Frequency effects in the processing of verbs and argument structure: Evidence from adults with and without acquired aphasia

By

Elizabeth Courtney Anderson

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Sheffield
Faculty of Medicine, Dentistry & Health
Department of Human Communication Sciences

December 2017

For nouns elicit ideas that leave profound and lasting traces;…whilst verbs, whose function is to denote movements, denoting what happens before and what afterwards, referring to the invisible present, which is very difficult to be understood even by philosophers.

G.B. Vico, Naples, 1744

From Denes and Dalla Barba (1998)

Dedicated to the loving memory of John David Gersdorf

17 January 1938 - 1 December 2017

Acknowledgements

I owe my deepest gratitude to my supervisors, Dr Ruth Herbert and Professor Patricia Cowell, for their continued enthusiasm and guidance throughout this project. I consider myself very fortunate to have in my supervisors such wonderful mentors, educators and women in science.

To the committee who granted me doctoral funding through a Faculty Fee Scholarship and teaching assistantship - for making possible the enriching experiences of the last four years.

To Dr Stuart Cunningham and Daniel Wilson - many thanks for the many hours you spent assisting me with data capture and processing.

To Dr Dagmar Divjak and Dr Traci Walker - for generously volunteering your time and input to contribute to my development as a linguist.

To Sarah Allen, Amy Barrett, Andrew Buddery, Hannah Gill, Beth Kelly and Lydia Seevathean - for providing materials and reliability judgements that improved the rigour of this project.

To Dr Caroline Haw and members of the Aphasia Research Group - for inspiring the argument about treatment generalisation developed in Section 9.3.2.

To all my participants - for helping me learn as much about life as I learned about research during this project. Thank you for your examples of perseverance and joy.

To Mum, Dad, Kathy and Colin - for all your support, in kind emotional, financial and edible.

And finally to Luke - for your endless patience and uncompromising love. You are what is truly most important.

Tansley, Matlock, December 2017

Abstract

Introduction

In usage-based approaches to language, grammar is viewed as an emergent phenomenon that derives from humans' repeated exposure to individual instances of particular linguistic expressions (Bybee, 2006). Goldberg's (1995) construction grammar is a version of usage-based grammar that treats language as an inventory of form-meaning pairings, termed constructions. Usage-based approaches to language predict that factors of language use, such as frequency of occurrence, affect processing at every level of the linguistic system, from sounds to sentences. This approach is gaining increasing recognition in the field of aphasiology, where sentence-level frequency effects have historically been described in terms of deficits (Gahl & Menn, 2016). The current research adopts a usage-based approach to language and contributes new data on the topic of verb and sentence processing in typical adults and adults with acquired aphasia.

Aims

This research investigated the effects of two frequency-based properties of verbs on language processing in adults, including the frequency of a verb as a single word, termed lexical frequency, and the frequency of a verb in a particular syntactic construction, termed construction frequency. Specifically, this project aimed: (1) to examine the effect of construction frequency and lexical frequency on sentence processing in adults; (2) to explore whether the pattern of performance from adults with acquired aphasia was similar to or divergent from the performance of typical adults; and (3) to consider how residual linguistic capabilities in participants with aphasia affected their performance in experimental tasks.

Methods

In Phase 1, 20 typical adults and four adults with acquired aphasia took part in a verbal fluency task in which they named verbs that could occur in eight unique syntactic constructions. Noun phrases were encoded as pronouns, so no semantic activation was available from the lexemes contained in sentence stimuli, and a blank space stood in place of the verb. For example, a sentence corresponding to the conative construction was presented as *you* ___ *at us*.

In Phase 2, 90 typical adults and 14 adults with acquired aphasia took part in a grammaticality judgement task and a sentence completion task. Participants silently read sentences like those in Phase 1 and were subsequently presented with a written verb. In the grammaticality judgement task, participants decided whether or not the verb could occur in

the sentence stimulus. In the sentence completion task, participants replaced the blank space in the sentence stimulus with the given verb and produced the entire sentence aloud. Participants' number of target responses and response times were measured in each task.

The frequency of verbs in Phase 2 varied along two dimensions. These independent variables included construction frequency and lexical frequency, each of which had two levels, namely, high frequency and low frequency. These four groups resulted in a factorial design, where conditions differed with respect to levels of construction frequency and lexical frequency.

Results

In Phase 1, the number of times typical participants generated verbs in response to syntactic constructions was more strongly related to verbs' construction frequency than lexical frequency, for most constructions. Sentence stimuli successfully elicited verbs from participants with aphasia.

In Phase 2, typical participants showed an effect of construction frequency in the grammaticality judgement task and an effect of lexical frequency in the sentence completion task. These effects were moderated by construction and interactions. In general, group-level results from participants with aphasia were consistent with findings from typical participants. Some individuals with aphasia showed frequency effects to a greater or lesser extent than typical participants.

Conclusion

Results suggest that at the sentence level, the frequency of verbs as single words and the frequency of verbs in particular syntactic contexts affects language processing, depending on task demands. Findings confirm the predicted effect of linguistic experience on language use. Importantly, this project extends the number of investigations of pathological language undertaken in a usage-based linguistic framework. Results from participants with aphasia are discussed with reference to treatments for sentence processing deficits in aphasia, item selection for those treatments and theories of agrammatism.

Publications arising

Student conference presentations

Anderson, E., Herbert, R., & Cowell, P. (2014). Verbs, argument structure and grammatical constructions: Relationships in neurologically intact and impaired speakers, Linguistics and English Language Postgraduate Conference, University of Edinburgh, May 2014.

Anderson, E., Herbert, R., & Cowell, P. (2015). Verbs and argument structure in neurologically intact speakers and speakers with aphasia, Department of Human Communication Sciences PGR Conference, University of Sheffield, 7 July 2015.

Anderson, E., Herbert, R., & Cowell, P. (2016). Frequency effects in the processing of verbs and argument structure constructions: A pilot study with participants with acquired aphasia, Department of Human Communication Sciences PGR Conference, University of Sheffield, 28 June 2016.

International conference presentations

Anderson, E., Herbert, R., & Cowell, P. (2015). Argument structure constructions elicit verbs from speakers with aphasia, British Aphasiology Society Biennial International Conference, University College London, 9-11 September 2015.

Anderson, E., Herbert, R., & Cowell, P. (2016). The role of frequency in the association between verbs and argument structure constructions, 6th UK Cognitive Linguistics Conference, University of Bangor, 19-22 July 2016.

Anderson, E., Herbert, R., & Cowell, P. (2016). Frequency affects processing of verbs and argument structure in acquired aphasia: Results from a pilot investigation, 17th Science of Aphasia Conference, Venice, Italy, 25-30 September 2016.

Anderson, E., Herbert, R., & Cowell, P. (2016). Frequency effects in sentence production in acquired aphasia: A preliminary investigation, Academy of Aphasia 54th Annual Meeting, Llandudno, Wales, 16-18 October 2016.

Peer-reviewed articles

Anderson, E., Herbert, R., & Cowell, P. (forthcoming). The role of frequency in the association between verbs and argument structure constructions. Online proceedings of UK-CLA Meetings, available from http://uk-cla.org/proceedings

Invited talks

Anderson, E., Herbert, R., & Cowell, P. (2017). Single-word and context frequency effects of verbs in aphasia, British Aphasiology Society Research Update Meeting, Leeds Beckett University, 27 April 2017.

Anderson, E., Herbert, R., & Cowell, P. (2017). Construction Grammar in aphasia research, Aphasia Research Group, University College London, 14 June 2017.

Table of contents

List of tables

List of figures

1 Linguistic structure in usage-based approaches to language

This research adopts a usage-based approach to language, as articulated by Bybee (2010) and Goldberg (2006). This chapter describes how usage-based approaches to language emerged in the late twentieth century and how they contrast with previous scholarship in the field of linguistics. Usage-based approaches to language recognise frequency of occurrence as a driving force behind language structure. These approaches view language as an emergent phenomenon based on humans' application of domain-general processing capacities to specific instances of language use. The processing mechanisms that shape language structure are described in the first section.

One approach to grammar that is compatible with usage-based approaches to language is based on constructions, which are pairings of specific linguistic forms with particular communicative functions. Constructions are linguistic units that language users recognise as independent due to their high frequency of occurrence, and their functions can be semantic or pragmatic (Goldberg, 2006). The association between form and meaning in language was recognised as long ago as the fifth century BC. In Plato's *Cratylus,* Socrates observed that the meaning of words cannot be predicted from their phonetic shape (Sedley, 2013). In the *Categories*, Aristotle proposed ten categories to classify *ta legomena*, 'things that are said'. One interpretation of this work holds that these categories are based on linguistic structures (Studtmann, 2013). Though the study of language based on form-meaning correspondences fell into disfavour with many linguists in the mid-twentieth century (see below), it can be said to be a view of language that has classical roots.

This chapter introduces a version of construction grammar (Goldberg, 1995) that specifies how constructions combine to create sentences, and the chapter concludes by reviewing evidence that language users can access the semantics of constructions greater than the single word.

1.1 Foundations of usage-based approaches to language

This section provides an account of three dimensions of usage-based approaches to language: how the field developed over the past fifty years, how language may derive from domain general cognitive functions and how language structure can emerge from repeated exposure to linguistic forms.

1.1.1 Contrast between usage-based approaches and generative grammar

The crux of linguistic theory in usage-based approaches to language contrasts in several fundamental ways with earlier thinking in the field. With the publication of *Syntactic Structures* in 1957, Chomsky and his views of transformational generative grammar have influenced linguistic endeavour from the mid twentieth century onwards. Several important differences exist between the research programme advanced in Chomsky's generative grammar and the usage-based approach.

Generative grammar focuses primarily on describing knowledge of language in an abstract conception of the mind (Chomsky, 1980). The theory is based on the distinction between language use and language knowledge. Externalised language (E-language) refers to the actual use of language and its manifestation as a social phenomenon. In contrast, Internalised language (I-language) refers to the system of language that is represented in the mind, irrespective of any particular aspects of the system's use. The grammar of an E-linguist is an ordered description of the facts of language use, while the grammar of an I-linguist is argued to reflect properties of the human mind (Cook & Newson, 1996). The tension between the investigation of E-language and I-language is historic. As Cook and Newson (1996) commented:

> The opposition between these two approaches in linguistics has been long and acrimonious; neither side concedes the other's reality…An E-linguist collects samples of actual speech or actual behaviour; evidence is concrete physical manifestation. An I-linguist invents possible and impossible sentences; evidence is whether speakers know if they are grammatical. The E-linguist despises the I-linguist for not looking at 'real' facts; the I-linguist derides the E-linguist for looking at trivia (p. 22).

Usage-based approaches to language fit comfortably within Cook and Newson's description of E-language, as these approaches focus on describing instances of language use. As reviewed in the next section, many researchers in the usage-based approach take language processing to be a function of domain-general cognitive processes in humans.

Related to the distinction between E-language and I-language in generative grammar is the difference between competence and performance. E-language corresponds to a speaker's performance, or her actual use of language in a given context. The representation of I-language in the mind corresponds to a speaker's language competence, or her linguistic knowledge. Competence can be further divided into the core and periphery. The core is composed of linguistic rules that can be applied to language and result in linguistic forms, and the periphery contains knowledge of exceptions to these main patterns of language. The periphery contains, for example, idioms, marked forms and non-standard uses of structural forms (Cook & Newson, 1996). Usage-based approaches make fewer distinctions between rule-based language description and exceptions. Rather, the cognitive mechanisms that underpin the processing of linguistic structures that can be considered as the periphery apply equally well to the processing of core structures.

This distinction between competence and performance is carried over into the generative approach to syntax. The theory of autonomous syntax holds that the form of sentences is independent of their meaning and use. Syntactic operations occur in isolation from semantic influences and pragmatic functions (Nuyts, 1995). Syntax in this view is not influenced by usage events, which are relegated to E-language.

Beginning in the 1970s, functional linguists in the fields of language change and typology began to recognise that grammar is crucially related to context and discourse. Many researchers came to view language structure, or grammar, as conventions based on repeated instances of language use (Bybee, 2007). Langacker (1988) first employed the term ‘usage-based’ to describe grammar as a network of constructions of varying abstractions, driven by bottom-up processes of human cognition. Today, most usage-based linguists accept that language qualifies as a complex, adaptive system (Beckner et al., 2009), where language arises from interactions among individuals in a community based on past behaviour and competing motivations. Beckner et al. (2009) described this system:

> Cognition, consciousness, experience, embodiment, brain, self, human interaction, society, culture, and history are all inextricably intertwined in rich, complex, and dynamic ways in language. Everything is connected. Yet despite this complexity, despite its lack of overt government, instead of anarchy and chaos, there are patterns everywhere. Linguistic patterns are not preordained by God, genes, school curriculum, or other human policy. Instead, they are emergent (p. 18).

A multitude of current linguistic theories can be described as emergentist (MacWhinney, 2015). These approaches share three main assumptions, both with each other and with scientific endeavours in the areas of physical, biological and social research. First, the approaches are based on Darwinian evolution in terms of competition and selection. Second, they recognise complex systems as hierarchically structured, where complexity at higher levels is not entirely predictable from the structure of lower levels. Third, they appreciate that processes operate on various timescales that compete in the present moment (MacWhinney, 2015).

Despite the major differences between generative grammar and usage-based approaches, both research programmes consider language to be a component of cognition and seek to account for the potentially infinite nature of linguistic utterances, that is, both approaches seek to describe - in the technical sense - a generative grammar (Goldberg, 2006). The following section provides an overview of the relationship between language and cognition that many researchers accept in the usage-based approach.

1.1.2 Domain-general cognitive processing in language

Most usage-based linguists hold that linguistic structure emerges from the application of domain-general cognitive processes to language. Bybee (2006, 2010) identified processes that can account for language structure, including chunking, analogy and categorisation. She explains how the operation of these processes in the domain of language results in the structure, or grammar, of a language. In this view, grammar is taken to be an emergent property derived from the multiple and various ways in which speakers use language.

The processes of chunking, analogy and categorisation are attributes of human cognition that can be applied in the domain of language, giving rise to basic linguistic units known as constructions.

1.1.2.1 Chunking

Chunking is a process that results in effective sequence learning (Solopchuk, Alamia, Olivier & Zénon, 2016) across various domains in human cognition. Chunking has been cited as the mechanism involved in the performance of non-linguistic activities such as the memorisation of musical sequences (Janata & Grafton, 2003), the perception of dance (Bronner & Shippen, 2015), the execution of action sequences (Graybiel, 1998), the recall of phone numbers (Fonollosa, Neftci & Rabinovich, 2015) and the development of typing (Yamaguchi & Logan, 2016). In all these instances, humans build up combinations of elements based on their increasing experience or familiarity with information units that occur in a particular domain.

Chunking is a frequency-based process that operates over linguistic and non-linguistic information alike. In language, chunks are formed over highly frequent combinations of linguistic units, such as adjacent phonemes, morphemes and words. Chunking results in the hierarchical organisation of language, because a chunk can both contain smaller units and simultaneously form part of a larger unit. Chunks are associated with ease in their production and accessibility compared to less frequently attested combinations of linguistic items (Bybee, 2010).

Bybee and Schiebman (1999) demonstrated how the production of the word *don't* can be indicative of chunking in language. They measured the amount of phonetic reduction in the articulation of 138 instances of the word *don't* in various contexts. The authors found that articulations of *don't* were more reduced in the linguistic contexts in which they were most frequently used. For example, the full vowel /əʊ/ was reduced to a schwa after the pronoun *I*, which was the most frequent pronoun in the dataset, and before frequently occurring verbs like *know*, *think* and *have*. The authors noted the relationship between reduction of phrases and their pragmatic use, as the expression *I don't know* can function to yield the floor of a conversation to another speaker. The frequency and function of the word *don't* can account for its phonetic reduction, which indicates its status as part of a linguistic chunk.

A similar explanation applies to the analysis of the utterance *let's go out*. The utterance is hierarchical because it contains the chunks *let's* and *go out*, each of which can occur in other contexts. *Let's* can occur with a variety of elements in order to make a suggestion, such as *let's dance*, *let's eat* or *let's get going*. The phrase *go out* can refer pragmatically to travelling to a place of leisure. The positive connotation of *go out* arises from the context in which it is used, independently of the individual words *go* and *out*. In turn, the word *go* represents chunking of the phonemes /g/ and /əʊ/. Chunking over phonemes and words results in the structure of the utterance *let's go out*.

1.1.2.2 Analogy

Gentner (1983) described analogy as structure-mapping between domains in which 'a relational structure that normally applies in one domain can be applied in another domain' (p. 156). Analogy operates on the similarity of the relationship between objects in different scenarios, irrespective of their attributes (Gentner, 1983). Analogy is an integral component of human intelligence and problem solving. Humans can use analogy to produce solutions to novel problems immediately after experiencing an analogous situation (Gick & Holyoak, 1980) and even across intervening years (Chen, Mo & Honomichl, 2004). Analogy also results in better learning outcomes for children (Tse, Fong, Wong & Masters, 2017), younger adults (Komar, Chow, Chollet & Seifert, 2014; Lam, Maxwell & Masters, 2009; Tse, Wong, Whitehall, Ma & Masters, 2016) and older adults (Tse, Wong & Masters, 2017).

Linguistic analogy refers to the use of a novel item in an existing construction and contributes to explaining language productivity (Bybee, 2010). The phrase *let's go out* can be recognised as an instance of a more general linguistic pattern containing the word *let's* followed by a verb phrase, or what could be called the *let's X* construction, the English hortative construction (van der Auwera, Dobrushina & Goussev, 2013). The form *let's X* encodes a speaker's suggestion and emphasises the positive interpretation of an action. A listener unfamiliar with dance crazes of the twentieth century would nonetheless be able to interpret the word *jive* as such a positive action in the utterance *let's jive*. The means by which the listeners uncovered such a meaning is the process of analogy to other elements in the *let's X* construction, such as *go out.*

1.1.2.3 Categorisation

Categorisation is a well-attested psychological process. Categorisation has been cited as the process by which humans classify objects, predict features of newly encountered items, maintain social stereotypes and even form consumer preferences (Markman & Ross, 2003), as well as identify environmental sounds (Guastavino, 2007) and negotiate space (Hund & Plumert, 2005).

Categorisation can also constrain language. Some linguistic patterns contain schematic positions in which a variety of lexical items can occur, such as the second element in the *let's X* construction discussed above. In Bybee's (2010) view, 'constructions contain schematic positions that encompass sets of items that have been sorted into categories' (p. 57). Categories are built up from a language user's experience with particular utterances that occur in the same linguistic pattern. Categories have central members, known as prototypes, and other items show graded category membership with respect to their frequency and similarity to a prototype (Rosch & Mervis, 1975).

Frequency is integral to the identification of category prototypes. Taylor (2015) examined the *for NP on end* construction, where the noun phrase refers to a unit of time and the construction encodes a speaker's feeling of boredom at the passing of time. Taylor specified frequency as the defining factor of prototypicality for the noun phrase category that occurs in the construction. The terms *hours*, *days*, *weeks* and *months* were the most frequent noun phrases in the construction in data from the British National Corpus. Other, less frequent noun phrases, such as *many days* and *several weeks*, were related to these central members as more specific instances of the duration of time, but occurred only once.

The second element in the construction *let's X* refers to an action with positive connotations, as noted above. Items that can occur in this position can be viewed as a category. The most frequent item in the construction, according to the British National Corpus, is the verb *go*, so *go* can be considered a prototypical member of this category. The items *go out* and *jive* are less frequent members of the category but are semantically similar to the prototype *go* because they describe more specific forms of actions. *Go out* may be considered more central to the category than *jive*, because it is a more frequent phrase and more similar to the central member *go*, in terms of semantics and phonology.

1.1.3 Frequency and language structure

As reviewed in the previous sections, most usage-based approaches to language accept that frequency of occurrence is a driving force behind the shape of grammar. Frequency in language use is an explanatory factor in processes as diverse as phonological reduction, constituency and diachronic language change (Bybee & Hopper, 2001). This section reviews the various ways that frequency of occurrence has been quantified and how frequency has been observed to affect language structure, at the level of individual constructions and the grammatical system.

Usage-based linguists have operationalised frequency of occurrence in three main ways. First, token frequency refers to how often a linguistic unit occurs in language overall (Bybee, 2007). For example, the word *go* occurs in British English 881 times per million words (Leech, Rayson & Wilson, 2001). Second, type frequency refers to the number of distinct items that can occur in a linguistic pattern (Bybee, 2007). The British National Corpus lists a total of 398 verbs that can occur in the construction *let's X*, meaning the construction has a lower type frequency than the construction *don't X*, in which 1509 verbs occur. Third, contextualised frequency refers to the frequency of the association between linguistic items (Divjak & Caldwell-Harris, 2015). This frequency measure originated in the field of corpus linguistics and lexicography, where context can aid in differentiating among various word senses. The phrase *let's go* occurs 901 times in the British National Corpus, but *let's dance* occurs only ten times, so contextualised frequency counts indicate there is a stronger association between *let's* and *go* than *let's* and *dance*.

