	8.99	<.01	1	0.00	.97	0.06	.81	0.25	.62	1	0.82	.36	0.16	.69	0.13	.72
STR	1	3.42	<.10	1	0.82	.37	1.17	.19	0.31	.58	1	0.49	.49	0.08	.37	0.04	.84
SW	1	0.01	.91	1	0.99	.32	0.19	.66	0.01	.91	1	1.36	.24	0.97	.33	0.02	.89
WL	1	8.01	<.01	1	10.93	<.001	0.18	.67	0.27	.60	1	10.52	<.01	0.12	.73	0.06	.81

*Note.* Effects significant at the  $p < .05$  level are in bold font. Rel = Relatedness. Freq = Frequency.

Table D5

*Supplementary details for the group error analyses in the sentence production task for each of the three participant groups and across groups*

Groups	Sentence Errors							Combined Noun Errors						
	Frequency			Rel*Freq		Group*Freq		Frequency			Rel*Freq		Group*Freq	
	<i>df</i>	$\chi^2$	<i>p</i>	$\chi^2$	<i>p</i>	$\chi^2$	<i>p</i>	<i>df</i>	$\chi^2$	<i>p</i>	$\chi^2$	<i>p</i>	$\chi^2$	<i>p</i>
Nonfluent	1	0.07	.79	1.15	.28			1	0.70	.40	0.12	.73		
Fluent	1	0.21	.65	0.48	.49			1	0.56	.45	0.18	.67		
Controls	1	8.15	<.01	1.14	.29			1	0.35	.55	0.23	.63		
Nonfluent vs. Controls	1					2.28	.13	1					0.01	.93
Nonfluent vs. Fluent	1					0.07	.79	1					1.26	.26
Fluent vs. Controls	1					2.11	.15	1					1.04	.31

*Note.* Effects significant at the  $p < .05$  level are in bold font. Rel=Relatedness. Freq = Frequency.

Table D6

*Geometric means of participants with aphasia and the three participant groups for the different latency measures in the different conditions*

Participant	Sentence Initiation Time				Subject-Noun Initiation Time				Verb-object Noun Initiation Time			
	Relatedness		Frequency		Relatedness		Frequency		Relatedness		Frequency	
	Rel	Un	high	low	Rel	Un	high	low	Rel	Un	high	low
Nonfluent												
BY	2970 (1930-4570)	3094 (1993-4803)	2784 (1900-4081)	3555 (2095-5374)	3585 (2240-5738)	3652 (2273-5866)	3199 (2131-4803)	4167 (2524-6878)	681 (286-1624)	735 (291-1860)	566 (239-1338)	933 (389-2240)
DA	3483 (2236-5426)	3654 (2603-5130)	3531 (2463-5061)	3620 (2389-5484)	3982 (2733-5803)	4375 (3029-6319)	4034 (2748-5922)	4354 (3033-6251)	1437 (716-2885)	1935 (1003-3732)	1472 (750-2891)	1930 (977-3812)
JG	2999 (1805-4984)	2908 (1976-4280)	2876 (2011-4113)	3004 (1827-4939)	3312 (2050-5595)	3152 (1978-4987)	3177 (1961-5250)	3304 (2041-5264)	834 (386-1804)	829 (330-2085)	926 (326-2631)	763 (387-1503)
JHM	2230 (1525-3260)	2188 (1592-3009)	2174 (1493-3166)	2237 (1629-3071)	2494 (1694-3674)	2439 (1788-3346)	2440 (1659-3587)	2486 (1820-3395)	1240 (684-2248)	1005 (511-1975)	1036 (516-2080)	1173 (647-2126)
RP	7150 (4530-11285)	7084 (4653-10784)	6628 (4207-10443)	7233 (4611-11345)	7585 (4750-12112)	7515 (4969-11367)	6993 (4393-11130)	7644 (4850-12048)	1598 (712-3586)	1383 (643-2977)	1296 (564-2980)	1594 (748-3396)
Group mean	3481 (2248-5392)	3479 (244-5393)	3326 (2182-5074)	3625 (2315-5707)	3911 (2598-5888)	3917 (2566-5979)	3712 (2497-5518)	4124 (2666-6379)	1103 (765-1590)	1109 (741-1109)	1010 (695-1468)	1222 (825-1810)

Table D6 continued.

Participant	Sentence Initiation Time				Subject-Noun Initiation Time				Verb-object Noun Initiation Time			
	Relatedness		Frequency		Relatedness		Frequency		Relatedness		Frequency	
	Rel	Un	high	low	Rel	Un	high	low	Rel	Un	high	low
Fluent												
NP	2369 (1799-3120)	2402 (1719-3316)	2242 (1709-2939)	2533 (1819-3527)	2695 (2061-3523)	2736 (1984-3772)	2551 (1965-3311)	2886 (2098-3970)	739 (361-1513)	661 (294-1487)	631 (353-1131)	761 (308-1881)
STR	2311 (1848-2890)	2444 (1786-3344)	2299 (1715-3082)	2479 (1918-3203)	2647 (2025-3462)	2900 (2000-4204)	2762 (1897-4021)	2812 (2127-3718)	425 (230-785)	489 (268-891)	449 (246-819)	472 (254-876)
SW	2867 (1860-3690)	2909 (2207-4000)	2862 (1841-3674)	2924 (2194-3986)	3035 (2013-3903)	3078 (2351-4090)	3049 (1996-3752)	3069 (2345-4183)	569 (277-1103)	596 (297-1391)	520 (267-1203)	672 (303-1284)
WL	2356 (1717-2995)	2390 (1640-3079)	2411 (1614-3148)	2332 (1759-2907)	3778 (2314-6169)	3406 (2135-5433)	3657 (2195-6093)	3525 (2250-5524)	1372 (535-3523)	1597 (566-4513)	1261 (492-3231)	1752 (631-4863)
Group mean	2466 (2229-2728)	2527 (2300-2777)	2442 (2187-2726)	2558 (2324-2815)	3007 (2553-3542)	3020 (2758-3316)	2977 (2549-3477)	3061 (2767-3387)	704 (427-1160)	745 (441-1257)	657 (416-1037)	806 (462-1405)
Controls												
Group mean	1167 (1031-1320)	1131 (1003-1275)	1132 (1008-1270)	1175 (1028-1344)	1313 (1180-1461)	1272 (1142-1416)	1272 (1145-1414)	1321 (1183-1476)	174 (156-193)	162 (141-186)	162 (146-180)	185 (161-212)

*Note.* Figures in parentheses represent values one standard deviation above and below the mean, respectively (calculated from logged reaction time values, then presented in antilogged form).