Bybee and Thompson (1997) explained how token frequency and type frequency underpin several effects that influence the development of individual constructions over time. The reducing effect and conserving effect refer to syntactic changes that affect items with high token frequency. The reducing effect is evident in how the phrase *be supposed to* was originally a passive form that underwent grammaticalisation to become today's auxiliary, *s'posta*, with a meaning similar to *should*. The conserving effect is evident in how all Middle English verbs could invert with the subject of a sentence and could precede the negative marker *not*, while today only the high frequency auxiliary verbs retain these properties. Lastly, type frequency relates to the degree of productivity that syntactic constructions display. The English ditransitive construction demonstrates limited productivity that extends to certain verb classes more readily than others, given their type frequency. For example, verbs of transfer occur more productively in the ditransitive construction than verbs referring to manners of speaking, as *I emailed him the document* seems grammatically acceptable, but *I yelled him the slogan* may seem less so.

Like Langacker (1988), Diessel (2015) described the structure of language as a network architecture. Network models in cognitive science are designed to process data and learn from that processing, and so by definition they are usage-based. In Diessel's description, constructions are related to one another by four types of associative links. Taxonomic links describe the hierarchical nature of grammar, in that an utterance can instantiate a more abstract construction. In this section, the phrase *let's go out* has been described as a specific instance of the more general, or schematic, *let's X* construction. Horizontal links describe the relationship between constructions at the same level of abstraction, such as related words. Syntactic links capture the syntactic phrases that can occur in schematic constructions. Finally, lexical links relate individual words to schematic constructions. To illustrate, *go* is associated with a variety of constructions, such as *let's go*, as well as sentences like *we go through the field* and *she went at the parcel with a knife*, whose structures are, respectively, the intransitive motion construction and the conative construction (see below). Lexical links are frequency-driven in that they reflect probabilistic relationships between words and schematic constructions based on an individual's language experience. This network model can be taken to reflect the organisation of the mental lexicon.

This section introduced usage-based approaches to language as one view of current thinking in linguistics. Humans can utilise domain-general cognitive mechanisms in language processing, and frequency of occurrence plays a major role in both this processing and the shaping of language structure. The following section will introduce a particular version of usage-based grammar in more detail.

1.2 Construction grammar

This section presents an approach to grammar that is consistent with a usage-based approach to language. As reviewed in the previous section, language structure can be said to arise from the application of domain-general processing mechanisms to language. These mechanisms give rise to linguistic units referred to as constructions. The phrase *let's X*, discussed above, qualifies as such a construction. A more general view of constructions is adopted by Goldberg (1995, 2003) in her version of construction grammar, outlined below.

1.2.1 What is a construction?

Goldberg (2013) defined constructions as 'learned correspondences between form and function, at various levels of complexity and abstractions' (p. 435). Any linguistic pattern qualifies as a construction if an aspect of its form or function is not predictable from its component parts, or it occurs with sufficient frequency (Goldberg, 2006). Table 1.1 presents examples of constructions in English of increasing size and complexity (Goldberg 2006).

Table 1.1 English constructions, from Goldberg (2006)

Morpheme	*un-*, *-er*
Word	*abracadabra*, *beam*, *count*
Complex word	*jellybean*
Partially-filled complex word	Adj-*ly* (regular adverbs)
Filled idiom	*Bring home the bacon*, *cut it out*, *knock it off, search me*
Partially filled idiom	*Give* <someone> *a dressing down*, *make* <someone>*'s day*
Covariational conditional	The Xer the Yer *The closer you get, the better it looks*
Transitive	Subject Verb Object *Sam sent a letter*
Ditransitive	Subject Verb Object1 Object2 *Sam sent Alex a letter*
Passive	Subject aux VPpp (PP_{by}) *The letter was sent by Sam*

Note. Examples are author's own.

Constructions vary in their abstractness and lexical specification. Words and idioms are entirely lexically specified, while complex words can be only partially lexically specified. The last three constructions in Table 1.1 are abstract in that they are patterns that do not contain any particular content words. Some authors reserve the term 'construction' to refer to these schematic patterns (Diessel, 2015), but Goldberg (1995, 2006) applies the moniker to the range of linguistic units on the continuum from morpheme to abstract, schematic construction in order to emphasise the similarity among all types of linguistic data - they are all pairings between a linguistic form and a function derived from language use.

A variety of construction grammars have been articulated over the past thirty years. They differ with regard to the phenomena that they attempt to explain and the technical processes by which their goals are realised. One of the first proposals to position grammar as a meaningful system in the realm of social interaction and cognition was Langacker's (1987) exposition of cognitive grammar. Since then, Croft's (2001) radical construction grammar has taken grammatical constructions to be the starting point for exploring syntactic typology, and sign-based construction grammar promotes formalising a theory of construction grammar (Michaelis, 2010). Embodied construction grammar focuses on the dynamic nature of the relationship between linguistic form and real-word meaning (Bergen & Chang, 2005), while fluid construction grammar is concerned with the computational modelling of language (Steels, 2012).

Proponents of construction grammars agree that a language can be described as an inventory of form-meaning pairings, or constructions. There is no strict division between syntax and the lexicon (Goldberg, 1995). Rather, the contents of the mental lexicon may be described as a 'construct-icon' containing an inventory of all the linguistic constructions that a language user has experienced (Goldberg, 2003).

1.2.2 Constructions in language use

This section provides some examples of linguistic analyses in the construction grammar framework, where researchers have specified the relationship between a particular linguistic form and its function in language use.

1.2.2.1 What's X doing Y?

Kay and Fillmore (1999) identified a construction in English that they termed the *What's X doing Y?* construction. This construction is used to express surprise at an incongruent event. Examples are shown in (1).

(1a) What are you doing here?
(1b) What is the dog doing with his muddy paws on my clean carpet?
(1c) What are these pen marks doing in this library book?

The *What's X doing Y?* construction is defined by a number of formal morphosyntactic characteristics. The construction begins with non-referential *what* and contains the present participle form of *do* as a complement of copular *be*. However, *doing* is not required to refer to an action, and its *-ing* inflection is not necessarily interpreted in the progressive aspect, as (1c) demonstrates. The inclusion of *else* is not allowed in the construction, and neither is the negation of *be* or *do*. To illustrate, the utterances *what else are you doing here*, *what aren't you doing here* and *what are you not doing here* can only be used to refer to the action taking place and cannot express a meaning of surprise.

Kay and Fillmore (1999) noted that this construction encodes a pragmatic meaning of incongruity, in that the complement denoted by X is unexpected or unwelcome in the situation described by Y. The sentence in (1a) could be used innocently to enquire about the nature of one's activities at a specific location, but when it instantiates the *What's X doing Y?* construction it is equally fitting as a response to a trespasser. The interpretation of a speaker expressing the incongruity of a situation is the main reading of (1b) and (1c). This pragmatic meaning is associated with the lexico-syntactic form of the construction. The *What's X doing Y?* construction is a unique English construction defined by morphosyntactic properties that encode a specific meaning. It is an example of a partially-filled idiom (see Table 1.1).

1.2.2.2 Nominal extraposition

Michaelis and Lambrecht (1996) provided an analysis of the nominal extraposition construction, illustrated in (2) (examples from Michaelis & Lambrecht, 1996).

(2a) It's amazing the difference!

(2b) It's unbelievable the people who are verbally abusive to fat people.

Nominal extraposition takes the form it-<*predicate*>-<*definite NP*>. A non-referential pronoun *it* functions as the subject, which is not co-referential with the extraposed noun phrase. The predicate in nominal extraposition contains the verb *be* and an adjective phrase, and the predicate must be adjacent to the extraposed noun phrase. Finally, the noun phrase in nominal extraposition is definite and marked prosodically. Spoken aloud, the adjective and noun in the nominal extraposition construction are stressed. (2a) would be realised phonologically as *it's aMAzing the DIFFerence*. This combination of formal lexical, syntactic and prosodic features defines the nominal extraposition construction.

Nominal extraposition is an exclamative speech act. As such, it can only occur as a main clause. The post-verbal NP is interpreted as scalar. (2a) is not synonymous with the observation that *the difference was amazing*. Rather, nominal extraposition expresses how amazing the difference was (Michaelis & Lambrecht, 1996). Its meaning is akin to sentences like *it's amazing how much different this restaurant has become over the last twenty years* or *it's amazing how much my new mattress has reduced my back pain*. Though noun phrases do not intrinsically denote scalar properties, they are interpreted in this way in the context of nominal extraposition.

Nominal extraposition has two defining pragmatic features. First, its use is appropriate only when the noun phrase is identifiable to the listener and recoverable from context. The noun phrase refers to an aspect of the superordinate discourse topic. Michaelis and Lambrecht (1996) referred to the noun phrase as 'inactive' but 'accessible'. This accessibility is apparent in the grammatical definiteness of the noun phrase, which reflects the referent's cognitive status as identifiable. A speaker can assume that the referent of the noun phrase is shared by the listener. Second, the entire sentence is focussed. A sentence that instantiates nominal extraposition asserts the entire proposition, including the predicate and the post-verbal noun phrase. The focus domain is therefore the entire sentence. These two pragmatic properties are unique attributes of the nominal extraposition construction.

1.2.2.3 Expedia

The examples in the previous sections were taken from published analyses of English constructions and illustrate how constructions can be specified as the pairing of linguistic forms and usage-based functions. This section will show how a similar analysis can explain a more recent example of creative language. In May 2013, the UK advertising agency Ogilvy and Mather launched a marketing campaign for the travel company Expedia that encouraged customers to *travel yourself interesting*. The campaign involved advertisements with this motto that ran on television, in print media, online and via social media throughout the summer of 2013.

Expedia's motto has two formal attributes that result from the instantiation of two abstract constructions. First, the lack of an overt subject is due to the motto's function as an imperative. The imperative is a construction that pairs a linguistic form in which the subject is left unexpressed with the pragmatic function of issuing a command (Diessel, 2015). Second, Expedia's motto contains two post-verbal arguments, an object and an adjective. The presence of the object cannot be attributed to the semantics of the verb *travel*, which refers to an event in which an agent moves to a destination. In other circumstances, *travel* cannot grammatically occur with an object, as shown in (3):

(3a) We travelled to Florida

(3b) *We travelled ourselves to Florida

Rather, the post-verbal arguments in Expedia's motto originate from the resultative construction (Goldberg, 1995), which takes the form of a post-verbal, co-referential object and predicate, and encodes an event in which the action denoted by the verb brings about the result in the object, as shown in (4).

(4a) The chef diced the vegetables very small

(4b) The ceaseless alarm drove me crazy

The formal features of Expedia's motto include the lack of syntactic subject and the presence of two phrases following the verb *travel*. These features can be attributed to the two abstract constructions, the imperative and resultative. Taken together, they explain the semantics of the utterance as an instruction by Expedia to its customers that they should travel in order to become interesting.

1.2.3 Goldberg's (1995) construction grammar

The present research is based on Goldberg's (1995) construction grammar framework, because she provides an explicit account of how sentences arise from the combination of other constructions in a usage-based approach to language.

1.2.3.1 Sentence structure in construction grammar

The final three examples in Table 1.1 are called *argument structure constructions* (Goldberg, 1995). Argument structure constructions provide the means for sentence expression, as they associate event-level meaning with syntactic structure. Argument structure constructions are composed of argument roles, which correspond to semantic roles such as agent, patient and goal. The form and meanings of other argument structure constructions are shown in Table 1.2.

Table 1.2 Forms and meanings of argument structure constructions in English, from Goldberg (1995)

Construction	Form	Meaning
Caused motion	Subject Verb Object Oblique$_{path}$ *Sam sent a letter to France*	X causes Y to move to/from Z
Conative	Subject Verb Oblique$_{at}$ *Sam waved at the postman*	X directs action at Y
Intransitive motion	Subject Verb Oblique$_{path}$ *Alex ran down the path*	X moves to/from Y
Resultative	Subject Verb Object$_i$ Predicate$_i$ *Sam kissed Alex silly*	X causes Y to become Z
Way	Subject$_i$ Verb <possessive$_i$> way Oblique$_{path}$ *The letter slowly made its way to France*	X creates path Z and moves through it

Note. Examples are author's own. Alternate rows in grey for ease of reference.

Argument structure constructions encode meaning independently of the words they contain. Consider the sentences (5), which both instantiate the resultative construction:

(5a) Sam kissed Alex silly

(5b) Jay shot Kier dead

Clearly, the events described by (5a) and (5b) are very different. While (5a) describes a scene of passion and (5b) describes a crime, the two sentences are similar in that an agent causes a change of state in the patient. In (5a), Alex became silly as a result of being kissed, and in (5b) Kier became dead as a result of being shot. This resultative meaning can be attributed directly to the syntactic form that the sentences share: *subject-verb-object-predicate*. Goldberg (1995) refers to such a semantic event as a *scene* and maintains that argument structure constructions refer to humanly relevant scenes. Her *scene encoding hypothesis* states that the meaning of an argument structure construction is an event basic to human experience. Scenes are basic in the sense that they are highly frequent ways of interacting with the world.

1.2.3.2 The integration of verbs and argument structure constructions

A sentence results from the integration of a verb and an argument structure construction. Goldberg (1995) explains the meaning of verbs with recourse to frame semantics. This approach to lexical semantics recognises that the meanings of words encompass the richness of speakers' experience with the world and their culture (Petruck, 1996). It is knowledge of this finely detailed semantic content that explains English speakers' reliable differentiation between synonyms such as *amble* and *saunter* (Dąbrowska, 2009). Verbs also specify what Goldberg (1995) terms participant roles, or the actors in the event encoded by a verb.

In order to integrate with an argument structure construction, the participant roles of a verb fuse with the argument roles of a construction. This process is guided by the roles of the verb and construction that can be considered profiled. Profiled roles of lexical verbs are participant roles that are particularly salient in the scene whose semantics they encode. Profiled roles of argument structure constructions are argument roles expressed as direct grammatical relations, such as subjects and objects (Goldberg, 1995).

The fusion of participant roles and argument roles is constrained by two principles. The semantic coherence principle ensures that only semantically compatible roles can be fused. The participant role of the verb must be an instance of the argument role of the construction, or vice versa. The correspondence principle stipulates that profiled participant roles of the verb fuse with profiled argument roles of the construction (Goldberg, 1995). These principles constrain the integration of verbs and argument structure constructions, resulting in the formation of grammatical utterances.

To illustrate the integration of a verb and an argument structure construction, consider the transitive construction, which contains two profiled argument roles, an agent and a patient. Consider also the verb *kiss*, which specifies two profiled participant roles, a kisser and the kissed. The verb *kiss* can integrate with the transitive construction because the integration satisfies the semantic coherence principle. The participant roles of *kiss* are instances of the argument roles of the transitive construction: a kisser is a type of agent, and the kissed is a type of patient. The correspondence principle is also satisfied, because both participant roles of *kiss* are profiled, and these can fuse with profiled argument roles of the transitive construction. This integration results in, for example, the sentence *Sam kissed Alex*. Importantly, argument structure constructions can contribute roles to a sentence that are not specified by the verb (Goldberg, 1995), as demonstrated in the previous analysis of Expedia's motto, where the predicate *interesting* was an attribute of the resultative construction rather than the verb *travel*.

The result of the integration of a verb and an argument structure construction is a single event that reflects the semantics of both the verb and the construction. The relationship between the meaning of the verb and the meaning of the construction is called an *R-relation* (Goldberg, 1995). Commonly, the verb designates a specific instance of the general event to which an argument structure construction refers. This type of R-relation is found between *kiss* and the transitive construction. The verb *kiss* specifies the type of action taking place in the event referred to by the construction; the meaning of *kiss* is an instance of the meaning of the construction. The relation between *travel* and the resultative construction is different. In this case, *travel* is the means by which the event referred to by the construction is accomplished: that is, the result of becoming interesting is brought about by travelling. R-relations also constrain the integration of verbs and argument structure constructions (Goldberg, 1995).

1.2.3.3 Relationships among argument structure constructions

Because construction grammar treats linguistic units of varying degrees of complexity, from morphemes and lexemes to argument structure constructions, as the same basic type of data, relationships that are traditionally thought to hold between lexical items, such as polysemy, can also hold between argument structure constructions.

Polysemy describes an extension in meaning from the central sense of a construction. Just as individual words can be polysemous, or have various different meanings, so too can argument structure constructions. Goldberg (1995) gives the example of the ditransitive construction, whose central sense is an agent's successful transfer of an object to a recipient. However, the ditransitive construction can also be used to refer to events in which the transfer is potential, enabled, refused, intentional or in the future (Goldberg, 1995). Table 1.3 exemplifies these senses of the ditransitive construction.

Table 1.3 Senses of the ditransitive construction, from Goldberg (1995)

Sense	Example
Actual transfer (central sense)	Sam sent Alex a letter
Potential transfer	Sam promised Alex a letter
Enabled transfer	The prison guard allowed Alex the letter
Refused transfer	The prison guard denied Alex the letter
Intentional transfer	Sam wrote Alex a letter
Future transfer	A neighbour forwarded Alex the letter

The ditransitive construction is an example of constructional polysemy, because the same syntactic form is associated with a collection of distinct but related meanings.

Constructions can also be related to one another as subparts, as more fully specified versions of another construction and by metaphorical extension. These four relationships among argument structure constructions are consistent with Diessel's (2015) characterisation of grammar as a network model (see Section 1.1.3). In both conceptions, constructions that share features are related. Goldberg summarises these relations as linguistic, and by Diessel's interpretation these relationships also reflect the cognitive architecture of grammar.

1.3 Argument structure constructions

This section reviews evidence that language users can access the semantics of argument structure constructions independently of the lexical items they contain. This evidence informs the approach to argument structure taken in the current research, where argument structure constructions are considered to be linguistic units independent of particular verbs.

This review discusses a number of studies whose conclusions bear on the interpretation of argument structure constructions as semantic units, rather than an exhaustive account of the subject. Studies were selected to provide an overview of the range of methods that have been used and constructions that have been investigated to date.

1.3.1 Evidence from reader judgements

A number of recent studies have investigated the semantic contribution that argument structure constructions supply to sentences, via ratings and reader judgements.

Kaschak and Glenberg (2000) showed that the meaning of argument structure constructions can constrain the interpretation of verbs. They designed sentences that contained novel verbs derived from real nouns, such as *crutched*. Readers were presented with these verbs in the ditransitive and transitive constructions, as shown in (6):

(6a) Lyn crutched Tom her apple so he wouldn't starve	ditransitive
(6b) Lyn crutched her apple so Tom wouldn't starve	transitive

In a sentence-choice task, participants decided which of the sentences in (6) was consistent with the statement *Tom got the apple* or *Lyn acted on the apple*. In a meaning-choice task, participants selected a definition for the novel verb, either *to transfer using a crutch* or *to act on using a crutch.* In both tasks, participants were significantly more likely to associate the ditransitive sentence with the meaning of transfer, a result consistent with the semantics of the ditransitive construction.

In a second experiment, Kaschak and Glenberg (2000) included sentences like those in (6) at the conclusion of a narrative that described a scene of potential transfer. They asked participants to paraphrase the final sentence, or define the novel verb. Participants were significantly more likely to refer to transfer in their responses when the novel verb appeared in the ditransitive construction rather than the transitive construction. In their two experiments, Kaschak and Glenberg demonstrated the relationship between the ditransitive construction and the meaning of transfer, independently of lexical verbs. The semantics of the ditransitive argument structure construction influenced readers' interpretation of novel verbs.

Bencini and Goldberg (2000) explored the effect of a wider range of argument structure constructions on readers' interpretation of sentences. They produced a set of sixteen sentences, composed of the verbs *get*, *slice*, *take* and *throw* presented in the caused motion, ditransitive, resultative and transitive constructions. Participants were asked to sort the sixteen sentences into four groups based on the overall meaning of the sentence.

In their first experiment, seven of the seventeen participants sorted sentences based on the meaning of the argument structure construction rather than the verb. Ten participants produced mixed groupings, which were significantly more similar to groupings based on the meanings of the constructions than the verbs. In their second experiment, six of another set of seventeen participants produced groupings based on the meaning of the construction. These six participants were all able to describe the semantics of the constructions. One participant described the resultative construction as 'a person…doing something to an object and the object changes', and another participant described the transitive as 'one person…doing an action with an object' (p. 648). Bencini and Goldberg's study demonstrated that readers can recognise the meaning of a range of argument structure constructions.