Table D7

*Noun relatedness and frequency effects on sentence latency measures for participants with aphasia and supplementary details for the three participant groups*

Participant	Sentence Initiation Times							Subject-Noun Initiation Times							Verb-object Noun Initiation Times						
	Rel		Freq		R*F			Rel		Freq		R*F			Rel		Freq		R*F		
	<i>df</i>	<i>F</i>	<i>p</i>	<i>F</i>	<i>p</i>	<i>F</i>	<i>p</i>	<i>df</i>	<i>F</i>	<i>p</i>	<i>F</i>	<i>p</i>	<i>F</i>	<i>p</i>	<i>df</i>	<i>F</i>	<i>p</i>	<i>F</i>	<i>p</i>	<i>F</i>	<i>p</i>
Nonfluent																					
BY	1, 88	0.02	.90	6.98	<.01	0.08	.78	1, 88	0.05	.83	11.98	<.001	0.07	.79	1, 86	0.10	.76	8.77	<.01	0.13	.72
DA	1,108	0.40	.53	0.64	.43	1.13	.29	1,107	4.12	<.05	1.02	.31	0.25	.62	1,105	0.02	.88	10.3	<.01	0.12	.73
JG	1, 61	0.27	.60	0.04	.84	2.79	.10	1,61	0.56	.46	4.92	<.05	3.11	<.10	1, 63	0.11	.74	0.87	.35	1.91	.17
JHM	1,104	0.23	.63	0.19	.67	0.01	.91	1,104	0.35	.55	0.07	.79	0.05	.83	1,101	3.05	<.10	4.15	<.05	3.71	<.10
RP	1, 67	0.04	.85	3.12	<.10	0.19	.66	1, 67	0.03	.86	6.00	<.05	0.79	.38	1, 69	0.59	.45	0.21	.65	0.75	.39
Group	1, 684			5.03	<.05	1.86	.17	1, 684	0.02	.88	6.80	<.01	0.03	.87	1, 676			7.57	<.01	1.82	.18
Fluent																					
NP	1, 73	0.03	.86	4.50	<.05	0.30	.58	1, 72	0.05	.83	4.15	<.05	0.34	.56	1, 73	1.53	.47	1.49	.22	1.25	.27
STR	1, 86	7.42	<.01	7.96	<.01	1.55	.22	1, 85	5.44	<.05	4.05	<.05	1.24	.27	1, 87	1.24	.27	0.01	.91	0.20	.65
SW	1, 44	0.17	.68	0.16	.67	0.11	.75	1, 45	0.07	.81	0.06	.81	0.00	.96	1, 44	0.03	.86	4.52	<.05	2.26	.14
WL	1, 79	0.00	.97	0.08	.79	0.78	.38	1, 79	1.01	.32	0.42	.52	0.43	.51	1, 74	0.34	.56	5.79	<.05	0.12	.73
Group	1, 461			1.80	.18	1.41	.24	1, 462	0.01	.94	0.15	.70	1.06	.30	1, 458			4.76	<.05	0.02	.88
Controls	1, 1334			2.00	.16	0.72	.39	1, 1338	0.53	.47	0.79	.37	0.67	.41	1, 1295			26.84	<.001	0.08	.78

*Note.* Effects significant at the  $p < .05$  level are in bold font. Rel/R=Relatedness. Freq/F= Frequency. Values for group analyses that are not presented here are reported in the main text.

## References

- Abuom, T. O., & Bastiaanse, R. (2013). Production and comprehension of reference of time in Swahili-English bilingual agrammatic speakers. *Aphasiology*, 27(2), 157-177. doi: 10.1080/02687038.2012.699632
- Ahrens, K. (2003). Verbal integration: The interaction of participant roles and sentential arguments. *Journal of Psycholinguistic Research*, 32(5), 497-516. doi: 10.1023/A:1025452814385
- Akaike, H. (1976). An information criterion (AIC). *Mathematical Sciences*, 14(153), 5-9.
- Allum, P. H., & Wheeldon, L. R. (2007). Planning scope in spoken sentence production: The role of grammatical units. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(4), 791. doi: 10.1037/0278-7393.33.4.791
- Amunts, K., Schleicher, A., Bürgel, U., Mohlberg, H., Uylings, H., & Zilles, K. (1999). Broca's region revisited: Cytoarchitecture and intersubject variability. *Journal of Comparative Neurology*, 412(2), 319-341. doi: 10.1002/(SICI)1096-9861(19990920)412:2<319::AID-CNE10>3.0.CO;2-7
- Ashburner, J., Barnes, G., Chen, C., Daunizeau, J., Flandin, G., Friston, K., . . . Phillips, C. (2008). SPM8 manual. *Functional Imaging Laboratory, Institute of Neurology*.
- Baars, B. J., Motley, M. T., & MacKay, D. G. (1975). Output editing for lexical status in artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behaviour*, 14(4), 382-391. doi: 10.1016/S0022-5371(75)80017-X
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390-412. doi: 10.1016/j.jml.2007.12.005

- Baldo, J. V., & Dronkers, N. F. (2007). Neural correlates of arithmetic and language comprehension: A common substrate? *Neuropsychologia*, 45(2), 229-235. doi: 10.1016/j.neuropsychologia.2006.07.014
- Baldo, J. V., Wilkins, D. P., Ogar, J., Willock, S., & Dronkers, N. F. (2011). Role of the precentral gyrus of the insula in complex articulation. *Cortex*, 47(7), 800-807. doi: 10.1016/j.cortex.2010.07.001
- Baldo, J. V., Wilson, S. M., & Dronkers, N. F. (2012). Uncovering the neural substrates of language: A voxel-based lesion symptom mapping approach. In M. Faust (Ed.), *Advances in the neural substrates of language: Toward a synthesis of basic science and clinical research*. Oxford: Wiley-Blackwell.
- Barker-Collo, S. L. (2001). The 60-item Boston naming test: Cultural bias and possible adaptations for New Zealand. *Aphasiology*, 15(1), 85-92. doi: 10.1080/02687040042000124
- Barry, C., Morrison, C. M., & Ellis, A. W. (1997). Naming the Snodgrass and Vanderwart pictures: Effects of age of acquisition, frequency, and name agreement. *The quarterly Journal of Experimental Psychology*, 50A(3), 560-585. doi: 10.1080/783663595
- Bastiaanse, R. (1995). Broca's aphasia: A syntactic and/or a morphological disorder? A case study. *Brain and Language*, 48(1), 1-32. doi: 10.1006/brln.1995.1001
- Bastiaanse, R. (2008). Production of verbs in base position by Dutch agrammatic speakers: Inflection versus finiteness. *Journal of Neurolinguistics*, 21(2), 104-119.
- Bastiaanse, R., & Jonkers, R. (2012). Linguistic accounts of agrammatic aphasia. In R. Bastiaansen & C. Thompson (Eds.), *Perspectives on agrammatism* (pp. 17-33). London: Psychology Press.

- Bastiaanse, R., & Thompson, C. K. (2003). Verb and auxiliary movement in agrammatic Broca's aphasia. *Brain and Language*, 84(2), 286-305. doi: 10.1016/S0093-934X(02)00553-9
- Bastiaanse, R., & van Zonneveld, R. (2005). Sentence production with verbs of alternating transitivity in agrammatic Broca's aphasia. *Journal of Neurolinguistics*, 18(1), 57-66. doi: 10.1016/j.jneuroling.2004.11.006
- Bates, E., Wilson, S. M., Saygin, A. P., Dick, F., Sereno, M. I., Knight, R. T., & Dronkers, N. F. (2003). Voxel-based lesion-symptom mapping. *Nature Neuroscience*, 6(5), 448-450. doi: 10.1038/nn1050
- Beeke, S., Maxim, J., & Wilkinson, R. (2008). Rethinking agrammatism: Factors affecting the form of language elicited via clinical test procedures. *Clinical linguistics & phonetics*, 22(4-5), 317-323. doi: 10.1080/02699200801918911
- Beeke, S., Wilkinson, R., & Maxim, J. (2003). Exploring aphasic grammar 2: Do language testing and conversation tell a similar story? *Clinical linguistics & phonetics*, 17(2), 109-134. doi: 10.1080/0269920031000061786
- Beeke, S., Wilkinson, R., & Maxim, J. (2007). Individual variation in agrammatism: A single case study of the influence of interaction. *International Journal of Language & Communication Disorders*, 42(6), 629-647. doi: 10.1080/13682820601160087
- Bencini, G., & Roland, D. (1996). *Verb access difficulties in agrammatic aphasic narratives*. Paper presented at the 70th Annual Meeting of the Linguistic Society of America, San Diego, CA.
- Berndt, R. S. (2001). More than just words: Sentence production in aphasia. In R. S. Berndt (Ed.), *Handbook of Neuropsychology* (2 ed., Vol. 3, pp. 173-187). Amsterdam: Elsevier.



- Berndt, R. S., & Haendiges, A. N. (2000). Grammatical class in word and sentence production: Evidence from an aphasic patient. *Journal of Memory and Language*, 43(2), 249-273.
- Berndt, R. S., Haendiges, A. N., & Wozniak, M. A. (1997). Verb retrieval and sentence processing: Dissociation of an established symptom association. *Cortex*, 33(1), 99-114. doi: 10.1016/S0010-9452(97)80007-X
- Berndt, R. S., Mitchum, C. C., Haendiges, A. N., & Sandson, J. (1997). Verb retrieval in aphasia. 1. Characterizing single word impairments. *Brain and Language*, 56(1), 68-106.
- Biegler, K. A., Crowther, J. E., & Martin, R. C. (2008). Consequences of an inhibition deficit for word production and comprehension: Evidence from the semantic blocking paradigm. *Cognitive Neuropsychology*, 25(4), 493-527. doi: 10.1080/02643290701862316
- Bloem, I., & La Heij, W. (2003). Semantic facilitation and semantic interference in word translation: Implications for models of lexical access in language production. *Journal of Memory and Language*, 48(3), 468-488. doi: 10.1016/S0749-596X(02)00503-X
- Bloem, I., van den Boogaard, S., & La Heij, W. (2004). Semantic facilitation and semantic interference in language production: Further evidence for the conceptual selection model of lexical access. *Journal of Memory and Language*, 51(2), 307-323. doi: 10.1016/j.jml.2004.05.001
- Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review*, 89(1), 1.
- Bock, J. K. (1986a). Meaning, sound, and syntax: Lexical priming in sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12(4), 575. doi: 10.1037/0278-7393.12.4.575

- Bock, J. K. (1986b). Syntactic persistence in language production. *Cognitive psychology*, 18(3), 355-387. doi: 10.1016/0010-0285(86)90004-6
- Bock, J. K. (1987a). Coordinating words and syntax in speech plans. *Progress in Psychology of Language*, 3, 337-390.
- Bock, J. K. (1987b). An effect of the accessibility of word forms on sentence structures. *Journal of Memory and Language*, 26(2), 119-137. doi: 10.1016/0749-596X(87)90120-3
- Bock, J. K., & Cutting, J. C. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language*, 31(1), 99-127. doi: 10.1016/0749-596X(92)90007-K
- Bock, J. K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General*, 129(2), 177. doi: 10.1037/0096-3445.129.2.177
- Bock, J. K., & Levelt, W. J. M. (1994). Language production. Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945-984). San Diego, CA.
- Borovsky, A., Saygin, A. P., Bates, E., & Dronkers, N. F. (2007). Lesion correlates of conversational speech production deficits. *Neuropsychologia*, 45(11), 2525-2533. doi: 10.1016/j.neuropsychologia.2007.03.023
- Bradley, D. C., Garrett, M. F., Kean, M.-L., Kolk, H., & Zurif, E. (1979). Word recognition in agrammatic aphasia. *Brain and Language*.
- Bradley, D. C., Garrett, M. F., & Zurif, E. B. (1980). Syntactic deficits in Broca's aphasia. In D. Caplan (Ed.), *Biological Studies of Mental Processes* (pp. 269-286). Cambridge, MA: MIT Press.
- Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition*, 75(2), B13-B25. doi: 10.1016/S0010-0277(99)00081-5

- Breedin, S. D., & Saffran, E. M. (1999). Sentence processing in the face of semantic loss: A case study. *Journal of Experimental Psychology: General*, 128(4), 547. doi: 10.1037/0096-3445.128.4.547
- Breedin, S. D., Saffran, E. M., & Schwartz, M. F. (1998). Semantic factors in verb retrieval: An effect of complexity. *Brain and Language*, 63(1), 1-31. doi: 10.1006/brln.1997.1923
- Breen, K., & Warrington, E. K. (1994). A study of anomia: Evidence for a distinction between nominal and propositional language. *Cortex*, 30(2), 231-245. doi: 10.1016/S0010-9452(13)80195-5
- Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, 41(4), 977-990. doi: 10.3758/BRM.41.4.977
- Burchert, F., Meißner, N., & De Bleser, R. (2008). Production of non-canonical sentences in agrammatic aphasia: Limits in representation or rule application? *Brain and Language*, 104(2), 170-179. doi: 10.1016/j.bandl.2007.06.004
- Burchert, F., Swoboda-Moll, M., & De Bleser, R. (2005). Tense and agreement dissociations in German agrammatic speakers: Underspecification vs. hierarchy. *Brain and Language*, 94(2), 188-199. doi: 10.1016/j.bandl.2004.12.006
- Caramazza, A., & Hillis, A. E. (1989). The disruption of sentence production: Some dissociations. *Brain and Language*, 36(4), 625-650. doi: 10.1016/0093-934X(89)90091-6
- Caramazza, A., & McCloskey, M. (1988). The case for single-patient studies. *Cognitive Neuropsychology*, 5(5), 517-527. Retrieved from <http://www.tandfonline.com/loi/pcng20>

- Carroll, J. B., & White, M. N. (1973). Age-of-acquisition norms for 220 picturable nouns. *Journal of Verbal Learning and Verbal Behaviour*, 12, 563-576. doi: 10.1016/S0022-5371(73)80036-2
- Caspers, S., Geyer, S., Schleicher, A., Mohlberg, H., Amunts, K., & Zilles, K. (2006). The human inferior parietal cortex: Cytoarchitectonic parcellation and interindividual variability. *NeuroImage*, 33(2), 430-448. doi: 10.1016/j.neuroimage.2006.06.054
- Chang, F., Dell, G. S., & Bock, J. K. (2006). Becoming syntactic. *Psychological review*, 113(2), 234-277. doi: 10.1037/0033-295X.113.2.234
- Chen, S., & Bates, E. (1998). The dissociation between nouns and verbs in Broca's and Wernicke's aphasia: Findings from Chinese. *Aphasiology*, 12(1), 5-36. doi: 10.1080/02687039808249441
- Cleland, A. A., & Pickering, M. J. (2003). The use of lexical and syntactic information in language production: Evidence from the priming of noun-phrase structure. *Journal of Memory and Language*, 49(2), 214-230. doi: 10.1016/S0749-596X(03)00060-3
- Cloutman, L. L. (2005). *The effects of semantic context on word retrieval performance in aphasia*. (Doctoral Dissertation, Victoria University of Wellington, New Zealand).
- Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers*, 25, 257-271. Retrieved from <http://link.springer.com/article/10.3758/BF03204507>
- Crawford, J. R. (2004). Psychometric foundations of neuropsychological assessment. In L. H. Goldstein & J. E. McNeil (Eds.), *Clinical neuropsychology: A practical guide to assessment and management for clinicians* (pp. 121-140). Chichester, UK: Wiley.