Goldwater and Markman (2009) investigated the contribution of the semantics of the passive and middle constructions to readers' interpretations of novel verbs. Their study was based on the previous finding by Mauner and Koenig (2000) that participants judged more sentences in the middle construction, like (7a), than the passive construction, like (7b), to be nonsensical:

(7a) The clocks had sold quickly, but no one sold them   middle

(7b) The clocks were sold quickly, but no one sold them   passive

Goldwater and Markman reasoned that the replication of this finding for sentences containing novel verbs would constitute evidence for the middle and passive as independent linguistic constructions.

Goldwater and Markman (2009) introduced their participants to novel verbs derived from real nouns. For example, the verb *to sauce* was presented in the sentence *those tomatoes sauce easily* to refer to the process of turning a fruit or vegetable into a sauce. Participants then judged whether sentences containing these verbs in the middle construction, like (7c), and the passive construction, like (7d), were nonsensical.

(7c) The ripe tomatoes had sauced expertly to complement the pasta at the gala dinner   middle

(7d) The ripe tomatoes were sauced expertly to complement the pasta at the gala dinner   passive

As predicted, participants judged more sentences in the middle construction than the passive construction to be nonsensical, showing that readers interpreted the constructions containing novel verbs in the same way as the constructions containing real verbs. Goldwater and

Markman's (2009) results add the middle and passive to the inventory of argument structure constructions that have been shown to influence readers' sentence interpretations.

In addition to influencing readers' interpretations of the event-level meaning of sentences, argument structure constructions can also affect how readers understand the nouns that sentences contain. Kako (2006) carried out a series of experiments in which native English speakers rated nouns in terms of how likely they were to have certain semantic properties that were associated with agents and patients. Agents can typically be considered sentient, mobile or the causer of a change of state, while patients can typically undergo changes of state and be causally affected by others (Dowty, 1991). Kako (2006) found that participants consistently judged syntactic subjects to be more agent-like than syntactic objects, and objects to be more patient-like than subjects. The semantic properties associated with a noun can be attributed to its position in the transitive argument structure construction.

Kako's (2006) final experiment is particularly noteworthy. He investigated participants' judgements of sentences that contained intransitive and nonsense verbs presented in the transitive and intransitive constructions. Participants were asked how likely it was that the verb in each sentence had properties associated with real transitive verbs, such as involving something being created, physically changing or making physical contact with something else. Participants rated verbs in the transitive construction as more transitive than verbs in the intransitive construction, despite all verbs in the experiment being exclusively intransitive or nonsense. Importantly, in the transitive construction, nonsense verbs were rated as more transitive than intransitive verbs. These findings indicate that participants did not treat intransitive verbs in the transitive construction as completely transitive. Rather, they attempted to integrate the semantics of the intransitive verb and the transitive construction in a meaningful way. Overall, Kako's (2006) work showed that the transitive construction influenced readers' interpretations of both nouns and verbs.

Taken together, the studies reviewed in this section show that argument structure constructions contribute to the meaning of sentences independently of the words the sentences contain, including verbs. The next section will examine evidence from online methodologies.

1.3.2 Evidence from priming paradigms

The studies in this section provide further evidence that argument structure constructions are independent, meaningful units of language. While the evidence reviewed in the previous section relied on readers' explicit judgements, the research presented in this section reflects participants' unconscious processing of language.

1.3.2.1 Argument structure constructions prime verbs related to their meaning

Johnson and Goldberg (2013) investigated whether the syntactic structure of argument structure constructions facilitate the processing of lexical items related to the construction. The authors composed written sentences that took the form of the caused motion, ditransitive, removal and resultative constructions and used non-words to represent open class lexical items. For each construction, Johnson and Goldberg identified three related verbs. A high-frequency associate was a verb that occurred in the construction with high frequency. A low-frequency associate was a verb that occurred in the construction with low frequency. A semantically related non-associate was a verb that was not attested in the construction but had a meaning related to the semantics of the construction. The constructions, sentence stimuli and verb associates employed in the study are shown in Table 1.4.

Table 1.4 Materials from Johnson and Goldberg's (2013) priming study

Argument structure construction	Sentence stimulus	High-frequency associate	Low-frequency associate	Semantically related non-associate
Caused motion	*He lorped it on the molp*	put	placed	decorated
Ditransitive	*He daxed her the norp*	gave	handed	transferred
Removal	*She vakoed it from her*	took	removed	ousted
Resultative	*She jorped it miggy*	made	turned	transformed

Participants read a sentence stimulus aloud and then made a lexical decision on a subsequently presented verb. Participants were faster at recognising verbs when they were preceded by a related argument structure construction. By-participant analyses revealed significant priming effects for all three types of associates, and by-item analyses revealed significant priming effects for high-frequency and low-frequency associates. Johnson and Goldberg (2013) concluded that speakers accessed the meaning of argument structure constructions automatically. This meaning can only be attributed to the structure of the construction, because the sentence stimuli did not contain any real open class words. Results also indicated that readers have implicit knowledge of the verbs associated with particular argument structure constructions, providing evidence for Diessel's (2015) lexical links between words and syntactic constructions in his network architecture of language.

1.3.2.2 Structural priming

The studies reviewed so far in this section have provided evidence for argument structure constructions as meaningful linguistic units from tasks involving sentence comprehension.

The current section considers evidence from structural priming of sentence production as indicating speakers' sensitivity to argument structure constructions.

Structural priming refers to speakers' tendency to repeat or reuse recently encountered syntactic structures. Levelt and Kelter (1982) undertook one of the first investigations of this phenomenon. They found that speakers' replies were likely to match the surface form of questions they were asked. Shopkeepers in the Netherlands were asked the Dutch equivalent of either *what time does your shop close* or *at what time does your shop close*. Responses were more likely to take the form of, for example, *five o'clock* if the form of the question lacked a preposition, or *at five o'clock* if the form of the question contained a preposition. Similar findings were obtained under laboratory conditions.

Bock (1986) developed an experimental method to investigate structural priming of complex syntactic structures. In her task, participants were asked to repeat sentences and then spontaneously produce sentences to describe pictures, under the guise of a memory test. Researchers analysed the form of participants' picture descriptions. Structural priming was present if the form of participants' spontaneous productions matched the form of the sentences they repeated. This paradigm tested for structural priming in the description of events that could be encoded by either an active or passive sentence, such as *the player hit the ball* or *the ball was hit by the player*, or by a ditransitive sentence that contained two noun phrases, or a preposition phrase, such as *Sam sent Alex a letter* or *Sam sent a letter to Alex.* Bock (1986) found evidence of structural priming for these structures, despite a lack of shared lexical material between prime and target sentences. This finding lends support to the construction grammar framework, which recognises the caused motion, ditransitive, transitive and passive as independent argument structure constructions.

Bock's (1986) study identified a number of semantic influences in operation during structural priming. In the three experiments reported in the study, structural priming of the passive construction was influenced by the animacy of the agent in the prime sentence. In the first experiment, participants' production of passive sentences was highly correlated with target pictures depicting non-human agents. In the second experiment, there was no effect of priming of the passive construction when participants described events with human agents. In the third experiment, significantly more passive constructions were used to describe pictures of events with non-human agents than human agents. Bock, Loebell and Morey (1992) investigated the role of such conceptual influences in more detail and found that the animacy of syntactic subjects function as effective primes in structural priming. That is, sentences with animate subjects prime sentence productions that contain animate subjects. The same held for inanimate subjects, and the effect was found for both active and passive sentence structures.

Findings from other studies run counter to the conclusion that conceptual or semantic factors influence structural priming. Bock and Loebell (1990) investigated the effect of prime sentences such as those in (8).

(8a) The wealthy widow gave her Mercedes to the church

(8b) The wealthy widow drove her Mercedes to the church

(8c) The 747 was alerted by the control tower

(8d) The 747 was landed by the control tower

(8a) and (8b) both take the form *subject-verb-object-PP*, but the preposition phrase in (8a) encodes a beneficiary and the preposition phrase in (8b) encodes a location. (8c) and (8d) both take the form *subject-*be_{aux}*-*$verb_{PP}$*-*PP_{by} , but the *by*-phrase in (8c) encodes the agent of a passive construction and the *by*-phrase in (8d) encodes a location. Despite these semantic differences, the authors found that the two versions of each sentence form primed the production of sentences of the same form to the same degree. Bock and Loebell (1990) interpreted these findings as evidence of syntactic representations independent of conceptual information.

In response to this study, Hare and Goldberg (1999) investigated the effect of what they termed a 'fulfilling frame', shown in (9), as a prime sentence.

(9) The officers provided the soldiers with guns

(9) has the same form as the sentences in (8a-b), namely, they all take the form *subject-verb-object-PP*. However, the order of the thematic roles in (9) parallels the order of roles in the ditransitive construction, namely *agent-(verb)-recipient-patient*. Hare and Goldberg (1999) reasoned that if sentences like (9) primed the production of the ditransitive construction, this would constitute evidence for a semantic influence in structural priming. Consistent with this hypothesis, 83% of responses following a fulfilling frame prime were ditransitive productions. That is, participants were not simply matching the syntactic structure of the prime in their responses, but were sensitive to the semantic roles of the construction as well.

Chang, Bock and Goldberg (2003) noted that Hare and Goldberg's (1999) results could be explained by the association of animacy with recipients and inanimacy with patients in the experimental and filler materials included in the study. To rectify this confound, they utilised the English load-spray alternation, in which two semantically synonymous structures differ in thematic role order but not in animacy, as in (10).

(10a) The maid rubbed polish onto the table theme-location

(10b) The maid rubbed the table with polish location-theme

Both sentences in (10) have the form *subject-verb-NP-PP*. The post-verbal arguments are both inanimate, but in (10a) they take the order of theme and location, while in (10b) they take the order of location and theme. Chang et al. (2003) observed priming based on the order of the thematic roles present in the prime sentence.

The construction grammar framework can accommodate the range of findings from the structural priming paradigm. Structural priming has been observed to operate over the linear order of structural syntactic forms as well as meaningful semantic roles. Constructions are defined as the association between such structural forms and meaningful elements, so priming could occur over either dimension of argument structure constructions (Casenhiser & Bencini, 2015). Crucially, structural priming occurs when prime and target sentences contain different lexical verbs. This points to the existence of an abstract structural form that is independent of particular lexical items. In construction grammar, this form is defined as an argument structure construction.

Studies that claim structural priming occurs outside the influence of semantics overlook two important considerations. First, the closed class items contained in a construction may lead to a high degree of shared phonological content between prime and target. For example, (8c) and (8d) have three points of phonological similarity, including an auxiliary, the morpheme of the past participle of the verb and the preposition *by*. These points of similarity could explain the observed priming effects, despite the difference in meaning between the sentences. Second, studies in structural priming do not account for the effect of past language experience on participants' in situ performance. Findings from Bock (1986) and Bock et al. (1992) could be explained by the fact that the passive construction frequently encodes non-animate agents, and syntactic subjects often encode animate real-word entities. Though this suggestion is currently empirically unsubstantiated, the effects of language use and frequency of language structures were not accounted for in studies of structural priming.

Before concluding this section, more recent studies show that structural priming also occurs in the domain of sentence comprehension. The comprehension of a sentence with a particular syntactic structure eases subsequent comprehension of the same structure, as measured by eye movements and EEG (Tooley & Traxler, 2010), and influences the analysis of sentences with multiple interpretations (Pickering, McLean & Branigan, 2013).

1.3.3 Evidence from neurolinguistics

The final piece of evidence for the independence of argument structure constructions as linguistic units that will be discussed in this section comes from brain imaging methodologies.

Multi-voxel pattern analyses (MVPA) is an approach that is gaining popularity in the evaluation of fMRI data (Coutanche, 2013). MVPA is a type of analysis that allows the simultaneous measurement of brain activity in a variety of locations. Whereas conventional imaging techniques account for information in only one brain region at a time, MVPA preserves important information about the patterns of brain responses. MVPA is more sensitive and specific than traditional fMRI analyses (Haynes & Rees, 2006).

MVPA has been used successfully to investigate linguistic representations in the domains of semantics (Wang, Baucom & Shinkareva, 2013) and speech comprehension (Abrams et al., 2013). Allen, Pereira, Botvinick and Goldberg (2012) employed the technique to investigate the processing of semantically similar but syntactically distinct grammatical constructions. They compared the ditransitive and caused motion constructions, as in (11):

(11a) She threw him something	ditransitive
(11b) She threw something to him	caused motion

The brain regions under investigation included Broca's area (Brodmann Area (BA) 44 and 45), Wernicke's area (posterior BA 22), anterior BA 22 and BA 47.

On each trial, participants read three sentences of either single construction type. After reading these sentences, they determined whether a probe sentence was the same as the previous sentences in the trial, had the same meaning but was not matched word-for-word, or neither. A control task presented three sets of scrambled words from ditransitive or caused motion sentences. The probe for the control task required participants to decide whether a word matched a word in the sets, had the same meaning as a word in the sets, or neither.

MVPA successfully distinguished between the ditransitive and caused motion constructions. The same words appeared in each construction, so this finding does not depend on the lexical items contained in the sentence stimuli. Results from the control task were not significant, so the difference between the ditransitive and caused motion constructions was not attributable to differences in length or morphology between the sentences. Neither construction was more frequent than the other, so frequency could not account for the finding, either.

Follow-up analyses on the regions of interest revealed that the anterior portion of BA 22 and BA 47 could distinguish between the constructions in combination, but neither area could do so individually or for the control task. Importantly, BA 22 and BA 47 have been implicated in semantic processing (Booth et al., 2002; Chou et al., 2006; Howard et al., 1992; Price et al., 1992). Allen et al. (2012) argued that their finding supports the claim of construction grammar that argument structure constructions differ syntactically as well as semantically. Because there was no difference between the two constructions on behavioural measures, including accuracy and response times to the probe question on each trial, this finding likely reflects a genuine qualitative difference between the two constructions, rather than the differential complexity of processing one construction over the other. In particular, the ditransitive construction is associated with the semantics of transfer, while the caused motion construction encodes a path of motion.

Johnson, Turk-Browne and Goldberg (2013, 2016) used fMRI to identify neural areas that are involved in the learning of argument structure constructions. The authors presented adult participants with video clips of a novel event that is not encoded by an English construction, that of sudden appearance. Participants saw videos in which a character suddenly appeared near another object. The event was described by a construction that took the form *verb-NP-NP*, where the first noun phrase encoded the theme and the second noun phrase encoded a locative. In one condition, participants heard this construction when viewing events of appearance. In another condition, participants heard the words of the construction in a random order when viewing events of appearance. Upon testing, higher accuracy was correlated with less activity in ventral striatal areas, which have been associated with estimating the difference between predictions and outcomes. The authors concluded that successful learning is defined by the ability to accurately predict upcoming structures. Over the course of the experiment, participants in the first condition displayed increasing activation in the posterior precuneus, an area that has been associated with pattern learning (Casenhiser & Bencini, 2015). In sum, Johnson and colleagues showed that real-world events can be mapped onto grammatical structures. Argument structure constructions are learnable via neural mechanisms associated with pattern prediction.

*

This chapter introduced usage-based approaches to language as a strand of current thinking in the field of linguistics. These approaches consider frequency of occurrence to be a driving force behind language structure, which can arise from domain-general cognitive mechanisms. The basic unit of language can be termed a construction, and constructions are form-meaning pairings that vary in complexity and abstractness. The final section of this chapter reviewed evidence that language users recognise sentence-level argument structure constructions as meaningful linguistic structures independently of the lexical items they contain.

The next chapter introduces the acquired language disorder aphasia and reviews evidence that frequency of occurrence influences language processing in typical adults and adults with acquired aphasia.

2 Language processing in typical adults and adults with acquired aphasia

This chapter reviews how frequency affects language processing in typical adults and adults with acquired aphasia. The first section provides a background to the acquired language impairment aphasia and describes deficits in the processing of verbs and argument structure that it can cause, as well as comments on how these deficits have been approached in aphasia research to date. The second section reviews the effect of frequency on language processing in typical and aphasic populations. The chapter concludes with the aims of the current investigation.

2.1 Acquired aphasia

The term aphasia refers to 'a family of clinically diverse disorders that affect the ability to communicate by oral or written language, or both, following brain damage' (Goodglass, 1993, p. 1). Aphasia can result from a variety of neural insults including infection, trauma, tumours, lack of blood supply or haemorrhage at any stage of life, but aphasia most commonly results from stroke in older adults (Goodglass, 1993). About one-third of stroke survivors experience some degree of aphasic symptoms (Brady, Kelly, Godwin & Enderby, 2012). In the United Kingdom, this equates to a total of over 400,000 individuals (Stroke Association, 2017).

Impairment to left-hemisphere neural areas, notably perisylvian regions, most commonly causes the language impairments observed in aphasia (Dell & Chang, 2014). Records of acquired speech disorders have been evident throughout medical history since ancient times (c. 3500 BC; Goodglass, 1993), but the association between left-hemisphere brain damage and the linguistic deficits that characterise aphasia was first noted in the nineteenth century, in the work of Marc Dax (1836), Paul Broca (1861) and Carl Wernicke (1874) (Goodglass, 1993). Around this time (1871), Heymann Steinthal advocated for a more detailed description of language in aphasia than physicians provided and distinguished between single-word and sentence-level deficits in aphasia (Code, 2013).

2.1.1 Language in aphasia

2.1.1.1 Classifications of language in aphasia

Researchers throughout the decades have proposed various systems of classification to describe different types of aphasia. A main goal of most classification systems was linking the locus of neuropathology to language behaviour; however, such diagnoses proved generally unreliable. Agreement between outcomes on two of the most influential systems of aphasia classification, the Western Aphasia Battery (Kertesz, 1982) and the Boston Diagnostic Aphasia Examination (Goodglass & Kaplan, 1972), and patients' actual symptomatology can be as low as 30% (Crary, Wertz & Deal, 1992), and the subtests included in such assessments may not provide appropriately sufficient or specific information to inform treatment (Byng, Kay, Edmundson & Scott, 1990). Furthermore, the advent of contemporary brain imaging techniques has rendered the goal of deriving lesion locations from language behaviour largely obsolete (Code, 2013). Nonetheless, the Boston Diagnostic Aphasia Examination (Goodglass & Kaplan, 1972) remains in use today and provides information for classifying the linguistic profiles of individuals with aphasia into types. This system will be briefly described below.

The most major distinction among types of aphasia is based on the fluency of an individual's language output. Fluent speech is similar to unimpaired speech in its rate, utterance length, melodic contour and production ease (Damasio, 1998). 'Press of speech' refers to fluent production that sounds faster than normal, or 'hyper-fluent', though the speech rate of individuals with fluent aphasia has been observed to be the same as typical speakers (Edwards, 2005). Four types of aphasia are characterised by fluent speech. The most mild form of fluent aphasia is anomic aphasia, where the dominant impairment is word finding difficulties (Damasio, 1998). Word finding difficulties, or anomia, are however one of the most pervasive and persistent symptoms of aphasia, and they are experienced to some degree by nearly all speakers with aphasia (Goodglass & Wingfield, 1997). Conduction aphasia refers to a language impairment characterised primarily by a deficit in repetition and the production of phonemic paraphasias but relatively preserved auditory comprehension. In transcortical sensory aphasia, repetition is preserved but auditory comprehension is impaired (Damasio, 1998). Wernicke's aphasia may be considered one of the most severe types of fluent aphasia, as auditory comprehension impairments can be significant (Damasio & Geschwind, 1984) and conversational strategies such as repair can be atypical (Ferguson, 1998).

In contrast, non-fluent speech refers to a slow speech rate, short utterance length, loss of melodic contour and effortful production (Damasio, 1998). Three types of aphasia are characterised by non-fluent speech. Individuals with Broca's aphasia have relatively well preserved auditory comprehension (Damasio & Geschwind, 1984). This aphasia type is commonly associated language output that can be described as 'agrammatic', or lacking in function words, inflectional morphology and the variety of syntactic structures attested normally (Goodglass, 1997). In transcortical motor aphasia, non-fluent speech is coupled with a mild impairment in auditory comprehension. Global aphasia refers to an almost complete inability to participate in verbal communication (Damasio, 1998).

2.1.1.2 Verb processing in aphasia

The present research investigates the processing of verbs and argument structure constructions. This and the following sections review how the processing of verbs and argument structure can be disrupted in individuals with aphasia. As an introduction, some differences between the processing of nouns and verbs are discussed.

Differences between nouns and verbs

Nouns and verbs are recognised as distinct grammatical classes in linguistics due to differences in semantics, syntax and morphology. Semantically, a noun typically refers to an object, whereas a verb tends to refer to an action. Section 1.2.3.2 explained how sentences arise from the integration of verbs and argument structure constructions, and therefore verbs have a closer relationship to syntactic structure than do nouns. With regard to morphology, verbs in English occur with a greater number of inflections than do nouns. While a regular verb, such as *dance*, can inflect for tense (*danced*), person (*she dances*) and aspect (*dancing*), a regular noun, such as *dog*, can inflect only for number (*dogs*).

Though nouns and verbs are presented as distinct grammatical classes, their differences are not entirely dichotomous. Rather, phonological, conceptual and syntactic differences may converge to give the appearance of distinct syntactic categories (Black & Chiat, 2003). After accounting for the conflation between nouns and verbs as semantic categories referring to objects and actions, Vigliocco, Vinson, Druks, Barber and Cappa (2011) concluded that nouns and verbs are processed by a shared neural system.

Nonetheless, the processing of verbs is reported to be more demanding than the processing of nouns, as revealed by studies that compare the processing of actions and objects. Response times from typical speakers are longer to verbs than nouns (Szekely et al., 2005), and both typical adults and adults with aphasia produce the names of verbs more slowly and less accurately than nouns (Mätzig, Druks, Masterson & Vigliocco, 2009). Studies in aphasia indicate a dissociation between noun and verb processing. Luzzatti et al. (2002) reported overall worse performance in verb naming than noun naming in a group of 58 Italian adults with aphasia. Within this group, 26 individuals showed differential performance in the processing of verbs and nouns, as six participants showed superior performance on verbs and twenty showed superior performance on nouns. Differences in the processing of nouns and verbs may be due to differences in the collection of semantic, morphological and syntactic attributes that each word class attracts, and these differences are manifested as processing differences between the two classes of words.

Verb processing in aphasia

Verb retrieval difficulties are well attested in aphasia. In line with the findings of Luzzatti et al. (2002), many studies that compare differences between verb and noun processing report greater demands associated with verbs. Two individuals with fluent aphasia, HW and SJB, were worse when naming verbs than nouns (Caramazza and Hillis, 1991). In writing, GOS made a greater number of errors when spelling verbs compared to adjectives and function words (Baxter & Warrington, 1985).

Selective deficits in verb retrieval are commonly associated with agrammatism and non-fluent aphasia, while selective deficits in noun retrieval are associated with anomia and fluent aphasia (Druks, 2002). Agrammatic speakers show verb deficits in production tasks (Miceli, Silveri, Villa & Caramazza, 1984; Kim & Thompson, 2000; Zingeser & Berndt, 1990), as well as in both production and comprehension tasks (McCarthy & Warrington, 1985; Miceli, Silveri, Nocentini & Caramazza, 1988).

However, verb impairments are not necessarily limited to a single type of aphasia. In one group of five speakers with selective deficits in verb production, three were diagnosed with agrammatic Broca's aphasia and two with Wernicke's aphasia (Berndt, Mitchum, Haendiges & Sandson, 1997a). Williams and Canter (1987) found that groups of individuals with Broca's, Wernicke's, anomic and conduction aphasia all scored significantly lower on tests of action naming than object naming. Similarly, groups of agrammatic and anomic speakers both showed poorer performance on action than object naming (Bastiaanse & Jonkers, 1998). Additionally, Wambaugh, Doyle, Martinez and Kalinyak-Fliszar (2002) considered an intervention targeting verb retrieval to be appropriate for individuals with anomic, Wernicke's and conduction aphasia (Wambaugh et al., 2001). Though verb retrieval deficits may be more pronounced in speakers with non-fluent compared to fluent aphasia (Conroy, Sage & Lambon Ralph, 2006), they can affect speakers with aphasia regardless of its classification.

Difficulties in verb production do not necessarily entail difficulties in verb comprehension. Berndt, Mitchum, Haendiges and Sandson (1997b) explored differences in the production and comprehension of nouns and verbs in a sample of 11 adults with aphasia. Five individuals showed more difficulty producing verbs than nouns but only two showed evidence of verb comprehension difficulties. A single participant, HY, performed significantly worse on verbs than nouns in comprehension tasks, but showed the opposite pattern in production, where he correctly produced more verbs than nouns.

In sum, differences between nouns and verbs may arise due to differences in phonology, semantics and syntax between the two grammatical classes, and these differences converge to make verb processing more demanding than noun processing across different types of aphasia.

2.1.1.3 Argument structure processing in aphasia

Argument structure refers to the set of syntactic phrases included in a grammatical sentence. Section 1.3 described how argument structures are interpreted as independent, meaningful linguistic units in the construction grammar framework. In research on aphasia, investigations of argument structure have explored the intransitive, in sentences that contain only a subject and verb (e.g. *Sam sneezed*); transitive, in sentences that contain a subject and an object (e.g. *Sam loves Alex*); passive, in sentences where the syntactic subject encodes a theme (e.g. *The letter was sent by Sam*); ditransitive, in sentences where post-verbal indirect and direct objects are both encoded as noun phrases (e.g. *Sam sent Alex a letter*); and prepositional dative, which corresponds to the caused motion construction discussed in Chapter 1 (e.g. *Sam sent a letter to Alex*).

In aphasia, specific difficulties with argument structure can be independent of verb processing difficulties. YR (Whitworth, Webster & Howard, 2015) showed unimpaired access to and production of nouns and verbs as single words but poor production of argument structure in connected speech. Similarly, JM (Webster, Franklin & Howard, 2004) displayed errors in generating accurate argument structure in sentence production tasks with coincidental near-normal performance in single-word processing tasks. In contrast to these two individuals, IB (Jensen, 2000) scored poorly in a test of verb retrieval and often omitted verbs in the sentence context but was nonetheless able to generate appropriate nouns in the transitive construction. The performance of these three individuals shows that difficulties in processing verbs and argument structures can occur independently.

This section continues with a survey of research on the production and comprehension of argument structure in aphasia.

Production of argument structures

What may appear as a paucity of argument structure in the spontaneous speech of adults with aphasia may not be attributable to total loss of these constructions. Using the structural priming paradigm, described in Section 1.3.2.2, Saffran and Martin (1997) investigated the transitive, passive, ditransitive and prepositional dative constructions in two speakers with agrammatism and three speakers with fluent aphasia. Participants produced significantly fewer instances of the transitive construction following passive primes, and more passive productions following passive primes. Hartsuiker and Kolk (1998) investigated the same set of constructions in a larger study with twelve speakers with Broca's aphasia and reported a wider range of effects: participants demonstrated priming effects for the passive, ditransitive and prepositional dative constructions. Hartsuiker and Kolk concluded that priming was the result of an unconscious, automatic, facilitatory process. Taken together, these studies demonstrate that speakers with aphasia can have preserved but unconscious knowledge of argument structure constructions that may be inaccessible as a result of their language impairment.

Aphasic impairments in sentence production can arise from an inability to correctly realise thematic roles as syntactic phrases (see also information on the mapping hypothesis in Section 2.1.2.1). Barbieri, Basso, Frustaci and Luzzatti (2010) investigated the ability of seven Italian adults with aphasia to produce transitive sentences with prepositional adjuncts. In a picture description task, target productions were sentences like *the man slices the bread with the knife*. Compared to typical participants, four participants with non-fluent aphasia and one participant with fluent aphasia were significantly impaired in relating thematic roles to sentence structure. Though they produced target verbs accurately, participants frequently generated sentences in which the instrument was realised as the direct object of the verb rather than the object of the preposition, such as *the girl sews a thread for the button* rather than *a woman sews a button with a thread*. Participants demonstrated disruption in linking the participants of a scene to their argument roles in the transitive construction.

The relationship between verbs and the number and type of argument structures with which they are associated can affect language processing in aphasia. Thompson, Lange, Schneider and Shapiro (1997) examined six types of verbs, including those that could only occur with an agent, like *smile*; an agent and a patient, like *open;* or an agent, patient and location, like *lean*. Another set of verbs could be used with multiple unique argument structures, including optional patients, like *eat*; optional goals, like *teach*; and optional sentence complements, like *know*. In a naming task, speakers with aphasia demonstrated a hierarchy of verb difficulty such that verbs like *give* and verbs that took sentence complements were the most difficult to name; verbs like *open* were less difficult to name; and verbs like *smile* were named more correctly than any other verb type. In a sentence production task, participants with and without aphasia correctly produced sentences with verbs that could occur in only one argument structure more often than verbs that could occur with optional arguments. Overall, speakers with agrammatism found verbs with fewer arguments and less variable argument structure easier to produce, and these factors influenced sentence production in typical speakers, as well. This hierarchy of verb difficulty was confirmed in a subsequent study with seven agrammatic speakers (Kim & Thompson, 2000). Consistent with these findings, DiLallo, Mettler and DeDe (2017) reported that participants with aphasia produced more intransitive than transitive verbs in spontaneous speech.

Thompson (2003) described the processing difficulties that accompany the hierarchy of verb complexity as the argument structure complexity hypothesis. She noted that verb complexity can refer to the number as well as the type of arguments encoded by a verb, as she observed a verb production asymmetry between unergative and unaccusative verbs. Unergative verbs occur with syntactic subjects that correspond to semantic agents, like *laugh*, and unaccusative verbs occur with syntactic subjects that correspond to semantic themes, like *melt*. A group of eight adults with agrammatism showed significantly poorer naming of unaccusative than unergative verbs, and the pattern was evident in each individual.

Research on argument structure production in aphasia suggests that argument structures may be inaccessible but not entirely lost to speakers with aphasia, that difficulties may arise between mapping thematic roles onto syntactic structure and that difficulties increase with the number and type of argument structures with which verbs are associated.

Comprehension of argument structures

Argument structure complexity affects sentence comprehension, as well as production. Shapiro and Levine (1990) examined the processing of a variety of verb types in readers with Broca's aphasia and fluent aphasia. Participants with Broca's aphasia showed increased processing of verbs that could occur in a variety of argument structures compared to verbs that occur in only a single argument structure. Typical participants showed the same effects as these non-fluent participants. In contrast, participants with fluent aphasia did not appear to be sensitive to this argument structure information in reading.

This section described how the production and comprehension of argument structure can pose difficulty for adults with aphasia independently of verb processing impairments. The next section continues with a survey of grammatical abilities in aphasia, as most investigations of verb and argument structure processing in aphasia are undertaken with reference to this area.

2.1.2 Grammatical abilities in aphasia

Two terms refer specifically to the nature of grammatical difficulties in fluent and non-fluent aphasia. Agrammatism describes a pattern of non-fluent speech characterised by the omission of function words, inflectional morphology, limited syntactic structure and overuse of infinitival verb forms (Goodglass, 1997). The term can also encompass impaired comprehension of sentences that are thought to depend solely on syntactic analysis for interpretation, such as the passive construction (Berndt, 1998). Paragrammatism describes a pattern found in fluent speech that involves atypical use of syntactic structures but comparably minor deficits in the production of morphology and function words (Goodglass, 1997).

2.1.2.1 Theories of agrammatism

Numerous researchers have proposed theories to explain the nature of agrammatic deficits in aphasia. This section continues by briefly reviewing some of these theories and the data on which they are based. The following section discusses some criticisms of the usefulness of the term agrammatism.

Morphological processing deficits

Deficits in the processing of inflectional morphology and function words is one of the trademark signs of agrammatism. Goodenough, Zurif and Weintraub (1977) observed that agrammatic individuals displayed impaired understanding of articles and concluded that agrammatism may be attributable to a loss in processing function words, because they have very low semantic content. Goodglass and Berko (1960) investigated the production of inflectional morphology in aphasia and uncovered a pattern in which possessive forms posed the most difficulty for agrammatic speakers and plural forms the least difficulty. Friederici, Schönle and Garrett (1982) noted impairments in the comprehension and production of prepositions in agrammatism. Differences in the processing of morphemes, including function words, may underpin agrammatic language impairment.

Linear assignment of thematic roles

Caplan and Futter (1986) examined thematic role assignment during the comprehension of auditory sentences in the agrammatic listener SP. SP assigned the thematic roles of agent, theme and goal to noun phrases depending on their linear position in a sentence. By this method, SP correctly interpreted prepositional dative sentences like *the monkey gave the frog to the elephant* but incorrectly interpreted ditransitive sentences like *the monkey gave the elephant the frog*, indicating *elephant* as the theme and *frog* as the goal. The authors concluded SP's performance was the result of strategic assignment of thematic roles to sentence structure.

Mapping hypothesis

The relationship, or mapping, between semantic roles and syntactic structure can be disrupted in agrammatism, as Caplan and Futter (1986) demonstrated. Schwartz, Linebarger, Saffran and Pate (1987) identified problems in mapping between syntactic structure and thematic roles as the cause of sentence comprehension difficulties in agrammatism. Their conclusion accords with that of Dominey (2002), who observed that language breakdown in agrammatism may lead to a compensatory strategy in which the mapping between the semantic agent and syntactic subject is overused in sentence comprehension. In Dominey's study, a group of nine adults with agrammatism responded more accurately in a sentence-to-picture matching task to canonical sentences, where the order of nouns in the sentence reflected the agent, object and recipient, and he successfully replicated their performance in a neurocomputational model by implementing a bias that assigned the first noun the thematic role of agent. Alternatively, Kielar, Meltzer-Asscher and Thompson (2012) concluded that the inability to detect argument structure violations in aphasia, for sentences like *Anne sneezed the doctor and the nurse*, could be due to impaired integration of arguments into sentences as a result of insufficient access to the lexical specification of a verb.

Bastiaanse and van Zonneveld (2005) reached a conclusion similar to that of Dominey (2002) regarding overuse of the mapping between an agent and syntactic subject in sentence production by speakers with Broca's aphasia. Verbs with what the authors termed alternating transitivity can be realised in both transitive and intransitive sentences, but the semantic role of the syntactic subject is different in each sentence type. The transitive sentence *the server breaks the glass* is unergative, because the subject corresponds to an agent. The intransitive sentence *the glass breaks* is unaccusative, because the subject corresponds to a theme. Speakers with Broca's aphasia made significantly more errors when producing unaccusative sentences than unergative sentences. Their main error was producing transitive sentences for unaccusative targets, in which they incorrectly applied the mapping of a semantic agent to the syntactic subject.

Trace-deletion hypothesis

Grodzinsky (1986) formulated the trace-deletion hypothesis within the framework of Chomsky's (1981) government and binding theory in order to explain the agrammatic deficit in comprehending reversible passive sentences, such as *Alex was kissed by Sam*. In Chomsky's (1981) theory, the subject of a passive sentence is derived from its original position as an object of a verb. This movement operation leaves a trace in the original post-verbal position, so the representation of the passive sentence above has the form *$Alex_i$ was kissed [t_i] by Sam*. According to the theory of government and binding, only the trace can be assigned a thematic role in the passive. Agrammatic difficulty with reversible sentences could be explained by the deletion of the trace, which is necessary to assign the subject of the sentence the thematic role of theme.

Tree pruning hypothesis

Friedmann and Grodzinsky (1997) articulated the tree pruning hypothesis in order to account for the observation that, crosslinguistically, speakers with agrammatic aphasia show more severely impaired production of grammatical morphemes related to tense than agreement of person, number or gender. Their case study of an agrammatic Hebrew speaker confirmed this pattern. The tree pruning hypothesis was devised with reference to the generative minimalist programme (Chomsky, 1992). In this view, agreement is subordinate to tense in the hierarchical structure of phrases, which is in turn subordinate to complementisers, e.g. *that* and *which*, essential for marking embedded clauses. The hypothesis holds that impairment on one level necessarily involves impairment at superordinate levels of the hierarchical phrase structure. Since agreement is at the lowest of the hierarchy, it can be selectively preserved compared to tense.

2.1.2.2 Criticisms of the concept of agrammatism

The theories of agrammatism outlined above assume that agrammatism can be analysed as a unitary phenomenon; that is, they seek to explain a single pattern of performance, which is assumed to be indicative of agrammatism. However, this assumption may be flawed in light of other findings. Analyses of fluent and non-fluent speech patterns have revealed types of sentence production common to both (Bird & Franklin, 1996), and individuals who show patterns of performance that constitute a double dissociation between the speech patterns and comprehension deficits considered agrammatic (Berndt, 1998). Further, Schwartz et al. (1987) found that individuals with fluent conduction aphasia showed the same ‘agrammatic’ pattern of sentence comprehension as individuals with agrammatic speech output.

The mapping hypothesis (Schwartz et al., 1987), the trace-deletion hypothesis (Grodzinsky, 1986) and Caplan and Futter’s (1986) linear order hypothesis account for the finding that agrammatic comprehension of canonical sentences types, where thematic roles are assigned to nouns following the order of agent-theme-goal, is superior to comprehension of noncanonical sentence types, where thematic roles deviate from this order. This pattern of performance, however, is not consistent for all individuals with agrammatism. Berndt, Mitchum and Haendiges (1996; see also Caramazza, Capitani, Rey & Berndt, 2001) carried out a meta-analysis in which they identified 42 individuals diagnosed with agrammatic aphasia who took part in a sentence comprehension task involving active and passive sentences. The researchers performed within-participant analyses in order to examine whether individuals performed at chance on active or passive sentences and identified three patterns of performance: better than chance comprehension of both sentence types, at chance comprehension of both sentence types and superior comprehension of active sentences compared to passive sentences. In another study, BM (Druks & Marshall, 1995) showed superior comprehension of passive sentences compared to active sentences. It appears, therefore, that the label agrammatism has been used to refer to several distinct patterns of sentence comprehension, and that a typically agrammatic pattern of comprehension can be found in speakers with, and those without agrammatic language production. These differences weaken the usefulness of the term.

2.1.3 The status of the usage-based approach to language in linguistic aphasiology

Chomsky's theories of transformational generative grammar have had a major influence on research in aphasiology. Two accounts of agrammatism mentioned in Section 2.1.2.1, the trace-deletion hypothesis and the tree pruning hypothesis, were formulated specifically with reference to Chomsky's proposed frameworks. However, the application of generative grammar to acquired language disorders is problematic. Chomsky (1980) maintains that 'knowledge of English' (p. 5) is preserved in individuals with aphasia who are both fully recovered and who show signs of language impairment, but the 'capacity' (p. 5) of the fully recovered individual has been restored, whereas it is still lacking in the individual who shows impaired language. In terms of the parlance introduced in Section 1.1.1, aphasia appears to affect a person's linguistic performance, but not competence. Chomsky (1980) notes that this conclusion 'would have to be based on evidence, of course, but not necessarily evidence from behaviour' (p. 5). It remains unclear whether Chomsky's descriptions of generative grammar are applicable to the 'capacity' that he argues is impaired in aphasia, and even more of a mystery as to what type of evidence Chomsky would consider to have a bearing on the issue, if not evidence from language behaviour.

Recently, the field of aphasiology has witnessed a rise in the application of usage-based approaches to the explanation of acquired language impairments. In 2014, Zimmerer, Dąbrowska, Romanowski, Blank and Varley advocated for the construction grammar framework over an account based on generative grammar in their analysis of sentence comprehension in a man with primary progressive aphasia. In 2016, a special issue of the journal *Aphasiology* (volume 30, number 11) was devoted to the topic of frequency in aphasic language processing. In this issue, Gahl and Menn (2016) demonstrated the applicability of usage-based approaches to sentence-level deficits in aphasia. They endorsed construction grammar as having the potential 'to broaden the scope of the discussion beyond the patterns of "agrammatic" deficits that have dominated the study of sentence-level deficits in aphasia in the past' (p. 1373). Specifically, they concluded that:

> If there is one thing that research on "normal" sentence processing has made clear, it is that language processing...[is] shaped by past experience—that is, by usage in real-world contexts...[T]he effects of past linguistic experience are evident in the language of people with aphasia, just as they are in the language of all other speakers (pp. 1372-1373).

The current work responds to this appeal for research on sentence-level processing in aphasia following a usage-based approach. In particular, the present work examines the role of frequency in the processing of verbs and argument structure constructions, where frequency is taken to index what Gahl and Menn (2016) refer to as 'effects of past linguistic experience'. The following sections survey how the frequency of linguistic units at varying levels of complexity affects language processing in typical adults and adults with aphasia.

2.2 Frequency effects in language processing

Frequency effects in language processing have been attested at linguistic grain sizes from the phoneme (e.g. Levitt & Healy, 1985) to the sentence (e.g. Conklin & Schmitt, 2012). This section reviews findings that frequency affects language processing at the level of the single word, multi-word phrase and syntactic construction. Each subsection first reviews evidence from language comprehension and then from production, beginning with a discussion of data from typical adults and then data from adults with aphasia.