- Crawford, J. R., Garthwaite, P. H., & Gray, C. D. (2003). Wanted: Fully operational definitions in single-case studies. *Cortex*, 39(2), 357-370. doi: 10.1016/S0010-9452(08)70117-5
- Damian, M. F., & Dumay, N. (2007). Time pressure and phonological advance planning in spoken production. *Journal of Memory and Language*, 57(2), 195-209. doi: 10.1016/j.jml.2006.11.001
- de Roo, E., Kolk, H. H. J., & Hofstede, B. (2003). Structural properties of syntactically reduced speech: A comparison of normal speakers and Broca's aphasics. *Brain and Language*, 86(1), 99-115. doi: 10.1016/S0093-934X(02)00538-2
- Deese, J. (1978). Thought into speech: Linguistic rules and psychological limitations in processing information determine how we put our ideas into words. *American Scientist*, 66(3), 314-321. Retrieved from <http://www.jstor.org/>
- Deese, J. (1980). Pauses, prosody, and the demands of production in language. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech, studies in honour of Frieda Goldman-Eisler*. De Gruyter.
- Del Viso, S., Igoa, J. M., & Garcia-Albea, J. E. (1991). On the autonomy of phonological encoding: Evidence from slips of the tongue in Spanish. *Journal of Psycholinguistic Research*, 20(3), 161-185. doi: 10.1007/BF01067213
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93(3), 283-321. doi: 10.1037/0033-295X.93.3.283
- Dell, G. S., Oppenheim, G. M., & Kittredge, A. K. (2008). Saying the right word at the right time: Syntagmatic and paradigmatic interference in sentence production. *Language and cognitive processes*, 23(4), 583-608. doi: 10.1080/01690960801920735

- Dell, G. S., & Reich, P. A. (1981). Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behaviour*, 20(6), 611-629. doi: 10.1016/S0022-5371(81)90202-4
- Diggle, P. J. (1988). An approach to the analysis of repeated measurements. *Biometrics*, 959-971. Retrieved from <http://www.jstor.org/stable/2531727>
- Donovan, K., Siegert, R., McDowall, J., & Abernethy, D. (1999). Clustering and switching in verbal fluency in Parkinson's disease. *New Zealand Journal of Psychology*, 28, 61-66. Retrieved from [http://www.psychology.org.nz/NZJP\\_TOC#NZJP-1999](http://www.psychology.org.nz/NZJP_TOC#NZJP-1999)
- Dragoy, O., & Bastiaanse, R. (2010). Verb production and word order in Russian agrammatic speakers. *Aphasiology*, 24(1), 28-55. doi: 10.1080/02687030802586902
- Dragoy, O., Stowe, L. A., Bos, L. S., & Bastiaanse, R. (2012). From time to time: Processing time reference violations in Dutch. *Journal of Memory and Language*, 66(1), 307-325. doi: 10.1016/j.jml.2011.09.001
- Dronkers, N. F. (1996). A new brain region for coordinating speech articulation. *Nature*, 384(6605), 159-161. doi: 10.1038/384159a0
- Dronkers, N. F., Baldo, J. V., Wilkins, D. P., Ludy, C. A., Arevalo, A., & Knight, R. (2009). *Lesion localization of chronic aphasia syndromes*. Paper presented at the 47th Meeting of the Academy of Aphasia, Boston.
- Dronkers, N. F., & Larsen, J. (2001). Neuroanatomy of the classical syndromes of aphasia. In R. S. Berndt (Ed.), *Handbook of Neuropsychology* (2 ed., Vol. 3, pp. 19-30).
- Dronkers, N. F., & Ludy, C. A. (1998). Brain lesion analysis in clinical research. In H. A. Whitaker & B. Stemmer (Eds.), *Handbook of Neurolinguistics* (pp. 172-187). Academic Press.

- Dronkers, N. F., Wilkins, D. P., Van Valin Jr., R. D., Redfern, B. B., & Jaeger, J. J. (2004). Lesion analysis of the brain areas involved in language comprehension. *Cognition*, 92(1-2), 145-177. doi: 10.1016/j.cognition.2003.11.002
- Druks, J., & Masterson, J. (2000). *An object and action naming battery*. Hove: Psychology Press.
- Eickhoff, S. B., Amunts, K., Mohlberg, H., & Zilles, K. (2006). The human parietal operculum. II. Stereotaxic maps and correlation with functional imaging results. *Cerebral Cortex*, 16(2), 268-279. doi: 10.1093/cercor/bhi106
- Eickhoff, S. B., Schleicher, A., Zilles, K., & Amunts, K. (2006). The human parietal operculum. I. Cytoarchitectonic mapping of subdivisions. *Cerebral Cortex* 16(2), 254-267. doi: 10.1093/cercor/bhi105
- Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., & Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage*, 25(4), 1325-1335. doi: 10.1016/j.neuroimage.2004.12.034
- Faroqi-Shah, Y., & Thompson, C. K. (2003). Effect of lexical cues on the production of active and passive sentences in Broca's and Wernicke's aphasia. *Brain and language*, 85(3), 409-426. doi: 10.1016/S0093-934X(02)00586-2
- Faust, M. E., Balota, D. A., Spieler, D. H., & Ferraro, F. R. (1999). Individual differences in information-processing rate and amount: Implications for group differences in response latency. *Psychological Bulletin*, 125(6), 777-799. doi: 10.1037/0033-2909.125.6.777
- Ferreira, F. (1991). Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language*, 30(2), 210-233. doi: 10.1016/0749-596X(91)90004-4

- Ferreira, F., & Engelhardt, P. E. (2006). Syntax and production. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (pp. 61-91). Academic Press.
- Ferreira, F., & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, 46(1), 57-84. doi: 10.1006/jmla.2001.2797
- Ferreira, V. S., & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology*, 40(4), 296-340. doi: 10.1006/cogp.1999.0730
- Ferreira, V. S., & Firato, C. E. (2002). Proactive interference effects on sentence production. *Psychonomic Bulletin & Review*, 9(4), 795-800. doi: 10.3758/BF03196337
- Ford, M. (1982). *Sentence planning units: Implications for the speaker's representation of meaningful relations underlying sentences*. Cambridge, MA: MIT press.
- Ford, M., & Holmes, V. M. (1978). Planning units and syntax in sentence production. *Cognition*, 6(1), 35-53. doi: 10.1016/0010-0277(78)90008-2
- 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(2), 245-265. doi: 10.1080/02643290342000528
- Friedmann, N. (2006). Speech production in Broca's agrammatic aphasia: Syntactic tree pruning. In K. Amunts & Y. Grodzinsky (Eds.), *Broca's region*, 63-82. Oxford University Press.
- Friedmann, N., & Grodzinsky, Y. (1997). Tense and agreement in agrammatic production: Pruning the syntactic tree. *Brain and Language*, 56(3), 397-425. doi: 10.1006/brln.1997.1795