2.2.1 Frequency effects in single-word processing

2.2.1.1 Lexical frequency effects in language comprehension

Data from typical adults

The frequency of occurrence of single words, termed lexical frequency, affects the recognition of single words. Typical adults show shorter response times to high frequency words than low frequency words in lexical decision tasks (Balota, Cortese, Sergent-Marshall, Spieler & Yap, 2004; Brysbaert, Lange & Van Wijnendaele, 2000; Gardner, Rothkopf, Lapan & Lafferty, 1987; Gordon & Caramazza, 1982; Grainger, 1990; Morrison & Ellis, 1995). Eye-tracking studies have shown that typical readers spend less time looking at high frequency words than low frequency words, including both nouns and verbs, in sentence contexts (Inhoff & Rayner, 1986; Juhasz & Rayner, 2003; Just & Carpenter, 1980; Rayner & Duffy, 1986). In an eye-tracking investigation of spoken word processing, Dahan, Magnuson and Tanenhaus (2001) showed that lexical frequency had an immediate effect on word recognition. Given a visual array of images of objects in combination with the auditory presentation of a word, participants were more likely to fixate on, and spend less time looking at, objects with high frequency names compared to low frequency names. A time course analysis revealed that these frequency effects affected the earliest stages of lexical access, based on the relationship between eye movement behaviour and the acoustic signal.

Data from adults with aphasia

Like typical adults, adults with aphasia have been shown to produce fewer errors and shorter response times to high frequency words than low frequency words on lexical decision tasks (Gerratt & Jones, 1987). Other studies have investigated lexical frequency effects in the context of sentence comprehension.

DeDe (2012) reported that a group of eight adults with aphasia showed longer response times to low frequency than high frequency words embedded in a sentence context. Results from the individual participants were more varied. In a self-paced listening task, seven of the eight participants produced longer response times to low- than high-frequency words, but one participant showed the reverse pattern, displaying longer response times to high frequency words. In a self-paced reading task, seven of the eight participants showed frequency effects to a greater extent than typical participants. Four of these participants produced longer response times to low frequency items, but three participants showed reverse frequency effects. Participants' patterns of performance did not relate to their age, severity of language impairment or aphasia type and could not be explained by deficits in single-word comprehension. Huck, Thompson, Cruice and Marshall (2017) reported an effect of lexical frequency on sentence reading, using eye-tracking with a group of 17 adults with aphasia. Specifically, Huck et al. investigated the effects of both lexical frequency and predictability on word reading in sentences. They found a significant effect of word frequency for typical adults and for adults with aphasia, but the effect of lexical frequency was significant in aphasia only for unpredictable words. The effect of lexical frequency was related to the nature of the language impairment in individuals with aphasia. The magnitude of the frequency effect was greater for participants with more severe language processing difficulties.

Reverse frequency effects in aphasic language comprehension

DeDe (2012) reported that some participants showed evidence for reverse frequency effects in self-paced reading and listening tasks; however, she cautions against over interpreting these findings, noting that only one participant showed consistent reverse frequency effects across both tasks. Hoffman, Rogers and Lambon Ralph (2011) explored why individuals with acquired aphasia may not show faster processing of high frequency than low frequency words, as expected based on data from typical adults. They observed that aphasia can impair the system of semantic control that ensures relevant meanings of words are accessed in communication; e.g. that the botanical meaning of *bark* is accessed in a conversation about gardening, rather than meaning of the noise a dog makes. Hoffman et al. showed that words with high lexical frequency appeared in more semantically diverse contexts than words with low lexical frequency. After the effect of semantic diversity was taken into account, participants with aphasia showed a typical frequency effect in a synonym judgement task, suggesting that reverse frequency effects in aphasia may actually result from difficulties in processing the greater number of semantic associations for words with high lexical frequency.

2.2.1.2 Lexical frequency effects in language production

Data from typical adults

The effect of lexical frequency on language production in typical speakers has been attested across a variety of tasks. It has long been known that naming latencies from typical speakers are shorter for pictures of objects with high frequency names than objects with low frequency names (Oldfield & Wingfield, 1965). Research has demonstrated that this is a reliable effect (Griffin & Bock, 1998; Huttenlocher & Kubicek, 1983; Lachman, 1973; Lachman, Shaffer & Hennrikus, 1974). Jescheniak and Levelt (1994; Wingfield, 1968) confirmed that frequency effects in picture naming cannot be attributed to processes of object recognition or articulation. Szekely et al. (2005) specifically examined the picture naming of actions and found that response times from typical speakers were actually longer to high frequency verbs than low frequency verbs. The authors admitted that this finding may reflect strategic responding, as participants produced high frequency 'light verbs' in response to pictures that were difficult to name.

In written word naming tasks, typical adults name high frequency written words faster than low frequency written words (Balota et al., 2004; Brysbaert, 1996; Brysbaert et al., 2000; Forster & Chambers, 1973; Grainger, 1990) and articulate high frequency words with shorter durations than low frequency words (Wright, 1979). In spontaneous speech, Gahl (2008) found that typical participants produced high-frequency homophones, such as *time*, with shorter durations than their low-frequency counterparts, such as *thyme*. This finding extends Wright's (1979) results in written word naming to a non-experimental setting and controls for the phonological form of the words under investigation. Hesitations in spontaneous speech are more likely to occur before low frequency words in the speech of typical adults (Beattie & Butterworth, 1979; Levelt, 1983).

Evidence from a variety of tasks with typical speakers converges to suggest that lexical frequency affects the production of single words, with high frequency words associated with ease of production compared to low frequency words.

Data from adults with aphasia

Newcombe, Oldfield and Wingfield (1965) confirmed that pictures of objects with high frequency names elicit faster naming responses than objects with low frequency names for adults with aphasia. Similarly, EST (Kay & Ellis, 1987) was diagnosed with severe anomic aphasia and showed more accurate word retrieval for high than low frequency object names. In group studies, high frequency targets are associated with fewer errors from speakers with aphasia (Kittredge, Dell, Verkuilen & Schwartz, 2008) and elicit fewer non-word responses in picture naming and repetition tasks (Nozari, Kittredge, Dell & Schwartz, 2010).

Tasks other than picture naming have demonstrated the effect of lexical frequency on the production of single words in speakers with aphasia. HH (Raymer et al., 1997) was diagnosed with anomia and evinced a frequency effect in written word naming accuracy. In repetition tasks, adults with aphasia responded more quickly when repeating high frequency words compared to low frequency words (Bose, van Lieshout & Square, 2007; Varley, Whiteside & Luff, 1999). In spontaneous speech, a trilingual speaker with aphasia produced longer pauses before words he rated as less frequently used across his three languages (Goral, Levy, Swann-Sternberg & Obler, 2010).

Studies that specifically examined the production of verbs have reported more mixed findings on the effects of lexical frequency. In a picture naming study of actions, Kemmerer and Tranel (2000) found that lexical frequency was not a significant factor at the group level for 19 participants who showed impaired or borderline performance in verb naming. However, in a case series analysis, four participants with aphasia experienced greater difficulties with low frequency verbs than high frequency verbs, and two participants with aphasia showed the opposite pattern, showing greater difficulty with high frequency than low frequency verbs. In a group of 54 speakers with aphasia, Bastiaanse, Wieling and Wolthuis (2016) found an effect of lexical frequency only for the naming of nouns, after controlling for other properties known to affect word production. Unlike Kemmerer and Tranel, Bastiaanse et al. did not provide a case series analysis. Therefore it remains possible that some individuals in their study may have been sensitive to lexical frequency, as Kemmerer and Tranel (2000) reported in their case series analysis.

Beyond picture naming, Breedin, Saffran and Schwartz (1998) investigated verb retrieval in aphasia by asking participants to listen to a short story and respond by producing a verb that was mentioned in the story. Like two participants reported in Kemmerer and Tranel (2000), all eight participants with aphasia showed greater success in verb retrieval for low frequency verbs.

Reverse frequency effects in aphasic language production

Like Hoffman et al.'s (2011) explanation of reverse frequency effects in aphasic comprehension, accounts of reverse frequency effects in aphasic single-word production make reference to semantic characteristics of lexemes. Breedin et al. (1998) explained the reverse lexical frequency effect they observed in verb production as an artifact of semantic complexity, where high frequency verbs like *do* can be described as 'semantically light' or 'empty' in meaning, but low frequency verbs like *dance* are associated with richer, or more complex, semantic representations. This is consistent with the conclusion of Marshall, Pring, Chiat and Robson (2001) regarding noun production in JP, who was diagnosed with jargon aphasia and demonstrated superior performance in producing, and to a lesser extent understanding, low frequency nouns, such as *collie*, than high frequency nouns, such as *dog*. The authors explained that low frequency items may have fewer semantic neighbours than high frequency items. This may result in the language processing system more readily converging on a highly distinctive, but low frequency, target for production.

Note also that the researchers who reported no or reverse frequency effects in verb production did not control for the number of argument structures with which verb targets were associated, and verbs with more complex argument structures are more difficult to name (Thompson et al., 1997; see Section 2.1.1.3).

In sum, frequency effects have been reported across a range of tasks that depend on language output processes, where typical adults and adults with aphasia produce high frequency words more quickly and accurately than low frequency words. However, research that specifically examines verb production in aphasia has returned less consistent results regarding frequency, as some adults with aphasia show superior performance in producing low frequency verbs.

2.2.1.3 Age-of-acquisition: Another usage-based effect on single-word language processing

A discussion of single-word processing in the usage-based approach could be considered remiss without mention of age-of-acquisition. Age-of-acquisition (AoA) refers to the age at which words are acquired, and it is often highly correlated with measures of lexical frequency. It remains an open question as to which factor influences typical and aphasic processing to a greater extent. One reanalysis of Oldfield and Wingfield's original dataset from typical speakers identified AoA, rather than frequency, as the main predictor of object naming latencies (Morrison, Ellis & Quinlan, 1992). AoA effects have also been found in verb processing (Columbo & Burani, 2002; Morrison, Hirsch & Duggan, 2003), as well as in a variety of tasks in addition to naming (Juhasz, 2005). Nickels and Howard (1995) showed that groups of fluent and non-fluent speakers with aphasia showed AoA effects even when lexical frequency was controlled. AoA was a significant predictor of naming accuracy at the group level, and eight of the 27 participants in the study showed an effect of AoA.

The effect of language experience can influence language processing in multiple ways. The frequency of occurrence of a single word, as well as the age at which it was acquired, both influence language processing. At the single-word level, the effect of AoA is easy to investigate, as databases provide AoA measures for single words (e.g. databases created by Bird, Franklin & Howard (2001) and Kuperman, Stadthagen-Gonzalez & Brysbaert (2012)). However, at linguistic units greater than the single word, the effect of AoA is more difficult to quantify. AoA must currently be approached as a factor of language experience that can be investigated at the level of the single word, but is less easily applied to the investigation of more complex linguistic structures, like the argument structure constructions examined in the current work.

2.2.2 Frequency effects in multi-word processing

Recent research has explored the effect of the frequency of multi-word phrases on language processing. This section discusses evidence for multi-word frequency effects in language comprehension and production.

The units in a multi-unit sequence can be referred to as *grams* in an *N-gram*, where N stands for the number of units in the sequence. Thus, in the following section, a bigram refers to a sequence of two words, a trigram refers to a sequence of three words and a quadgram to a sequence of four.

Most research in this section reports data from typical adults. To date, research that investigates the effect of frequency on language processing in aphasia has largely been directed at the single-word level, as reviewed in the previous section. There have been no published studies on multi-word frequency effects in aphasic language comprehension, for example. In the following sections, most studies refer to data from typical adults, and research on aphasia is discussed under the heading ‘data from adults with aphasia’, as in the previous sections.

2.2.2.1 Multi-word frequency effects in language comprehension

Data from typical adults

Word monitoring paradigms have been used to investigate the effect of bigram frequency on the recognition of prepositions. Vogel Sosa and MacFarlane (2002) extracted two-word phrases from a corpus of naturally occurring spoken language that contained a word followed by *of*, such as *sort of* and *piece of*. Participants' response times to identify *of* in high frequency phrases such as *sort of* were significantly longer than their response times to *of* in lower frequency phrases such as *piece of*. The authors interpreted their findings as evidence for the holistic storage of high frequency bigrams, because these longer response times indicated that participants had to decompose composite two-word phrases in memory.

Kapatsinski and Radicke (2009) employed a similar task in which they asked participants to detect the particle *up* in verb-*up* phrases of varying frequencies. Results showed a U-shaped curve in accuracy and response times: participants were increasingly faster and more accurate with increasing phrase frequency from the lowest frequency phrases, such as *eke up*, to high frequency phrases, such as *stand up*. However, for what the authors termed 'ultra-high frequency' phrases, such as *set up*, participants' accuracy fell and response times rose compared to high frequency phrases. The authors deemed these findings consistent with Vogel Sosa and MacFarlane's (2002) conclusion that very high frequency phrases are stored as whole forms.

Multi-word frequency effects have also been reported for two-word phrases containing adjective-noun combinations. Sonbul (2015) explored the effect of the frequency-based collocation strength between synonymous combinations such as *fatal mistake* (a high frequency collocation), *awful mistake* (a low frequency collocation) and *extreme mistake* (no collocation). Typical participants read phrases in a sentence context in an online task and rated the typicality of each phrase in an offline task. High frequency collocations were associated with faster first-pass reading times and higher typicality ratings. In a study of adjective-noun combinations in Italian, speakers were able to accurately judge the frequency of collocations on a four-point scale when collocations occurred with high frequency (Siyanova-Chanturia & Spina, 2015).

Jacobs, Dell, Benjamin and Bannard (2016) examined the effect of multi-word frequency of adjective-noun phrases on recognition memory. In their experiments, participants read phrases and subsequently decided whether or not they had previously witnessed them. Participants were likely to report that they had witnessed high frequency phrases, regardless of whether they actually had. The authors concluded that phrase-level frequency influences the language representation of these phrases.

In addition to two-word combinations of adjectives and nouns, two-word combinations consisting of verbs and nouns have also been studied. McDonald and Shillcock (2003) measured reading times of verb-noun bigrams that had high or low transitional probabilities. Participants read phrases in a sentence context, and the authors reported shorter fixation durations from eye-tracking on nouns in phrases with high compared to low transitional probabilities.

Siyanova-Chanturia, Conklin and van Heuven (2011) investigated three-word sequences in the form of binomial expressions, such as *bride and groom*, and their reversed forms, such as *groom and bride*. The two versions of each phrase were matched in lexical frequency and semantic content, but differed in frequency as multi-word expressions. Participants read high frequency binomial phrases in a sentence context faster than their low frequency counterparts.

High frequency four-word phrases are recognised faster and remembered better than low frequency four-word phrases. Arnon and Snider (2010) presented typical adults with four-word phrases that differed in the last word and thus in overall frequency, such as *don't have to worry* and *don't have to wait*. Pairs of phrases were matched for the frequency of the final word, bigram and trigram. Participants decided whether written phrases were acceptable in English, in the manner of a lexical decision task, and responded significantly more quickly to high frequency phrases than low frequency phrases. Tremblay and Baayen (2010) investigated four-word phrases in an immediate free recall task. Participants were shown a set of six phrases and then asked to immediately recall those phrases. High frequency phrases were accurately recalled more often than low frequency phrases. Online differences in the processing of high and low frequency phrases were evident in EEG recordings of P1 and N1 responses.

Data from adults with aphasia

No studies have directly investigated the effect of multi-word frequency in aphasic language comprehension. Huck et al. (2017), introduced in Section 2.2.1.1, explored the related effect of predictability in sentence reading. They used eye-tracking to explore reading of sentences that contained high and low frequency nouns in predictable contexts, like (12a), and unpredictable contexts, like (12b).

(12a) Claire loves flowers and wants to be a florist learning how to make nice bouquets predictable

(12b) Anna was able to get a reduced ticket for the show because she is a florist working there unpredictable

In a norming study, typical adults were given the beginning of sentences and generated potential sentence endings. They supplied predictable nouns 84% of the time and unpredictable nouns less than 1% of the time. Therefore, what Huck et al. (2017) counted as predictability is roughly equivalent to the frequency that the noun was produced in the sentence context. Participants with aphasia spent more time unpredictable nouns than predictable nouns, suggesting that probabilistic measures at linguistic units larger than the single word can affect sentence comprehension in aphasia.

In sum, the evidence reviewed in this section demonstrates an advantage in the processing of language input for high frequency combinations of words compared to low frequency combinations. Compared to typical participants, this topic has not been explored with participants with aphasia to the same extent, where the focus of research has mainly been directed at the level of single-word processing (see Section 2.2.1) or syntactic processing (see Section 2.2.3).

2.2.2.2 Multi-word frequency effects in language production

Data from typical adults

In Chapter 1, phonetic reduction in the articulation of the word *don't* in high frequency contexts was discussed as being indicative of the psychological process of chunking (Bybee & Schiebman, 1999; see Section 1.1.2.1). The relationship between frequency-based factors and phonetic reduction has been summarised as the probabilistic reduction hypothesis, which states that word forms are more reduced in contexts in which they are more predictable (Jurafsky, Bell, Gregory & Raymond, 2001). These predictable contexts can be high frequency single words, as well as probabilistic measures in two- and three-word sequences.

Gregory, Raymond, Bell, Fosler-Lusssier and Jurafsky (1999) explored the effect of probabilistic measures on the shortening of words ending with the consonants /t/ and /d/. Given a corpus of spoken productions from spontaneous speech, they examined the realisation of word-final taps, word-final consonant deletion and word duration. A significant predictor of all three shortening processes was mutual information, a measure related to bigram frequency that also accounts for the lexical frequency of each word in a phrase. For word-final consonant deletion, trigram probabilities were also significant predictors. For word duration, the probability of the word given the previous word was a significant predictor. Lexical frequency affected consonant deletion and duration, with consonant deletion more likely to occur in high frequency words, and high frequency words produced with shorter durations.

Another study by these researchers examined probabilistic factors affecting the reduction of function words (Jurafsky et al., 2001). Using the same corpus of spontaneous speech, they reported that the higher the conditional probabilities of a function word, or the word's probability given the preceding word, the following word or both the preceding and following words, the greater amount of reduction in the production of function words, in terms of duration and vowel production. These findings are consistent with the probabilistic reduction hypothesis and show that frequency-based measures in multi-word phrases affect the production of single words.

The effect of multi-word frequency on the production of phrases, rather than single words, was attested by Janssen and Barber (2012). They elicited the production of two- and three-word phrases that contained noun-adjective and noun-noun combinations from typical adults. For example, they asked participants to describe a visual array that contained an image of a chair in the colour blue by producing the phrase 'blue chair' or 'a blue chair' in two experiments conducted in Spanish and French. Speakers' naming latencies decreased with the increasing frequency of phrases. Jacobs, Dell and Bannard (2017) also explored adjective-noun phrases. The authors employed a free recall task, which they argued incorporated the process of language production, and found that high frequency phrases were more likely to be recalled in their entirety than low frequency phrases.

Arnon and Cohen Priva investigated frequency effects of three-word phrases in spontaneous speech. They reported shorter phonetic durations in the spontaneous production of high frequency word sequences, regardless of whether the sequences formed single constituents (Arnon & Cohen Priva, 2013). In a further study, Arnon and Cohen Priva (2014) extracted three-word sequences that contained a noun as the second word from a corpus of spontaneous speech and measured the phonetic duration of the noun in each phrase. In their first experiment, Arnon and Cohen Priva (2014) found that nouns in the three-word sequences had shorter durations in high frequency trigrams. In their second experiment, Arnon and Cohen Priva (2014) showed that trigram frequency affected phonetic duration even when the analysis accounted for the predictability of the surrounding context: the higher the predictability of the surrounding context, the shorter the duration of the noun in the three-word phrases. Arnon and Cohen Priva concluded that phrase frequency affects duration, in addition to predictability.

Tremblay and Tucker (2011) explored a number of probabilistic measures in the reading aloud of four-word sequences. They attempted to isolate the effects of variables that are usually highly correlated and thus not disambiguated in other work. Participants' speech onset latencies were affected by the probability of occurrence of multi-word phrases included in the four-word sequences. Probabilities of occurrence index how frequent a particular multi-word phrase is in relation to phrases that begin with the same words. For example, *at the age of* is the most frequent phrase in the collection of four-word phrases, *at the age X*, because it has the highest probability of occurrence in that set of phrases. Trigrams influenced speech onset latencies to a greater degree than single-, three- or four-word phrases. The duration of participants' spoken responses was most influenced by factors at the single-word level.

Though Tremblay and Tucker (2011) concluded that 'unigrams, bigrams, trigrams, and quadgrams all affected both recognition and production' of four-word sequences (p. 321), the frequency of four-word phrases was not a significant main effect in their statistical model. In response, Arnon and Cohen Priva (2013, Study 1) replicated the frequency effect that Arnon and Snider (2010) reported in the comprehension of four-word phrases. Using Arnon and Snider's materials, which consisted of four-word phrases that differed only in their final words and thus overall frequency, participants were presented with a written phrase for a brief period of time and then produced the phrase aloud. Phrase frequency had a significant effect on the phonetic duration of participants' productions, with shorter durations observed for high frequency phrases.

Data from adults with aphasia

Literature relating to multi-word frequency effects in aphasia is scanter than for typical adults. In an investigation of noun production, Gregory, Varley and Herbert (2012) performed a post-hoc analysis that suggested their results were consistent with an interpretation based on multi-word frequency effects. The authors aimed to investigate whether determiners that were consistent with the count or mass status of a noun primed its production. They predicted naming latencies of nouns would be facilitated by an auditory prime consisting of a congruent determiner, for a phrase like *each book*, and inhibited by an incongruent determiner, for a phrase like *some book*. They included the determiner *that* as a neutral prime. However, they found both the neutral determiner and the congruent determiner primed noun production to a similar extent, as opposed to the incongruent determiner. This finding could be explained by the frequency of co-occurrence between the determiner and noun.

Kohen, Milsark and Martin (2011) produced data on sentence production that is compatible with usage-based accounts of language. The authors conducted a study of repetition in agrammatism, where speakers with aphasia repeated sentences that contained verbs followed by particles, as in (13a), or verbs followed by prepositions, as in (13b).

(13a) The driver is turning off the lights verb-particle

(13b) The driver is turning off the road verb-preposition

The authors predicted that verb-particle phrases would be easier for speakers with aphasia to repeat, due to their simpler semantic and structural complexity compared to verb-preposition phrases. These predictions are consistent with those in the construction grammar framework described in Section 1.2. Verb-particle phrases can be recognised as unique constructions due to their unitary meaning, but verb-preposition phrases are not semantically related in this way and therefore do not constitute independent constructions. To illustrate, the verb-particle phrases in (14) are not synonymous with individual lexical verbs.

(14) Kier is standing for election Jay is giving the campaign £1000

*Kier is standing up for election *Jay is giving up the campaign £1000

As predicted, verb-particle phrases elicited more accurate repetitions from speakers with aphasia. Though this study does not directly reflect the influence of frequency, the data do exemplify the influence of usage-based properties on language production in aphasia.

2.2.3 Frequency effects in syntax

Most frequency effects in syntax have been studied as verb bias. Verb bias is a lexical property of a verb that refers to the probability that it appears in certain syntactic structures over others (Gahl & Garnsey, 2004; Lapata, Keller & Schulte im Walde, 2001). Example (15) illustrates how the verb *remember* can occur with three unique types of complements.

(15a) Sam remembered the appointment direct object

(15b) Sam remembered the appointment was at four o'clock sentence complement

(15c) Sam remembered to confirm the appointment infinitive complement

Though *remember* can occur in three argument structures, it most frequently occurs with a direct object, and therefore has a direct object bias (Trueswell, Tanenhaus & Kello, 1993). Other verbs are biased to occur in transitive or intransitive sentence structures, as shown in (16), where the verb *walk* has an intransitive bias.

(16a) Kier walked through the park intransitive

(16b) Kier walked the puppy transitive

Verb bias is usually explored by comparing conditions in which the bias of the verb is similar to the syntactic structure of the sentence in which it occurs to conditions in which the bias of the verb differs from the syntactic structure of the sentence. These conditions can be described as a ‘match’ or ‘mismatch’ between the bias of a verb and the sentence structure.

Gahl and Garnsey (2004) noted that effects of verb bias refer to structural regularities in language processing, rather than frequency effects that depend on the identity of particular words. With reference to (16), it is not the case that the verb *walk* is most often followed by the preposition *through*, but that *walk* most frequently occurs in the intransitive construction. In this way, the frequency effects described in this section differ from the multi-word frequency effects discussed in Section 2.2.2, which depended on the specific lexical items contained in the sequences.

2.2.3.1 Syntactic frequency effects in sentence comprehension

Data from typical adults

Researchers have explored the effect of verb bias on sentence comprehension in typical adults by investigating verbs that have a direct object or sentence complement bias. When these verbs are presented in sentences that contain sentence complements, the effect of verb bias is evident in readers’ responses to the post-verbal noun phrase, whose function is ambiguous between a direct object and the subject of a sentence complement, i.e. *the appointment* in (15a) and (15b). Trueswell et al. (1993) reported that readers misanalysed noun phrases following direct object-bias verbs when those noun phrases actually functioned as the subject of subject complements; however, readers did not misanalyse noun phrases following verbs with a sentence complement bias in this way. This effect was replicated by Garnsey, Pearlmutter, Myers and Lotocky (1997).

Trueswell et al. (1993) also examined sentences in which the sentence complement was optionally preceded by the complementiser *that*. In these sentence types, a sentence complement may appear directly after the verb (e.g. *Sam remembered the appointment was at four o’clock*), or after the complementiser *that* (e.g. *Sam remembered <u>that</u> the appointment was at four o’clock*). Participants’ reading times correlated with the tendency of verbs to occur with sentence complements containing a complementiser. Trueswell and Kim (1998) obtained similar results in a priming study. Prime verbs had either a direct object or sentence completion bias and were presented before the main verb in sentences that contained sentence complements. Prime verbs with a direct object bias caused greater processing difficulties when participants read sentence complements than prime verbs with a sentence complement bias.

Wilson and Garnsey (2009) noted that the studies discussed above all reported increased processing difficulty when readers encountered a sentence complement after a direct object-bias verb. They demonstrated verb bias effects in sentences where the ambiguous noun phrase is in fact a direct object. Readers displayed slower reading times of direct objects in sentences that contained verbs with a sentence complement bias.

Street and Dąbrowska (2014) investigated verb bias with respect to active and passive sentences. In their study, adults read sentences and identified the agent in each case. Participants responded to passive sentences more quickly when the sentences contained verbs that were strongly associated with the passive, rather than the active, construction. This result accords with findings on the processing of verbs with a sentence complement or direct object bias: in each case, sentence processing is eased when verbs occur in syntactic contexts in which they are most frequently used.

Syntactic frequency effects can be unrelated to verb behaviour. Reali and Christiansen (2007) investigated differences in the processing of subject and object relative clauses. In a corpus study, they identified that when an embedded noun phrase contained a personal pronoun, such as *everybody* in (17a), object relative clauses were more frequent than subject relative clauses. However, when an embedded noun phrase contained an impersonal pronoun, such as *it* in (17b), subject relative clauses were more frequent than object relative clauses.

(17a) The puppy that everybody loves chewed the sofa — object relative clause
(17b) The puppy that chewed it was scolded — subject relative clause

In four self-paced reading studies, Reali and Christiansen found that adults were sensitive to these frequency differences. Embedded personal pronouns elicited longer reading times in subject relative clauses, where they were less frequent, while the impersonal pronoun *it* elicited longer reading times in object relative clauses, where it was less frequent. Results from this study demonstrate the effects of syntactic frequency on language comprehension in a construction unrelated to verb behaviour.

Data from adults with aphasia

Verb bias has been shown to affect sentence comprehension in adults with aphasia. Gahl (2002) investigated verbs whose bias matched or mismatched the sentence structure of active, passive and intransitive sentences. At the group level, five typical adults and 17 adults with aphasia responded significantly more accurately in a semantic plausibility judgement task when verb biases matched the sentence structure. Gahl et al. (2003) employed a similar task and found that a group of eight adults with aphasia showed an effect of verb bias for four sentence structures, including active and passive sentences and intransitive sentences with unergative and unaccusative verbs.

More recently, DeDe (2013a) explored the effect of verb transitivity bias, as shown in (15), on reading times in aphasia. Of ten participants with aphasia, nine showed longer reading times when there was a mismatch between verb bias and the intransitive sentence structure, and six showed longer reading times when there was a mismatch between verb bias and the transitive sentence structure. Participants with aphasia showed a greater effect of verb bias than typical participants in this study. Adults with aphasia also showed effects of verb bias in more structurally complex sentences that contained sentence complements (DeDe, 2013b).

2.2.3.2 Syntactic frequency effects in sentence production

Data from typical adults

Verb bias affects sentence production, in addition to sentence comprehension. Gahl and Garnsey (2004) investigated pronunciation variability in sentences that contained a verb with a direct object or sentence complement bias. They measured the rate of deletion of verb-final consonant /t/ or /d/, as well as the duration of verbs and the noun phrases that were ambiguous in function between direct objects and subjects of the sentence complements. Parallel to the findings on verb bias in sentence comprehension, they found significantly higher rates of final-consonant deletion and shorter word durations in contexts where the verb bias matched the sentence structure.

In addition, probabilistic measures associated with whole constructions affect spoken language production. The dative alternation is often considered to consist of two semantically equivalent forms, as was demonstrated in (11) as the ditransitive *she threw him something* and the caused motion *she threw something to him*. In fact, a speaker's choice between the two constructions is related to a variety of syntactic and semantic factors, such as the accessibility of the referents, definiteness, animacy and concreteness (Bresnan, Cueni, Nikitina & Baayen, 2007). Word duration also depends on the likelihood a particular syntactic construction. Based on a corpus of spontaneous speech, word durations were found to be shorter in more highly probable constructions (Kuperman & Bresnan, 2012; Tily et al., 2009).

Like Trueswell et al. (1993), researchers have measured speakers' inclusion of the complementiser *that* in sentences which contain verbs with direct object and sentence complement biases. They proposed that speakers may include the complementiser more often in sentences that contain verbs with a direct object bias, because the mismatch between verb bias and sentence structure could prove more difficult for a speaker to produce and a listener to comprehend. Two studies reported results consistent with this reasoning. In a corpus study on spontaneous spoken production, Jaeger (2010) reported that verbs which were more likely to occur with a sentence complement were less likely to occur with the complementiser *that*. In an experimental study, Ferreira and Schotter (2013) found that speakers were more likely to produce *that* following verbs with a direct object bias.

Data from adults with aphasia

Menn, Gahl, Holland, Ramsberger and Jurafsky (2003) extended Gahl et al.'s (2003) findings on verb bias in semantic plausibility judgements to repetition. SK was diagnosed with Broca's aphasia and repeated 98 sentences containing verbs in syntactic contexts that matched or mismatched their biases. Though she found it difficult to repeat sentences in their entirety, SK produced more complete responses when verb biases matched the syntactic context of the sentence. When she did produce errors, SK tended to alter the sentence syntax to match the verb bias.

DiLallo et al. (2017) examined the effect of verb bias in language production, based on data available from the *AphasiaBank* corpus. They considered a set of 22 verbs that had a transitive or intransitive verb bias and reviewed data from typical participants and participants with aphasia. Both groups of participants produced verbs in line with verb biases, in that they most often produced transitively-biased verbs in the transitive construction and intransitively-biased verbs in the intransitive construction. There was no significant difference between the two groups, indicating that participants with aphasia demonstrated the effect of verb bias to the same extent as typical participants. There was no significant difference between speakers with Broca's and Wernicke's aphasia. Participants with aphasia also generated significantly more production errors in sentence contexts where verbs did not match their verb bias.

2.3 The present research

This research explores the usage-based effect of frequency on the processing of verbs and argument structure. The work interrogates Diessel's (2015) lexical links in the network architecture of grammar, described in Section 1.2.3 as the association between individual words and syntactic constructions. This research was carried out in the construction grammar framework, and syntactic contexts were defined as argument structure constructions with unique meanings.

The present work examines two types of frequency: lexical frequency, which refers to the frequency of a verb as a single word, and construction frequency, which refers to the frequency of a verb in a particular argument structure construction. In line with the evidence discussed in Section 2.2, both lexical frequency and construction frequency were predicted to affect language processing, with processing advantages associated with high frequency items.

This research included both typical adults and adults with acquired aphasia as participants. As mentioned above, contextual frequency measures such as multi-word phrases have garnered less attention in investigations of atypical language processing, but findings from typical adults suggest such effects may play a substantial role in language processing at levels greater than the single word.

The research comprises two phases. In Phase 1, participants took part in a verbal fluency task. In Phase 2, a different set of participants took part in a computerised grammaticality judgement and a sentence completion task. The aims of each phase are detailed in the following chapters. Results of this work are anticipated to be of interest in psycholinguists and aphasiologists alike. The general discussion in Chapter 9 comments on findings with reference to their clinical implications for aphasia.

*

This chapter provided an overview of how acquired aphasia affects language processing and how frequency of occurrence influences language processing in typical adults and adults with aphasia at a variety of linguistic grain sizes. The thesis continues with a report of Phase 1.

3 Phase 1: Verbal fluency task

3.1 Introduction

The study reported in this chapter comprised Phase 1 of the project. Phase 1 explored speakers' knowledge of the association between verbs and argument structure constructions. Participants in Phase 1 included typical adults and adults with acquired aphasia. Results from this study contributed to the design of Phase 2.

Phase 1 builds on findings from Ellis, Brook O'Donnell and Römer (2014), who investigated the association between verbs and argument structure constructions. Ellis et al. asked typical adult participants to name verbs that they thought could occur in versions of the intransitive motion construction, which was presented in the form *pronoun* ___ *preposition determiner*. In the first experiment, participants supplied a single lexical item in response to each construction. In the second experiment, participants were given sixty seconds to supply verbs in response to each stimulus, in a computerised verbal fluency task. The authors analysed the effect of three usage properties of the verbs produced in response to each stimulus on the number of times participants generated the verbs in each task. These usage properties included the frequency of verbs in each construction; the faithfulness of verbs to each construction, compared to the number of other constructions that verbs could occur in; and the semantic prototypicality of verbs in each construction, as measured by their centrality in a semantic network containing all the verb types that could occur in a construction (Ellis & O'Donnell, 2012). Multiple regression analyses demonstrated the independent and significant effect of each of these three properties.

The present study explores the effects of two usage-based factors of verbs: lexical frequency and construction frequency. Lexical frequency refers to how often a single word is produced as a verb in British English. Construction frequency refers to the frequency with which verbs were produced in particular argument structure constructions in British English. Ellis et al. (2014) reported significant correlations between both these factors and the number of times participants generated a verb in response to a construction in their two experiments. Ellis et al. explored only the intransitive motion construction in typical speakers. Phase 1 broadens that research by including eight unique argument structure constructions, and including participants with acquired aphasia.

Both Ellis et al. (2014) and the current study employed a verbal fluency task. In verbal fluency tasks, participants generate members of a category within a given timeframe. This paradigm was considered appropriate for use with the populations of interest in Phase 1, because it has been successfully employed in the research context with older typical adults (e.g. Kahlaoui et al., 2012; Marsolais, Methqal & Joanette, 2015) and participants with aphasia (e.g. Arroyo-Anlló, Lorber, Rigaleau & Gil, 2011; Baldo, Schwartz, Wilkins & Dronkers, 2010).

Researchers have identified a number of characteristics that influence participants' responses to verbal fluency tasks, which researchers should be mindful of when employing this type of task. First, participants produce fewer items over the course of the time given for naming responses to a category, and responses tend to be produced in semantically-related clusters (Gruenewald & Lockhead, 1980). Second, words with high lexical frequency are named more often than words with low lexical frequency (Henley, 1969). Additionally, prototypical category members are produced before typical and non-typical category members (Kail & Nippold, 1984). Frequency and typicality are related, as researchers have taken items produced with high frequency in a verbal fluency task to represent the most typical instances of the category (Chang, 1986).

The aims of Phase 1 were:

- to examine the relationship between the number of times typical participants and participants with acquired aphasia generated verbs in response to particular argument structure constructions, and the verbs' lexical and construction frequency;
- to explore differences and similarities between the performance of typical adults and adults with acquired aphasia; and
- to identify whether the syntactic form of a construction alone, devoid of lexical semantic content, was sufficient to elicit verbs from participants with aphasia.

To summarise findings from Phase 1, the number of times typical participants produced verbs in the verbal fluency task was significantly correlated to verbs' construction frequency and lexical frequency, and for most constructions the strength of the relationship to construction frequency was stronger than to lexical frequency. Participants with aphasia successfully produced verbs in the task, and the verbs they produced multiple times in the task were higher in lexical frequency and construction frequency than verbs they produced only once in the task.

3.2 Method

3.2.1 Participants

Twenty typical adults and four adults with acquired aphasia took part in Phase 1. Typical participants were native speakers of British English between the ages of fifty and eighty years with no reported history of speech or language difficulties or psychiatric impairment. The group included ten men and ten women. All reported normal or corrected-to-normal vision and hearing. Table 3.1 shows characteristics of the group.

Table 3.1 Background information for typical participants in Phase 1

	M	*SD*	Minimum	Maximum
Age in years	63	6.4	52	74
Years in education	16	3.6	11	23

Note. Years in full-time education includes primary, secondary and higher education.

Table 3.2 Background information on participants with aphasia in Phase 1

Participant	Age	Gender	Handedness	Years in education	Profession
DS	70	M	R	14	Administrator
EF	61	F	R	13	Administrator
LM	73	M	A	15	Manual
TP	70	M	L	10	Professional

Note. Age in years; gender, male (M) or female (F); self-reported premorbid handedness, right (R), left (L) or ambidextrous (A); years in full-time education, including primary, secondary and higher education; profession.

Table 3.3 Characteristics of aphasia for participants in Phase 1

Participant	Time post-onset	Aetiology	Previous speech and language therapy
DS	10;0	CVA	One year of therapy following stroke
EF	2;9	CVA	Three months of therapy following hospital discharge
LM	0;11	CVA	Weekly sessions for one year following stroke
TP	10;2	CVA	Minimal therapy

Note. Time post-onset shown in years;months format.

Four speakers with acquired aphasia also took part in Phase 1. These adults were native speakers of British English who had received a diagnosis of acquired aphasia from a speech and language therapist and had been living with aphasia for at least six months prior to taking part in the study. There was no restriction on the location or type of brain damage that induced aphasia, though all participants had acquired aphasia as a result of a single CVA. Individuals with no residual expressive language were excluded from the study, in order to ensure that participants with aphasia retained sufficient ability to take part in the experimental task of Phase 1. Participants with aphasia reported no other developmental or acquired speech or language difficulties in addition to aphasia, and no history of psychiatric impairment. All participants with aphasia had normal or corrected-to-normal vision and hearing. At the time of participating in Phase 1, all participants were attending a client-led communication group for adults with aphasia at the University of Sheffield. None were receiving speech and language therapy provided by a health service.

Characteristics of the participants with aphasia are listed in Table 3.2. Details related to the clinical condition of participants in this group are provided in Table 3.3.

3.2.2 Ethics, recruitment and consent

The study received ethical approval from the University Research Ethics Committee of the Department of Human Communication Sciences at the University of Sheffield. Typical participants were recruited via printed posters and emails sent to departments within the University of Sheffield and local organisations for retired persons. Participants with aphasia were recruited from a communication group for adults with aphasia held at the University of Sheffield.

Written informed consent was obtained from all participants prior to their taking part in the study. Written materials for participants with aphasia, including an information sheet and consent form, were designed following the Stroke Association's accessible information guidelines (Herbert, Haw, Brown, Gregory & Brumfitt, 2012). In addition, an information sheet was created for family members and carers of participants with aphasia. If requested by the participants with aphasia or their family members or carers, a participant's family member or carer was present throughout the consent process. All participants were given a copy of the information sheet, containing the researcher's contact details, and a signed copy of the consent form. Project approval, information sheets and consent forms used for this study are provided in Appendix A.

3.2.3 Design

All participants took part in four language tasks in Phase 1. In the verbal fluency task, participants named verbs that could occur in eight unique argument structure constructions. This task was designed to investigate participants' knowledge of the relationship between verbs and argument structure constructions.

All participants took part in three additional language assessments: action-naming, function word processing and grammaticality judgement. Participants with aphasia completed these assessments so that their ability to generate verb forms and process written language could be compared to their performance in the novel verbal fluency task. Typical participants completed the tasks in order to provide normal data for the assessments. The assessments were administered after the verbal fluency task in order to avoid priming of verbs, which were responses of interest in the verbal fluency task. Full profiling of aphasia was not administered to ensure that testing could be completed within a single assessment session.

Each assessment session began with the verbal fluency task, then a short break if requested, then three related language assessments. All participants performed the action-naming task, followed by the function word processing task and finally the grammaticality judgement task.