- Fromkin, V. A. (1971). The non-anomalous nature of anomalous utterances. *Language*, 47(1), 27-52. Retrieved from <http://www.stanford.edu/~zwicky/Fromkin-non-anomalous.pdf>
- Garrett, M. F. (1975). The analysis of sentence production. In G. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 9, pp. 133-177). 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 Linguistic Series, 30.
- Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production* (pp. 177-220). London: Academic Press.
- Garrett, M. F. (1982). Production of speech: Observations from normal and pathological language use. In A. W. Ellis (Ed.), *Normality and pathology in cognitive functions* (pp. 19-76). London: American Press.
- Garrett, M. F. (1984). The organization of processing structure for language production. In D. Caplan, A. R. Lecours & A. Smith (Eds.), *Biological Perspectives on Language* (pp. 172-193). MIT Press.
- Geyer, S. (2003). *The microstructural border between the motor and the cognitive domain in the human cerebral cortex*. Wien: Springer.
- Geyer, S., Ledberg, A., Schleicher, A., Kinomura, S., Schormann, T., Bürgel, U., . . . Roland, P. E. (1996). Two different areas within the primary motor cortex of man. *Nature*, 382, 805-807. doi: 10.1038/382805a0
- Geyer, S., Schleicher, A., & Zilles, K. (1999). Areas 3a, 3b, and 1 of human primary somatosensory cortex: 1. Microstructure; organization and interindividual variability. *NeuroImage*, 10(1), 63-83. doi: 10.1006/nimg.1999.0440

- Geyer, S., Schormann, T., Mohlberg, H., & Zilles, K. (2000). Areas 3a, 3b and 1 of human primary somatosensory cortex: 2. Spatial normalization to standard anatomical space. *NeuroImage*, 11(6), 684-696. doi: 10.1006/nimg.2000.0548
- Goldman-Eisler, F., & Cohen, M. (1970). Is N, P, and PN difficulty a valid criterion of transformational operations? *Journal of Verbal Learning and Verbal Behaviour*, 9(2), 161-166. doi: 10.1016/S0022-5371(70)80045-7
- Goodglass, H. (1993). *Understanding aphasia*. Academic Press.
- Goodglass, H., Christiansen, J. A., & Gallagher, R. E. (1994). Syntactic constructions used by agrammatic speakers: Comparison with conduction aphasics and normals. *Neuropsychology*, 8(4), 598-613. doi: 10.1037/0894-4105.8.4.598
- Goodglass, H., Kaplan, E., & Barresi, B. (2001). *The assessment of aphasia and related disorders*. Lippincott Williams & Wilkins.
- Gordon, J. K., & Dell, G. S. (2003). Learning to divide the labor: An account of deficits in light and heavy verb production. *Cognitive Science*, 27(1), 1-40. doi: 10.1016/S0364-0213(02)00111-8
- Grefkes, C., Geyer, S., Schormann, T., Roland, P. E., & Zilles, K. (2001). Human somatosensory area 2: Observer-independent cytoarchitectonic mapping, interindividual variability, and population map. *NeuroImage*, 14(3), 617-631. doi: 10.1006/nimg.2001.0858
- Griffin, Z. M. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition*, 82(1), B1-B14. doi: 10.1016/S0010-0277(01)00138-X
- Griffin, Z. M., & Bock, J. K. (2000). What the eyes say about speaking. *Psychological Science*, 11(4), 247-279. doi: 10.1111/1467-9280.00255
- Hartsuiker, R. J., & Kolk, H. H. J. (1998). Syntactic facilitation in agrammatic sentence production. *Brain and Language*, 62(2), 221-254. doi: 10.1006/brln.1997.1905

- Heeschen, C., & Schegloff, E. A. (1999). Agrammatism, adaptation theory, conversation analysis: On the role of so-called telegraphic style in talk-in-interaction. *Aphasiology* 13(4-5), 365-405. doi: 10.1080/026870399402145
- Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences*, 4(4), 131-138. doi: 10.1016/S1364-6613(00)01463-7
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8(5), 393-402. doi: 10.1038/nrn2113
- Hofstede, B. T. M., & Kolk, H. H. J. (1994). The effects of task variation on the production of grammatical morphology in Broca's aphasia: A multiple case study. *Brain and Language*, 46(2), 278-328. doi: 10.1006/brln.1994.1017
- Holmes, C. J., Hoge, R., Collins, L., Woods, R., Toga, A. W., & Evans, A. C. (1998). Enhancement of MR images using registration for signal averaging. *Journal of Computer Assisted Tomography*, 22(2), 324-333. Retrieved from <http://journals.lww.com/jcat/pages/default.aspx>
- 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. doi: 10.1080/02643290601181890
- James, C. T., Thompson, J. G., & Baldwin, J. M. (1973). The reconstructive process in sentence memory. *Journal of Verbal Learning and Verbal Behaviour*, 12(1), 51-63. doi: 10.1016/S0022-5371(73)80060-X
- Jarvella, R. J., & Sinnott, J. (1972). Contextual constraints on noun distributions to some English verbs by children and adults. *Journal of Verbal Learning and Verbal Behaviour*, 11(1), 47-53. doi: 10.1016/S0022-5371(72)80058-6

- Kaan, E., & Schwaab, T. (2002). The brain circuitry of syntactic comprehension. *Trends in Cognitive Sciences*, 6(8), 350-356. doi: 10.1016/S1364-6613(02)01947-2
- Kay, J., Lesser, R., & Coltheart, M. (1992). *PALPA: Psycholinguistic assessments of language processing in aphasia*. Psychology Press.
- Kean, M.-L. (1977). The linguistic interpretation of aphasic syndromes: Agrammatism in Broca's aphasia, an example. *Cognition*, 5(1), 9-46. doi: 10.1016/0010-0277(77)90015-4
- Kean, M.-L. (1979). Agrammatism: A phonological deficit? *Cognition*, 7(1), 69-83. doi: 10.1016/0010-0277(79)90012-X
- Kertesz, A. (1982). *Western Aphasia Battery*. New York: Grune & Stratton.
- Kim, M., & Thompson, C. K. (2000). Patterns of comprehension and production of nouns and verbs in agrammatism: Implications for lexical organization. *Brain and Language*, 74(1), 1-25. doi: 10.1006/brln.2000.2315
- Kimberg, D. Y., Coslett, H. B., & Schwartz, M. F. (2007). Power in voxel-based lesion-symptom mapping. *Journal of Cognitive Neuroscience*, 19(7), 1067-1080. doi: 10.1162/jocn.2007.19.7.1067
- Kittredge, A. K., Dell, G. S., Verkuilen, J., & Schwartz, M. F. (2008). Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. *Cognitive Neuropsychology*, 25(4), 463-492. doi: 10.1080/02643290701674851
- Koenigs, M., Acheson, D. J., Barbey, A. K., Solomon, J., Postle, B. R., & Grafman, J. (2011). Areas of left perisylvian cortex mediate auditory-verbal short-term memory. *Neuropsychologia*, 49(13), 3612-3619. doi: 10.1016/j.neuropsychologia.2011.09.013
- Kolk, H. H. J. (1995). A time-based approach to agrammatic production. *Brain and Language*, 50(3), 282-303. doi: 10.1006/brln.1995.1049