3.2.4 Materials

3.2.4.1 Verbal fluency task

Stimuli for the verbal fluency task consisted of sixteen sentences, including two instances of eight unique argument structure constructions from Goldberg's (1995; Johnson & Goldberg, 2013) framework. In each sentence, a blank space stood in place of the verb. The sentence stimuli were composed entirely of function words, including pronouns and prepositions. In this way, the meaning of each sentence was generated uniquely by its syntactic form, rather than any semantic activation arising from the lexical items the sentences contained. Stimuli sentences therefore encoded the event-level meaning of argument structure constructions. Sentences began with the subject pronouns *I*, *you*, *we* and *they* because these pronouns elicit uninflected verb forms and can refer to a referent of either gender. The subject pronoun *it* was included in intransitive motion and passive constructions to allow participants the opportunity to supply responses limited to inanimate subjects.

Table 3.4 lists the constructions, their syntactic forms, meanings and corresponding sentence stimuli used in the verbal fluency task.

Table 3.4 Sentence stimuli for verb fluency task in Phase 1

Construction	Syntactic form	Meaning	Stimulus sentences
Caused motion	Subject Verb Object Oblique	X causes Y to move to Z	You ___ it to us I ___ it over there
Conative	Subject Verb Oblique_{at}	X directs motion at Y	You ___ at us They ___ at it
Ditransitive	Subject Verb Object1 Object2	X causes Y to receive Z	They ___ us some things I ___ you something
Intransitive motion	Subject Verb Oblique_{path}	X moves to/from Y	We ___ through there It ___ through there
Passive	Subject aux VPpp PP_{by}	X is acted on by Y	We were ___ by them It was ___ by them
Removal	Subject Verb Object Oblique_{source}	X causes Y to move from Z	I ___ it from you You ___ it from there
Transitive	Subject Verb Object	X acts on Y	We ___ them You ___ it
Way	Subject Verb <way> Oblique_{path}	X moves through path	They ___ their way to it We ___ our way there

Note. Alternating rows in grey for ease of reference.

Two lists were created. The lists differed in the order of presentation of the sentence stimuli. The order of the two versions of each argument structure construction was counterbalanced across the lists. Items were pseudo-randomly ordered with the constraint that sentence stimuli which instantiated the same construction were separated by at least three other sentences. Lists were assigned randomly to participants with the proviso that half the participants responded to one list and half responded to the other.

The same five example items and five practice items were included at the start of each list. The sentences used for these items were also composed of function words but did not take the form of any of the argument structure constructions under investigation. Missing verbs followed infinitival *to* or modal auxiliary verbs. The full set of stimuli for the verbal fluency task is provided in Appendix B. Three low-frequency verbs were supplied for each example item.

Each sentence stimulus was presented in simultaneous written and auditory forms. Written sentences were presented individually in black, bold, point-42 Helvetica type centred on A5-sized paper. The auditory forms of sentences were recorded by a female native speaker of British English on a Marantz PMD670, recording at stereo PCM (pulse code modulation) with a sample frequency of 22.1 KHz via a Sennheiser MD425 microphone. During the recording of the auditory sentence stimuli for the verbal fluency task, the speaker creating the recordings was asked to silently rehearse a verb in the gap in the sentences. In order to ensure the gaps in all stimuli sentences were acoustically similar, audio files were edited on the software Audacity: 500 milliseconds of white noise with an amplitude of 0.01 was inserted to represent the missing verb in the auditory stimuli.

3.2.4.2 Action-naming task

An action-naming task assessed participants' ability to produce verbs in isolation, i.e. without a sentence context. Participants were asked to name 20 black-and-white line drawings depicting lexical verbs. Items were taken from Druks and Masterson's (2000) *An Object and Action Naming Battery*.

Druks and Masterson (2000) classified the verb targets in their battery as transitive, unergative or unaccusative. Transitive verbs occur with a subject and an object, in a sentence like *the puppy licked the child*. Unergative verbs occur with a subject that corresponds to an agent, in a sentence like *the grocer sneezed*. Unaccusative verbs occur with a subject that corresponds to a theme, in a sentence like *the tap is dripping*. No ditransitive verbs were included in the battery due to the difficulty of representing such items pictorially. The numbers of each type of verb target included in the action-naming task, and the mean lexical frequency, naming agreement, age-of-acquisition, imageability and visual complexity ratings of each group are shown in Table 3.5. The full set of stimuli included in this task is provided in Appendix B.

Items were selected in order to maximise the range of verbs that formed the assessment. Thus, half the items included on the assessment occurred with an object, and half did not, and unaccusative verbs – despite being low in frequency – were also included.

Table 3.5 Number and mean ratings of verb targets included in action-naming task in Phase 1

	Transitive	Unergative	Unaccusative
Number of targets	8	8	4
Lexical frequency	150	164	41
Name agreement	99.4%	99.1%	97.5%
Age-of-acquisition	2.38	1.98	2.72
Imageability	4.05	4.44	4.01
Visual complexity	4.09	4.16	3.53

Note. Number of targets of each type; mean frequency per one million words; mean naming agreement in terms of percent of participants who agreed; mean age-of-acquisition, imageability and visual complexity rated on scale of 1-7. Ratings reported in Druks and Masterson (2000).

Items were pseudo-randomly ordered following two constraints. No more than two transitive, unergative or unaccusative verbs occurred in succession, and adjacent targets did not begin with the same phoneme. All participants saw the same list order.

3.2.4.3 Function word processing task

A novel assessment of function word processing was included to evaluate participants' ability to process written function words. Function words were used in the sentence stimuli in the verbal fluency task, and accessing the meaning of the sentences depended on function word processing. The assessment took the form of a lexical decision task. All the function words that appeared more than once in the set of sentence stimuli in the verbal fluency task were included in this task. Materials consisted of 14 real words and 14 non-words which were derived by changing one letter in each real word, following Herbert, Anderson, Best & Gregory (2014). Words were displayed as a list, centred on a sheet of A4 paper in lower case, black, bold, Helvetica type in point 22. The task began with five practice items, including two real words and three non-words.

The list was pseudo-randomly ordered so that no more than three words or non-words appeared consecutively, and non-words were not adjacent to the real words from which they were derived. All participants saw the same list order. The full set of stimuli included in this task is provided in Appendix B.

3.2.4.4 Grammaticality judgement task

A grammaticality judgement task was included to assess whether participants recognised each argument structure construction included in the verbal fluency task as an acceptable grammatical structure of English. Each argument structure construction included in the verbal fluency task corresponded to a grammatical and an ungrammatical sentence in the grammaticality judgement task, for a total of 16 items. Five practice items consisted of verbs following infinitival *to* and modal auxiliary verbs. The verbs in this task all occurred over 100 times per million (Leech et al., 2001), could not be used as auxiliary verbs and were monosyllabic or disyllabic. Verbs were presented in the past tense. Sentences that instantiated the same construction contained different words, and ungrammatical sentences were formed by changing the order of words in the verb phrase. The full set of stimuli included in this task is provided in Appendix B.

Each sentence was presented in auditory and written forms, following the creation of materials described above for sentences in the verbal fluency task (see Section 3.2.4.1). Items were pseudo-randomly ordered with the constraints that no more than three grammatical or ungrammatical sentences appeared in succession, and no adjacent sentences were based on the same argument structure construction. All participants saw the same list order.

3.2.5 Procedure

Participants met with the researcher individually in a quiet room.

3.2.5.1 Verbal fluency task

The verbal fluency task began with up to five example items and five practice items, as required by individual participants. Participants were informed that trials would end when a timer sounded and the researcher said ‘please stop’. Following the practice items, participants were presented with the sixteen experimental items, in one of two list orders.

On each trial, participants were presented with the written form of the sentence. At the same time, the recording of the sentence was played from a laptop via Dell AX210 USB speakers. Participants could listen to the recording as many times as they wished before responding, and the written sentence remained in view of participants throughout the response time. The researcher prompted participants to respond on each trial by saying ‘please say as many single words as you can think of that can fit in the sentence’. Timing began after the researcher finished this prompt and the participant indicated she was ready to proceed. Each trial ended after 30 seconds, which H. Kim, J. Kim, D. Kim and Heo (2011) suggested was an adequate response time for speakers with aphasia to respond to verbal fluency tasks.

3.2.5.2 Action-naming task

Each image was presented on a single sheet of A4 paper. Participants were asked to describe the pictures using a single word. Following Druks and Masterson (2000), the researcher asked participants ‘can you tell me what is happening here?’ or ‘can you tell me what he/she/it is doing?’ as each picture was presented. If participants named a non-target part of the picture, the question was repeated. If participants responded with a superordinate response, they were asked to be more specific. If a correct non-target response was produced, they were asked to provide an alternative.

3.2.5.3 Function word processing task

Participants read each function word silently and answered ‘yes’ or ‘no’ in response to the researcher’s prompt ‘is this a real word in English?’.

3.2.5.4 Grammaticality judgement task

Participants were presented simultaneously with the auditory and written forms of each complete sentence. The recording of the sentence was played from a laptop via Dell AX210 USB speakers. Written sentences remained in participants’ view until they responded, and participants could listen to the recording was often as they requested. They responded ‘yes’ or ‘no’ to the prompt ‘is this sentence okay in English?’.

3.2.5.5 Recording of data

The researcher transcribed participants’ spoken responses in situ and recorded any prompts given to participants. Responses were transcribed orthographically in the order in which they were produced. Non-words were transcribed in broad phonemic transcription. The entire assessment session was audio recorded on a Marantz PMD670, recording at stereo PCM (pulse code modulation) with a sample frequency of 44.1 KHz via a Sennheiser MD425 microphone. The researcher subsequently checked all written responses against the audio recordings to ensure that all data were accurately transcribed. The reliability of transcriptions is described in Section 3.2.5.7.

3.2.5.6 Coding of responses

All participant responses in the verbal fluency task were coded. The coding scheme is shown in Table 3.6. Responses scored as real verbs, phonological errors and rejections were entered into analyses.

All verbs produced in the verbal fluency task were scored for their grammaticality. For example, the production of the verb *visit* in response to the conative sentence stimulus *you ___ at us* was scored as ungrammatical.

In the action-naming task, participants’ final responses to items were scored following the procedure adopted in Herbert, Webster and Dyson (2012) for single-word picture naming. If any prompts were given, participants’ response following the prompt was scored. Because this test was administered to measure lexical retrieval abilities and not morphosyntactic processing, any inflection of the target verb was accepted as correct. In addition, the production of arguments with the target response was also accepted as correct. The response codes for the action-naming task are shown in Table 3.7.

Table 3.6 Response codes for Phase 1 verbal fluency task

Response code	Description	Example
Real verb*	Response matched any form of a real English verb	saw, given
Phonological error*	Response shared at least 50% of the phonemes of any form of a real English verb	/təʊl/ for *told*
Non-verb	Response was a real word of English that was not a verb	near, okay
Multiple words	Response contained multiple words to complete the stimulus sentence	you must be mad at us
Repetition	Response was previously produced on the same trial	find (real verb), found (repetition)
Rejection*	Response was produced but immediately rejected by participant	you played at - no
Non-word	Response was not a real word of English	/bʊʔ/
No response	Participants made no attempt to respond to the stimulus sentence, verbally acknowledged that they could not respond or were only able to read or repeat the stimulus sentence	no can't, I don't know

Note. Asterisks indicate verbs included in analysis. Alternating rows in grey for ease of reference.

Table 3.7 Response codes for Phase 1 action-naming task

Response code	Description	Example
Correct	Response contained any form of target verb	building, building bricks
Semantic error	Response bore a thematic relationship to target, e.g. belonged to same schema or involved same items as target	*lacing* in response to target *tying*
Phonological error	Response contained at least 50% of the phonemes of any form of the target verb; non-words included	*dricking* in response to target *dripping*
Unrelated word	Response was a single real word that bore no semantic relationship to the target	*licking* in response to target *touching*
Other	Responses contained narratives or gestures, or were visual errors or names of non-target parts of the picture	*taking his dog for a walk* in response to target *pulling*

Note. Alternating rows in grey for ease of reference.

3.2.5.7 Inter- and intra-rater reliability

Transcriptions and coding of the data from the verbal fluency task were subject to inter- and intra-rater reliability testing. Transcription agreement is reported in terms of percent agreement, and agreement for the coding of responses is quantified using Cohen's kappa (Cohen, 1960). This statistic denotes the degree of agreement between independent raters on nominal classifications, correcting for agreement that would be expected by chance (Hallgren, 2012). Kappa values range from -1 to 1: a value of 0.60 indicates moderate agreement, a value of 0.80 indicates substantial agreement and a value of 1.00 indicates perfect agreement (Landis & Koch, 1977; cited in Viera & Garrett, 2005). Calculations were performed in SPSS.

Two randomly selected trials from each typical participant were subject to testing. This equated to a total of 40 trials, or 12.5% of the trials in the dataset. A different subset of trials was subject to inter- and intra-rater testing. Inter-rater reliability judgements were performed by a native speaker of British English with a university degree in English language. Transcription agreement was 95%: transcriptions were the same for 270 of 283 cases. The coding was the same for 269 of these 270 items; $\kappa = 0.92$. Intra-rater transcriptions and coding judgements were performed eight weeks following the original. Transcription agreement was 99%: transcriptions were the same for 280 of 282 cases. There was 100% agreement on the coding of these 280 items; $\kappa = 1.00$.

Four randomly selected trials from each participant with aphasia were subject to inter- and intra-rater reliability testing. This equated to a total of 16 trials, or 25% of the trials in the dataset. A different subset of trials was subject to inter- and intra-rater testing. Inter-rater reliability judgements were performed by a native speaker of British English with a university degree in linguistics and experience transcribing the speech of adults with aphasia. Transcription agreement was 87%: transcriptions were the same for 40 of 46 cases. The coding was the same for 38 of these 40 items; $\kappa = 0.92$. Intra-rater transcriptions and coding judgments were performed twelve weeks following the original. Agreement on the transcription was 88%: transcriptions were the same for 46 of 52 cases. There was complete agreement on the coding of these 46 items; $\kappa = 1.00$.

An independent native speaker of British English confirmed the researcher's decisions on the grammaticality of verbs produced in response to each sentence stimulus in the verbal fluency task.

3.2.5.8 Frequency values for verbs produced in verbal fluency task

For each verb that participants produced in the verbal fluency task, two measures of frequency were derived. Both were based on the 100-million-word British National Corpus, which contains spoken and written texts primarily from the 1980s. The corpus is balanced in that it was constructed to accurately represent contemporary British English in terms of the types of texts it contains (Aston & Burnard, 1998).

Lexical frequency refers to how often an item's lemma occurred as a verb in the entire British National Corpus, in instances per million. Lemma frequency refers to how often all forms of a verb type occur in a corpus, e.g. the lemma frequency of the verb *dance* is 37 instances per million, accounting for the frequency of the forms *dance* (14 instances per million), *dancing* (13 instances per million), *danced* (8 instances per million) and *dances* (2 instances per million). Values were taken from the lists provided in Leech et al. (2001).

Construction frequency refers to how often a verb occurred in a particular argument structure construction. These values were derived from Brigham Young University's interface to the British National Corpus (Davies, 2004-). Each construction was defined as a query using the UCREL tagset (UCREL, 1993). By specifying the verb lemma in the query as a verb from the dataset, the raw frequency of each verb in a particular construction was extracted. Corpus results were manually filtered to ensure that results matched the target construction. The method for extracting construction frequency values is explained in more detail in Appendix C.

3.2.6 Data analysis

Data from typical adults were examined as correlations. Data from adults with aphasia were analysed with non-parametric tests. Mann-Whitney tests were used to compare two unrelated groups, and the Kruskal-Wallis test was used as the non-parametric equivalent to an independent one-way ANOVA. In keeping with standard methodology in research on frequency effects (e.g. Ellis et al., 2014), frequency values were log transformed and incremented by 0.01 before statistical tests were performed on data from participants with aphasia. Effect sizes for statistical tests are reported as r, following recommendations in Field (2009).

3.3 Results

This section reports results of Phase 1 from typical participants and participants with aphasia. Section 3.3.1 provides a summary of responses from both groups of participants. Section 3.3.2 contains analyses of data from typical participants in the verbal fluency task. Section 3.3.3 contains analyses of data from participants with aphasia, including results from the three language assessments in Section 3.3.3.4 .

3.3.1 Participant responses

Table 3.8 summarises the total number of verb types and responses that participants produced in the verbal fluency task. Verb types refer to unique verb lemmas, e.g. *look*, *love* and *play* are distinct verb types. Responses refer to the total number of replies participants produced in the entire verbal fluency task. The number of verb types is distinguished from

Table 3.8 Participant responses in Phase 1 verbal fluency task

	Typical participants				DS	EF	LM	TP
	M	*SD*	Min	Max				
Number of verb types	63	16	33	93	**19**	**12**	**10**	**11**
Number of responses	106	27	58	151	**25**	**30**	**12**	**31**
Proportion of responses which were grammatical	1.00	0.01	0.99	1.00	**0.84**	**0.70**	1.00	**0.61**

Note. Scores from participants with aphasia outside normal range in bold.

the number of responses, because participants produced the same verb in response to different constructions, or different versions of the same construction. Table 3.8 also shows the proportion of responses that were grammatical. Table 3.8 includes only responses coded as real verbs, phonological errors and rejections. A summary of participants' responses in terms of the coding scheme described in Section 3.2.5.6 is provided in Appendix D.

Typical participants were able to produce a variety of verb types in the verbal fluency task. Overall, these verbs were grammatical in the argument structure constructions to which they were produced. Compared to the group of typical participants, participants with aphasia performed outside the normal range in terms of the number of verb types and number of responses they produced in the verbal fluency task. DS, EF and TP also fell outside the normal range in terms of the grammaticality of their responses, but LM performed within the normal range on this factor.

3.3.2 Results from typical participants

3.3.2.1 Distributions of number of times of verbs generated in response to constructions in verbal fluency task

Figure 3.1 (over page) shows the distributions of the number of times all typical participants generated a verb in response to each of the eight argument structure constructions in the verbal fluency task. The maximum number of times a verb could be produced in the verbal fluency task per construction was 40, which represents all 20 participants generating the verb in response to each of the two stimulus sentences.

Each panel in Figure 3.1 depicts all verb types that typical participants produced in response to an argument structure construction. The y-axis shows the number of times the verb was produced, with a maximum of 40. Each bar on the x-axis represents a single verb type. The number of verb types produced in response to each construction is shown as *n* in each panel. Because of the great number of verb types, only some are labelled on the x-axis. For example, typical participants generated a total of 105 verb types in response to the caused

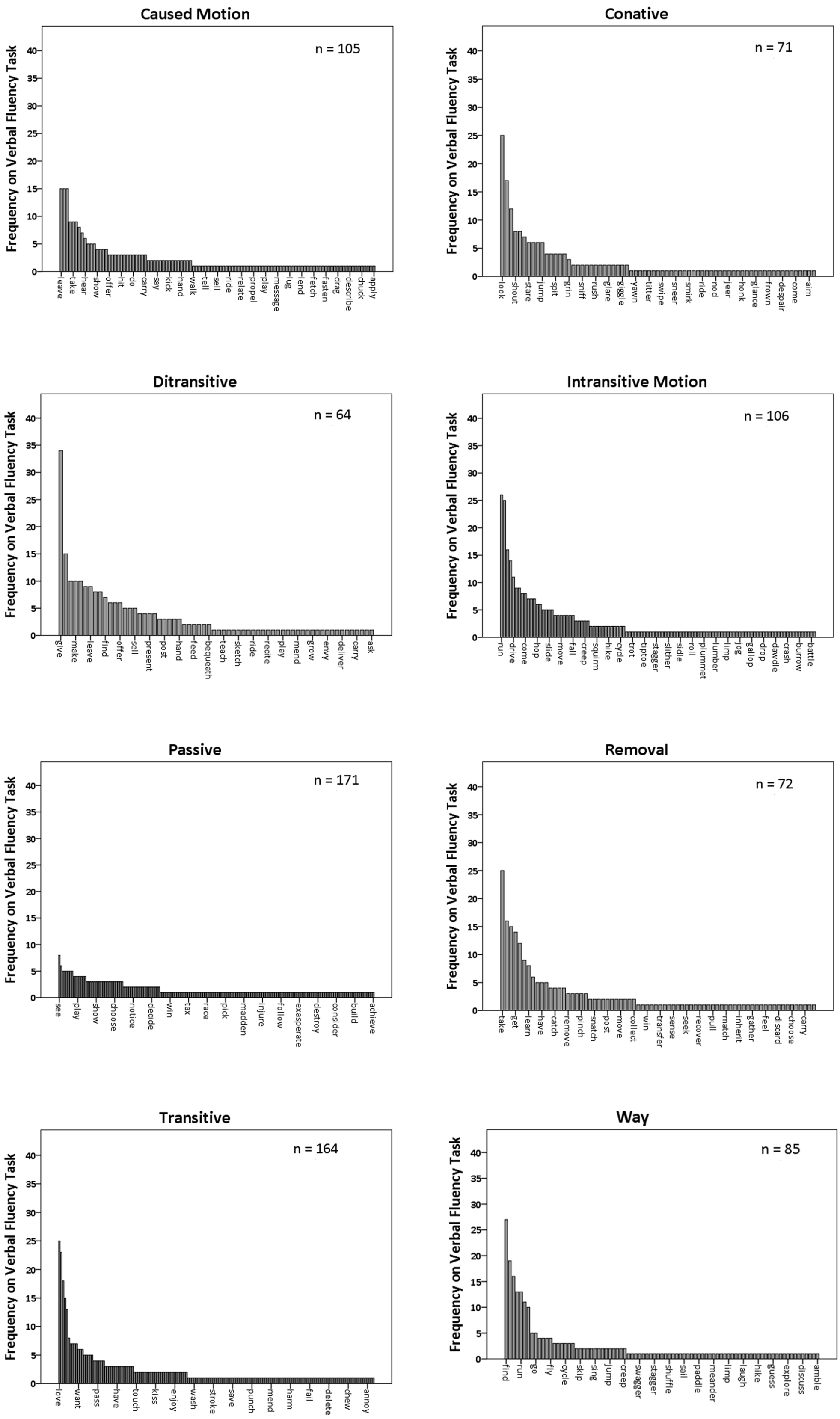


Figure 3.1 Distributions of number of times typical participants generated verbs in response to each argument structure construction in Phase 1 verbal fluency task

motion construction. The verb *leave* was produced the most number of times in the task, a total of 15.