- Kolk, H. H. J. (2006). How language adapts to the brain: An analysis of agrammatic aphasia. In L. Progovac (Ed.), *The syntax of nonsententials: Multi-disciplinary perspectives* (pp. 229-258). London: John Benjamins.
- Kolk, H. H. J. (2007). Variability is the hallmark of aphasic behaviour: Grammatical behaviour is no exception. *Brain and language*, 101(2), 99-102. doi: 10.1016/j.bandl.2007.04.002
- Kolk, H. H. J., & Van Grunsven, M. M. F. (1985). Agrammatism as a variable phenomenon. *Cognitive Neuropsychology*, 2(4), 347-384. doi: 10.1080/02643298508252666
- Konopka, A. E., & Bock, J. K. (2009). Lexical or syntactic control of sentence formulation? Structural generalizations from idiom production. *Cognitive Psychology*, 58(1), 68-101. doi: 10.1016/j.cogpsych.2008.05.002
- Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33(2), 149-174. doi: 10.1006/jmla.1994.1008
- Kurth, F., Eickhoff, S. B., Schleicher, A., Hoemke, L., Zilles, K., & Amunts, K. (2010). Cytoarchitectonic and probabilistic maps of the human posterior insular cortex. *Cerebral Cortex*, 20(6), 1448-1461. doi: 10.1093/cercor/bhp208
- Lee, J., Milman, L. H., & Thompson, C. K. (2008). Functional category production in English agrammatism. *Aphasiology*, 22(7-8), 893-905. doi: 10.1080/02687030701865670
- Lee, M., & Thompson, C. K. (2004). Agrammatic aphasic production and comprehension of unaccusative verbs in sentence contexts. *Journal of Neurolinguistics*, 17(4), 315-330. doi: 10.1016/S0911-6044(03)00062-9
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

- Levelt, W. J. M., & Kelter, S. (1982). Surface form and memory in question answering. *Cognitive Psychology*, 14(1), 78-106. doi: 10.1016/0010-0285(82)90005-6
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22(1), 1-38.
- Liang, K. Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1), 13-22. doi: 10.1093/biomet/73.1.13
- Linebarger, M. C., McCall, D., Virata, T., & Berndt, R. S. (2007). Widening the temporal window: Processing support in the treatment of aphasiac language production. *Brain and Language*, 100, 53-68. doi: 10.1016/j.bandl.2006.09.001
- Linebarger, M. C., Schwartz, M. F., Romania, J. R., Kohn, S. E., & Stephens, D. L. (2000). Grammatical encoding in aphasia: Evidence from a “processing prosthesis”. *Brain and Language*, 75(3), 416-427. doi: 10.1006/brln.2000.2378
- Loonstra, A. S., Tarlow, A. R., & Sellers, A. H. (2001). COWAT metanorms across age, education, and gender. *Applied neuropsychology*, 8(3), 161-166. doi: 10.1207/S15324826AN0803\_5
- Lukkien, C. L., Burmester, B. B., & Wilshire, C. E. (2012). The sentence production test for aphasia (manuscript submitted for publication).
- Mackay, J., & Mensah, G. A. (2004). Global burden of stroke. In J. Mackay, G. Mensah, S. Mendis & K. Greenlund (Eds.), *The WHO Atlas of Heart Disease and Stroke* (p. 5). Geneva: World Health Organization.
- Maher, L. M., Chatterjee, A., Gonzalez Rothi, L. J., & Heilman, K. M. (1995). Agrammatic sentence production: The use of temporal-spatial strategy. *Brain and Language*, 49(2), 105-124. Retrieved from <http://www.sciencedirect.com/science/article>

- Martin, N. (1996). Phonological facilitation of semantic errors in normal and aphasic speakers. *Language and Cognitive Processes*, 11(3), 257-282. doi: 10.1080/016909696387187
- Martin, R. C., Crowther, J. E., Knight, M., Tamborello II, F. P., & Yang, C. L. (2010). Planning in sentence production: Evidence for the phrase as a default planning scope. *Cognition*, 116(2), 177-192. doi: 10.1016/j.cognition.2010.04.010
- Martin, R. C., & Freedman, M. L. (2001). Short-term retention of lexical-semantic representations: Implications for speech production. *Memory*, 9(4-6), 261-280. doi: 10.1080/09658210143000173
- Martin, R. C., & He, T. (2004). Semantic short-term memory and its role in sentence processing: A replication. *Brain and Language*, 89, 76-82. doi: 10.1016/S0093-934X(03)00300-6
- Martin, R. C., Lesch, M. F., & Bartha, M. C. (1999). Independence of input and output phonology in word processing and short-term memory. *Journal of Memory and Language*, 41, 3-29. doi: 10.1006/jmla.1999.2637
- Martin, R. C., Miller, M., & Vu, H. (2004). Lexical-semantic retention and speech production: Further evidence from normal and brain-damaged participants for a phrasal scope of planning. *Cognitive Neuropsychology*, 21(6), 625-644. doi: 10.1080/02643290342000302
- Martin, R. C., & Romani, C. (1994). Verbal working memory and sentence comprehension: A multiple-components view. *Neuropsychology*, 8(4), 506-523. doi: 10.1037/0894-4105.8.4.506
- Martin, R. C., Shelton, J. R., & Yaffee, L. S. (1994). Language processing and working memory: Neuropsychological evidence for separate phonological and semantic

- capacities. *Journal of Memory and Language*, 33, 83-111. doi: 10.1006/jmla.1994.1005
- McCarthy, R. A., & Kartsounis, L. D. (2000). Wobbly words: Refractory anomia with preserved semantics. *Neurocase*, 6(6), 487-497. doi: 10.1080/13554790008402719
- Medina, J., Kimberg, D. Y., Chatterjee, A., & Coslett, H. B. (2010). Inappropriate usage of the Brunner-Munzel test in recent voxel-based lesion-symptom mapping studies. *Neuropsychologia*, 48(1), 341-343. doi: 10.1016/j.neuropsychologia.2009.09.016
- Meyer, A. S. (1996). Lexical access in phrase and sentence production: Results from picture-word interference experiments. *Journal of Memory and Language*, 35, 477-496.
- Meyer, A. S., Sleiderink, A. M., & Levelt, W. J. M. (1998). Viewing and naming objects: Eye movements during noun production. *Cognition*, 66(2), B25-B33. doi: 10.1016/S0010-0277(98)00009-2
- Meyer, A. S., & van der Meulen, F. F. (2000). Phonological priming effects on speech onset latencies and viewing times in object naming. *Psychonomic Bulletin & Review*, 7(2), 314-319. doi: 10.3758/BF03212987
- Miceli, G., Silveri, M. C., Romani, C., & Caramazza, A. (1989). Variation in the pattern of omissions and substitutions of grammatical morphemes in the spontaneous speech of so-called agrammatic patients. *Brain and Language*, 36(3), 447-492. doi: 10.1016/0093-934X(89)90079-5
- Miceli, G., Silveri, M. C., Villa, G., & Caramazza, A. (1984). On the basis for the agrammatic's difficulty in producing main verbs. *Cortex*, 20(2), 207-220. doi: Retrieved from <http://www.journals.elsevier.com/cortex>
- Michael, G. A. (2007). A significance test of interaction in 2xK designs with proportions. *Tutorials in Quantitative Methods for Psychology*, 3(1), 1-7. Retrieved from <http://recherche.univ-lyon2.fr/emc/IMG/pdf/Michael2007.pdf>



- Morrison, C. M., Chappell, 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(3), 528-559. doi: 10.1080/027249897392017
- Nespoulous, J. L. (2000). Invariance vs variability in aphasic performance. An example: Agrammatism. *Brain and language*, 71(1), 167-171. doi: 10.1006/brln.1999.2242
- Nespoulous, J. L., Dordain, M., Perron, C., Ska, B., Bub, D., Caplan, D., . . . Lecours, A. R. (1988). Agrammatism in sentence production without comprehension deficits: Reduced availability of syntactic structures and/or of grammatical morphemes? A case study. *Brain and Language*, 33(2), 273-295.
- Nickels, L. (2002). Theoretical and methodological issues in the cognitive neuropsychology of spoken word production. *Aphasiology*, 16(1-2), 3-19. doi: 10.1080/02687040143000645
- Novick, J. M., Kan, I. P., Trueswell, J. C., & Thompson-Schill, S. L. (2009). A case for conflict across multiple domains: Memory and language impairments following damage to ventrolateral prefrontal cortex. *Cognitive Neuropsychology*, 26(6), 527-567. doi: 10.1080/02643290903519367
- Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2010). Broca's area and language processing: Evidence for the cognitive control connection. *Language and Linguistics Compass*, 4(10), 906-924. doi: 10.1111/j.1749-818X.2010.00244.x
- Oldfield, R., & Wingfield, A. (1964). The time it takes to name an object. *Nature*, 202, 1031-1032. doi: 10.1038/2021031a0
- Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition*, 114(2), 227. doi: 10.1016/j.cognition.2009.09.007

- Peduzzi, P., Concato, J., Kemper, E., Holford, T. R., & Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology*, 49(12), 1373-1379. doi: 10.1016/S0895-4356(96)00236-3
- Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, 39(4), 633-651. doi: 10.1006/jmla.1998.2592
- Power, M. J. (1986). A technique for measuring processing load during speech production. *Journal of Psycholinguistic Research*, 15(5), 371-382. doi: 10.1007/BF01067720
- Rahman, R. A., & Melinger, A. (2009). Semantic context effects in language production: A swinging lexical network proposal and review. *Language and Cognitive Processes*, 24(5), 713-734. doi: 10.1080/01690960802597250
- Rahman, R. A., & Melinger, A. (2011). The dynamic microstructure of speech production: Semantic interference built on the fly. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(1), 149. doi: 10.1037/a0021208
- Raymer, A., & Kohen, F. (2006). Word-retrieval treatment in aphasia: Effects of sentence context. *Journal of Rehabilitation Research and Development*, 43(3), 367-378. doi: 10.1682/JRRD.2005.01.0028
- Rochon, E., Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (2000). Quantitative analysis of aphasic sentence production: Further development and new data. *Brain and Language*, 72, 193-218. doi: 10.1006/brln.1999.2285
- Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, 42, 107-142. doi: 10.1162/jocn.2010.21474
- Rorden, C., Bonilha, L., Fridriksson, J., Bender, B., & Karnath, H.-O. (2012). Age-specific CT and MRI templates for spatial normalization. *NeuroImage*, 61(4), 957-965. doi: 10.1016/j.neuroimage.2012.03.020

- Rorden, C., & Karnath, H.-O. (2004). Using human brain lesions to infer function: A relic from a past era in the fMRI age? *Nature Reviews Neuroscience*, 5(10), 812-819. doi: 10.1038/nrn1521
- Rorden, C., Karnath, H.-O., & Bonilha, L. (2007). Improving lesion-symptom mapping. *Journal of Cognitive Neuroscience*, 19(7), 1081-1088. doi: 10.1162/jocn.2007.19.7.1081
- Rudrauf, D., Mehta, S., Bruss, J., Tranel, D., Damasio, H., & Grabowski, T. J. (2008). Thresholding lesion overlap difference maps: Application to category-related naming and recognition deficits. *NeuroImage*, 41(3), 970-984. doi: 10.1016/j.neuroimage.2007.12.033
- Ruiter, M. B., Kolk, H. H. J., & Rietveld, T. C. M. (2010). Speaking in ellipses: The effect of a compensatory style of speech on functional communication in chronic agrammatism. *Neuropsychological rehabilitation*, 20(3), 423-458. doi: 10.1080/09602010903399287
- 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. doi: 10.1016/0093-934X(89)90030-8
- Saffran, E. M., & Martin, N. (1997). Effects of structural priming on sentence production in aphasics. *Language and Cognitive Processes*, 12(5-6), 877-882. doi: 10.1080/016909697386772
- Saffran, E. M., Schwartz, M. F., Linebarger, M. C., Martin, N., & Bochetto, P. (1988). *The Philadelphia comprehension battery*. Unpublished test battery.
- Saffran, E. M., Schwartz, M. F., & Marin, O. S. M. (1980). The word order problem in agrammatism: II. Production. *Brain and language*, 10(2), 263-280.

- Sahraoui, H., & Nespoulous, J. L. (2012). Across-task variability in agrammatic performance. *Aphasiology*, 26(6), 785-810. doi: 10.1080/02687038.2011.650625
- Salis, C., & Edwards, S. (2004). Adaptation theory and non-fluent aphasia in English. *Aphasiology*, 18(12), 1103-1120. doi: 10.1080/02687030444000552
- Sass, K., Heim, S., Sachs, O., Thedde, K., Muehlhaus, K., Krach, S., & Kircher, T. (2010). Why the leash constrains the dog: The impact of semantic associations on sentence production. *Acta Neurobiol. Exp. (Wars.)*, 70, 435-453. Retrieved from <http://www.ane.pl/home.php>
- Scheperjans, F., Eickhoff, S. B., Hömke, L., Mohlberg, H., Hermann, K., Amunts, K., & Zilles, K. (2008). Probabilistic maps, morphometry, and variability of cytoarchitectonic areas in the human superior parietal cortex. *Cerebral Cortex*, 18(9), 2141-2157. doi: 10.1093/cercor/bhm241
- Scheperjans, F., Hermann, K., Eickhoff, S. B., Amunts, K., Schleicher, A., & Zilles, K. (2008). Observer-independent cytoarchitectonic mapping of the human superior parietal cortex. *Cerebral Cortex*, 18(4), 846-867. doi: 10.1093/cercor/bhm116
- Schnur, T. T., Lee, E., Coslett, H. B., Schwartz, M. F., & Thompson-Schill, S. L. (2005). When lexical selection gets tough, the LIFG gets going: A lesion study of interference during word production. *Brain and Language*, 95(1), 12-13. doi: 10.1016/j.bandl.2005.07.008
- 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(2), 199-227. doi: 10.1016/j.jml.2005.10.002
- Schnur, T. T., Schwartz, M. F., Kimberg, D. Y., Hirshorn, E., Coslett, H. B., & Thompson-Schill, S. L. (2009). Localizing interference during naming: Convergent neuroimaging

- and neuropsychological evidence for the function of Broca's area. *Proceedings of the National Academy of Sciences*, 106(1), 322-327. doi: 10.1073/pnas.0805874106
- Schwartz, M. F., Faseyitan, O., Kim, J., & Coslett, H. B. (2012). The dorsal stream contribution to phonological retrieval in object naming. *Brain*, 135(12), 3799-3814. doi: 10.1093/brain/aws300
- Schwartz, M. F., & Hodgson, C. (2002). A new multiword naming deficit: Evidence and interpretation. *Cognitive Neuropsychology*, 19(3), 263-288. doi: 10.1080/02643290143000187
- Schwartz, M. F., Saffran, E. M., & Marin, O. S. M. (1980). The word order problem in agrammatism: I. Comprehension. *Brain and language*, 10(2), 249-262.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2), 461-464. Retrieved from <http://projecteuclid.org/DPubS?service=UI&version=1.0&verb=Display&handle=euclid.aos/1176344136>
- Scott, R. M., & Wilshire, C. E. (2010). Lexical competition for production in a case of nonfluent aphasia: Converging evidence from four different tasks. *Cognitive Neuropsychology*, 27(6), 505-538. doi: 10.1080/02643294.2011.598853
- Shapiro, L. P., & Levine, B. A. (1990). Verb processing during sentence comprehension in aphasia. *Brain and Language*, 38(1), 21-47. doi: 10.1016/0093-934X(90)90100-U
- Shapiro, L. P., Zurif, E., & Grimshaw, J. (1987). Sentence processing and the mental representation of verbs. *Cognition*, 27(3), 219-246. doi: 10.1016/S0010-0277(87)80010-0
- Slevc, L. R. (2011). Saying what's on your mind: Working memory effects on sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(6), 1530. doi: 10.1037/a0024350

- Smith, M. C., & Wheeldon, L. R. (1999). High level processing scope in spoken sentence production. *Cognition*, 73, 205-246. doi: 10.1016/S0010-0277(99)00053-0
- Smith, M. C., & Wheeldon, L. R. (2004). Horizontal information flow in spoken sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(3), 675. doi: 10.1037/0278-7393.30.3.675
- Stemberger, J. P. (1982). Syntactic errors in speech. *Journal of Psycholinguistic Research*, 11(4), 313-345. doi: 10.1007/BF01067585
- Stemberger, J. P. (1984). Structural errors in normal and agrammatic speech. *Cognitive Neuropsychology*, 1(4), 281-313. doi: 10.1080/02643298408252855
- Stemberger, J. P. (1985). An interactive activation model of language production. In A. Ellis (Ed.), *Progress in the psychology of language* (Vol. 1, pp. 143-186). Mahwah NJ: Erlbaum.
- Tabachnick, B. G., & Fidell, L. S. (2012). *Using Multivariate Statistics: International Edition*. Pearson.
- Thompson, C. K. (2003). Unaccusative verb production in agrammatic aphasia: A syntactic account of verb production deficits. *Journal of Neurolinguistics*, 16, 151-167. doi: 10.1016/S0911-6044(02)00014-3
- Thompson, C. K., & Choy, J. J. (2009). Pronominal resolution and gap filling in agrammatic aphasia: Evidence from eye movements. *Journal of Psycholinguistic Research*, 38(3), 255-283. doi: 10.1007/s10936-009-9105-7
- Thompson, C. K., Lange, K. L., Schneider, S. L., & Shapiro, L. P. (1997). Agrammatic and non-brain-damaged subjects' verb and verb argument structure production. *Aphasiology*, 11, 473-490. doi: 10.1080/02687039708248485
- 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(26), 15855-15860. doi: 10.1073/pnas.95.26.15855
- Thothathiri, M., Gagliardi, M., & Schwartz, M. (2012). Subdivision of frontal cortex mechanisms for language production in nonfluent aphasia. *Neuropsychologia*, 50, 3284-3294. doi: 10.1016/j.neuropsychologia.2012.09.021
- Thothathiri, M., Schwartz, M. F., & Thompson-Schill, S. L. (2010). Selection for position: The role of left ventrolateral prefrontal cortex in sequencing language. *Brain and Language*, 113, 28-38. doi: 10.1016/j.bandl.2010.01.002
- Trueswell, J. C., & Kim, A. E. (1998). How to prune a garden path by nipping it in the bud: Fast priming of verb argument structure. *Journal of Memory and Language*, 39(1), 102-123. doi: 10.1006/jmla.1998.2565
- Wagner, V., Jescheniak, J. D., & Schiefers, H. (2010). On the flexibility of grammatical advance planning during sentence production: Effects of cognitive load on multiple lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(2), 423. doi: 10.1037/a0018619
- Wheeldon, L. R., Ohlson, N., Ashby, A., & Gator, S. (2013). Lexical availability and grammatical encoding scope during spoken sentence production. *The Quarterly Journal of Experimental Psychology* (ahead-of-print), 1-21. doi: 10.1080/17470218.2012.754913
- Wheeldon, L. R., Smith, M. C., & Apperly, I. A. (2011). Repeating words in sentences: Effects of sentence structure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(5), 1051. doi: 10.1037/a0023435
- Williams, S. E., & Canter, G. J. (1982). The influence of situational context on naming performance in aphasic syndromes. *Brain and Language*, 17, 92-106. doi: 10.1016/0093-934X(82)90007-4

- Wilshire, C. E., & McCarthy, R. A. (2002). Evidence for a context-sensitive word retrieval disorder in a case of nonfluent aphasia. *Cognitive Neuropsychology*, 19(2), 165-186. doi: 10.1080/02643290143000169
- 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*, 5(4), 473-516. doi: 10.1080/02643298808253270
- Zingeser, L. B., & Berndt, R. S. (1990). Retrieval of nouns and verbs in agrammatism and anomia. *Brain and Language*, 39(1), 14-32. doi: 10.1016/0093-934X(90)90002-X
- 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*, 5(4), 473-516. doi: 10.1080/02643298808253270
- Zurif, E. B. (1982). The use of data from aphasia in constructing a performance model of language. In M. A. Arbib, D. Caplan & J. C. Marshall (Eds.), *Neural Models of Language Processes*. New York: Academic Press.