3.3.2.2 Correlations for verbs produced in verbal fluency task

Correlations between the number of times typical participants generated a verb in the verbal fluency task and verbs' lexical frequency and construction frequency are reported in Table 3.9. Data are collapsed across participants and across the two versions of each construction included in the verbal fluency task. For example, the group of typical participants produced a total of 105 verb types in response to the caused motion construction, to either stimulus sentence *you ___ it to us* or *I ___ it over there.* Spearman's rho correlations are shown in Table 3.9.

Table 3.9 Data from typical participants: Correlations between the number of times verbs were produced in response to a construction in verbal fluency task, and verbs' lexical frequency and construction frequency

	n	Lexical Frequency	Construction Frequency	Significant difference between lexical frequency and construction frequency
Caused motion	105	0.46***	0.42***	
Conative	71	0.29**	0.49***	
Ditransitive	64	0.32**	0.62***	•
Intransitive motion	106	0.29***	0.51***	•
Passive	171	0.40***	0.29***	
Removal	72	0.23*	0.57***	•
Transitive	164	0.39***	0.54***	•
Way	85	0.28**	0.34***	

Note. *$p < 0.05$. **$p < 0.01$. ***$p < 0.001$. All one-tailed.

For all eight argument structure constructions the strength of the correlation between the number of times participants produced verbs in the verbal fluency task and verbs' lexical frequency values fell between 0.23 and 0.46. These correlations were all significant ($0.001 < p < 0.05$). For all eight constructions, the correlation between the number of times participants produced verbs and verbs' construction frequency values fell between 0.29 and 0.62 (p-values < 0.001).

For six of the eight constructions, the relationship to construction frequency was stronger than to lexical frequency: only the caused motion and passive constructions did not follow this pattern. These differences reached significance for four of these six constructions: for the ditransitive, intransitive motion, removal and transitive constructions, the correlation between the number of times participants produced verbs in the verbal fluency task and verbs' construction frequency was significantly greater than verbs' lexical frequency.

The significance of the difference between correlations of lexical frequency and construction frequency was calculated following Baguley's (2012) implementation of Zou's (2007) method. Zou (2007) set out a method for determining the significance of the difference between overlapping dependent correlations, where two unique correlations share a common variable. His approach depends on the construction of confidence intervals for differences between correlations. Significance can be determined when the confidence intervals do not contain zero. The approach can be used for data with non-normal distributions (Zou, 2007) and Spearman's rho correlation coefficients (Zou, personal communication).

The caused motion and passive constructions were the only constructions to show a stronger correlation between the number of times participants generated verbs in the verbal fluency task and verbs' lexical rather than construction frequency. In order to further analyse this finding, correlations were calculated separately for each of the two versions of the stimulus sentences included in the verbal fluency task. Findings are shown in Table 3.10.

Table 3.10 Correlations between number of times verbs generated in verbal fluency task and verbs' lexical frequency and construction frequency for each version of the caused motion and passive constructions

	n	Lexical Frequency	Construction Frequency	Significant difference between lexical and construction frequency
Caused motion				
You ___ it to us	47	0.35**	0.50***	
I ___ it over there	71	0.54***	0.28**	•
Passive				
It was ___ by them	95	0.48***	0.37***	
We were ____ by them	97	0.17	0.15	

Note. *$p < 0.05$. **$p < 0.01$. ***$p < 0.001$. All one-tailed.

One version of the caused motion construction, the sentence stimulus ending with the preposition phrase *to us*, showed a stronger correlation for construction frequency than lexical frequency. The sentence stimulus containing the preposition phrase *over there* showed the opposite pattern: the correlation between the number of times participants generated verbs in the verbal fluency task was significantly stronger to verbs' lexical frequency than verbs' construction frequency.

For the passive sentence stimulus containing the subject pronoun *it*, the correlations between the number of times participants generated verbs in the verbal fluency task and verbs' lexical and construction frequency were both significant (p-values < 0.001). The correlation for lexical frequency was stronger than for construction frequency. For the passive sentence stimulus containing the subject pronoun *we*, neither correlation for lexical or construction frequency was significant.

3.3.3 Results from participants with aphasia

Analyses could not be performed on data from participants with aphasia in the same manner as data from typical participants, because participants with aphasia produced fewer verbs than typical participants in the verbal fluency task. The maximum number of times the group of participants with aphasia produced a verb type in response to a construction was three, therefore, it was not possible to calculate correlations within each construction.

Section 3.3.3.1 presents the verbs that each participant with aphasia produced in the verbal fluency task in response to each construction. In Section 3.3.3.2, frequency values of verbs that the group of participants with aphasia produced only once were compared to those of verbs produced more than once. Data were collapsed across participants and constructions. Analyses were carried out on a dataset containing all verbs that participants with aphasia produced, and again on a dataset containing only verbs coded as grammatical. In Section 3.3.3.3, the performance of each participant with aphasia was compared to the group of typical participants. This analysis allowed the identification of similarities and differences between individuals with aphasia and typical participants, in line with the second aim of the study. This analysis was also carried out first on a dataset containing all verbs and then on a dataset containing only verbs coded as grammatical. Section 3.3.3.4 presents participants' performance on the three additional language assessments, and Section 3.3.3.5 considers the effect of aphasia severity, as indexed by participants' performance on the language assessments, on the frequency of verbs produced in the task.

3.3.3.1 Lexical and construction frequency values of verbs produced by participants with aphasia

The verb types that participants with aphasia produced in response to the eight argument structure constructions in the verbal fluency task are shown in Table 3.11. For each verb, the number of times it was produced by the group of typical participants is shown (maximum 40), along with its lexical and construction frequency values. Verbs are listed in order of lexical frequency.

Table 3.11 Verbs produced by participants with aphasia (DS, EF, LM, TP) in verbal fluency task

	DS				EF			
	Verb	Typ Num	Lex Freq	Cx Freq	Verb	Typ Num	Lex Freq	Cx Freq
Caused motion	see	15	1920	11	have*	1	13655	14
	bring	5	439	932	give	15	1284	1230
	send	9	250	747	find*	4	990	6
	fly	0	90	35	want	3	945	2
	email	2	0	0				
Conative	see*	0	1920	0	see*	0	1920	0
	visit*	0	115	0	look	25	1151	7851
					play*	2	386	8
					shout	8	59	319
					wink	1	4	64
Ditransitive	tell	15	775	113	find	7	990	53
	bring	8	439	107				
	fetch	0	19	16				
Intransitive motion					be	8	42277	179
					find*	1	990	3
					want*	0	945	1
Passive	see	8	1920	285				
	play	4	386	188				
	stand	0	326	0				
	seat	0	11	1				
	sew	0	8	0				
Removal	buy	12	262	113	be*	0	42277	38
	receive	9	247	21	take	25	1797	434
	fetch	0	19	20	find*	5	990	5
					pick	2	150	3
Transitive	work	0	646	60	have	3	13655	1594
	play	2	386	168	see	13	1920	5683
	buy	3	262	592	give	5	1284	184
	sleep*	0	68	1	find	3	990	1568
	complain*	0	44	0	want	6	945	1237
	telephone	0	23	40	pick	1	150	44
Way	journey	0	2	0	look*	0	1151	0
					find	27	990	1038

Note. Verb types judged as ungrammatical marked with an asterisks. Values include number of times verbs produced by group of typical participants (Typ Num), lexical frequency (Lex Freq) and construction frequency (Cx Freq). Alternating rows in grey for ease of reference.

Table 3.11 (continued) Verbs produced by participants with aphasia (DS, EF, LM, TP) in verbal fluency task

	LM				TP			
	Verb	Typ Num	Lex Freq	Cx Freq	Verb	Typ Num	Lex Freq	Cx Freq
Caused motion	give	15	1284	1230	have	1	13655	14
	present	2	143	175	say	2	3344	81
					put	5	700	260
					sell	1	213	314
					kick	2	36	9
Conative	be	4	42277	185	say*	0	3344	1
					sell*	0	213	0
Ditransitive	tell	15	775	113	say*	0	3344	0
	owe	10	37	19				
Intransitive motion	be	8	42277	179	have*	0	13655	0
					say*	0	3344	9
					put*	0	700	5
					cut	2	184	380
Passive	stand	0	326	0	say	1	3344	154
					take	5	1797	688
					run	2	406	529
					sell	3	213	105
Removal	take	25	1797	434	have	5	13655	36
	steal	15	48	22	say*	0	3344	1
					tell*	0	775	10
					sell	1	213	2
					lay*	0	104	0
Transitive	see	13	1920	5683	have	3	13655	1594
					say	4	3344	806
					see	13	1920	5683
Way	find	27	990	1038	say*	1	3344	0
					sell*	0	213	0

Note. Verb types judged as ungrammatical marked with an asterisks. Values include number of times verbs produced by group of typical participants (Typ Num), lexical frequency (Lex Freq) and construction frequency (Cx Freq). Alternating rows in grey for ease of reference.

3.3.3.2 Analysis of frequency values of verbs produced by participants with aphasia

This analysis explored the lexical frequency and construction frequency values of verbs produced only once by the four participants with aphasia, and verbs produced more than once by the four participants. This analysis was designed to parallel the correlations calculated on data from typical participants. Since participants with aphasia did not often produce the same verb type multiple times in response to a construction, it was not possible to calculate correlations in same the manner as for data from participants without aphasia. Data are collapsed across participants and constructions.

The median lexical and construction values for these groups of verbs are shown in Table 3.12. Median values are presented due to the use of non-parametric statistical tests, which were selected because of the relatively small number of verbs included in the analysis, compared to the number of verbs included in the analysis of data from typical participants.

Table 3.12 Median lexical and construction frequency values of verbs produced once and verbs produced more than once by participants with aphasia in verbal fluency task

	Verbs produced once	Verbs produced more than once
All verbs		
n	58	18
Lexical frequency	324	1859
Construction frequency	22	146
Grammatical verbs only		
n	41	12
Lexical frequency	250	1217.5
Construction frequency	64	922

Lexical frequency and construction frequency values were log transformed and incremented by 0.01, and statistical tests were computed on the transformed values. Values were incremented in order to avoid loss of data after the log transformation, following Ellis et al. (2014), because the log of 1 is 0.

Comparisons were first carried out on a dataset containing all verbs that participants with aphasia produced in the task. Verbs produced more than once by the four participants with aphasia were significantly higher in lexical frequency than verbs produced only once (U = 215.00, z = -3.76, $p < 0.001$, r = -0.43). Verbs produced more than once were higher in construction frequency than verbs produced only once, and this difference was approaching significance (U = 381.00, z = -1.73, p = 0.084, r = -0.20).

A second set of comparisons was carried out on a dataset containing only the responses from participants that were coded as grammatical. Consistent with the above, verbs produced more than once by participants with aphasia were significantly higher in lexical frequency than verbs produced only once (U = 103.50, z = -3.03, p = 0.002, r = -0.42). In this dataset, the difference in construction frequency between the groups of verbs reached significance. Grammatical verbs that were produced more than once by the group of participants with aphasia were significantly higher in construction frequency than grammatical verbs produced only once (U = 101.00, z = -3.08, p = 0.002, r = -0.42).

3.3.3.3 Comparison of frequency values of verbs produced by participants with aphasia to verbs produced by typical participants

This analysis compared the lexical and construction frequency values of verbs produced by each participant with aphasia to the lexical and construction frequency values of verbs produced by all typical participants. In this analysis, frequency values were weighted by the number of times participants generated verbs in response to a construction in the verbal fluency task. To illustrate, the group of typical participants produced the verb *give* 34 times

Table 3.13 Median lexical and construction frequency values of verbs produced by typical participants and participants with aphasia in verbal fluency task

	Typical participants	DS	EF	LM	TP
All verbs					
n	2116	24	30	12	31
Lexical frequency	169	250	990	1137	3344
Construction frequency	63	21	121.5	182	10
Grammatical verbs only					
n	2116	21	20	12	18
Lexical frequency	169	262	990	1137	1248.5
Construction frequency	63	40	736	182	207

Note. Frequency values weighted by number of time verbs produced in verbal fluency task.

in response to two stimulus sentences, *they ___ us some things* and *I ___ you something*, the ditransitive construction. In the analysis, the frequency values for the verb *give* were weighted by 34. The dataset therefore contained each instance a verb was produced in the task, rather than each verb type.

This weighting was adopted because typical participants produced many verb types in the dataset only once. Figure 3.1 shows that about half of the verb types are in the tails of the distributions for each construction. If a weighting had not been employed, the construction frequency and lexical frequency of verbs that participants produced only once would influence the analysis to the same extent as verbs that participants produced many times. The weighting ensured that the values of lexical frequency and construction frequency included in the analysis represented the behaviour of typical participants.

Median values of the lexical and construction frequency of verbs produced by typical participants and each participant with aphasia are shown in Table 3.13. Data are collapsed across constructions.

Lexical frequency and construction frequency values were log transformed and incremented by 0.01. Statistical analysis of the transformed data compared the complete set of data from typical participants, collapsed across participants and constructions, to each set of data from the four participants with aphasia.

A Kruskal-Wallis test on the dataset containing all verbs showed a significant difference in lexical frequency between groups ($H(4) = 69.19$, $p < 0.001$). Mann-Whitney tests were used to compare the lexical frequency of verbs produced by each participant with aphasia to the lexical frequency of verbs produced by the group of typical participants. EF, LM and TP produced verbs that were significantly higher in lexical frequency than verbs produced by typical participants (for EF: $U = 13334.00$, $z = -5.46$, $p < 0.001$, $r = -0.12$) (for LM: $U = 6904.00$, $z = -2.73$, $p = 0.003$, $r = -0.06$) (for TP: $U = 12817.50$, $z = -5.83$, $p < 0.001$, $r = -$

0.13). There was no significant difference between the lexical frequency of the verbs produced by DS and typical participants (U = 25784.00, z = -0.22, p = 0.828, r = -0.004).

A Kruskal-Wallis test showed a significant difference in construction frequency between groups (H(4) = 11.43, p = 0.022). Mann-Whitney tests were used to compare the construction frequency of verbs produced by each participant with aphasia to the construction frequency of verbs produced by the group of typical participants. DS produced verbs that were significantly lower in construction frequency than verbs produced by typical participants (U = 20284.00, z = -2.01, p = 0.045, r = -0.04). EF and LM produced verbs higher in construction frequency than typical participants, but this difference did not reach significance for EF (U = 28281.50, z = -1.03, p = 0.305, r = -0.02) and was approaching significance for LM (U = 9180.00, z = -1.66, p = 0.097, r = -0.04). TP produced verbs that were lower in construction frequency than verbs produced by typical participants, and this difference was approaching significance (U = 26335.00, z = -1.89, p = 0.059, r = -0.04).

The analyses described above were performed on the dataset containing only the responses from participants that were coded as grammatical. The two datasets showed the same differences between the lexical frequency of verbs produced by participants with aphasia and typical participants. The two datasets differed in one comparison of the construction frequency of verbs produced by a participant with aphasia and typical participants. EF produced grammatical verbs that were significantly higher in construction frequency than verbs produced by typical participants (U = 11490.00, z = -3.52, p < 0.001, r = 0.08).

3.3.3.4 Performance on language assessments

In addition to the verbal fluency task, all participants took part in three background language assessments, including an action-naming task, function word processing task and grammaticality judgement task. The performance of the group of typical participants and each participant with aphasia is shown in Table 3.14.

Table 3.14 Performance on language assessments in Phase 1

	Typical participants					DS	EF	LM	TP
	N	*M*	*SD*	Min	Max				
Action-naming	20	0.995	0.02	0.95	1.00	0.95	**0.90**	**0.60**	**0.80**
Function word processing	28	0.97	0.05	0.82	1.00	1.00	1.00	0.89	0.96
Grammaticality judgement	16	1.00	0.00	1.00	1.00	1.00	1.00	1.00	**0.81**

Note. Number of items in each task (*N*); scores from participants with aphasia outside normal range in bold. Alternating rows in grey for ease of reference.

Participants with aphasia are listed in order from least impaired (DS) to most impaired (TP), based on their performance in these tasks.

In the action-naming task, eighteen of the twenty typical participants produced no errors. Participants with aphasia did not achieve the level of performance demonstrated by the group of typical participants. DS made a single semantic error, producing the response *lacing* to the target *tying*, whose image was a person tying shoes. EF's two errors consisted of naming a non-target part of an image, and a visual error. LM's errors in the action-naming task were phonological and visual errors. All his phonological errors were within one phoneme of the target; for example, he produced the response *polding* for the target *folding*. His remaining errors involved difficulty interpreting the visual stimuli; for example, he described the target *touching* as people receiving an electric shock. TP produced two semantic errors, an unrelated error that was a perseveration of a previously named verb, and the name of a non-target part of the stimulus. A breakdown of participants' non-target responses in terms of the coding scheme described in Section 3.2.5.6 is provided in Appendix D.

All typical participants performed at ceiling in the function word processing task, and all participants with aphasia performed within the normal range.

In the grammaticality judgement task, the score from one typical participant was discarded; due to technological failure, the audio recordings were not played. None of the nineteen typical participants made any errors in the grammaticality judgement task. TP was the only participant with aphasia to perform below ceiling in this task. He made three errors in which he judged ungrammatical sentences to be grammatical. However, a sign test showed that TP performed above chance in the grammaticality judgement task ($p = 0.01$).

Taken together, the results from the three language assessments demonstrate that participants with aphasia retained sufficient language function to meet the demands of the novel verbal fluency task. Though participants with aphasia performed more poorly than typical participants in the action-naming task, most errors were due to difficulty with the visual stimuli. Such images were not present in the verbal fluency task. EF's phonological errors were accounted for by the coding scheme used for the verbal fluency task, where such responses were entered into the data analysis. In the function word processing task, all participants with aphasia performed within the normal range. In the grammaticality judgement task, all participants with aphasia were able to correctly identify the argument structure constructions contained in the verbal fluency task as grammatical sentences.

3.3.3.5 Effect of severity of aphasia

This final analysis examined whether participants' severity of aphasia had an effect on the lexical frequency or construction frequency of the verbs they produced. This analysis investigated whether participants with more severe aphasia produced verbs that were higher in frequency than participants with milder forms of aphasia. To accomplish this aim, a Jonckheere-Terpstra trend test was adapted for use with the transformed data from participants with aphasia discussed in Section 3.3.3.3. The independent variable was defined as participants with aphasia, in order of severity based on their performance in the language assessments discussed in Section 3.3.3.4. The dependent variable was the lexical frequency and construction frequency of the verbs they produced.

When all verbs produced by participants with aphasia were included in the dataset, a significant trend was identified for lexical frequency values ($J = 2354$, $z = 3.92$, $p < 0.001$, $r = 0.40$). There was no significant trend for construction frequency ($J = 1760.5$, $z = 0.11$, $p = 0.457$, $r = 0.01$). The same pattern of findings emerged when only grammatical verbs were included in the analysis. Individuals with more severe cases of acquired aphasia produced verbs with higher lexical frequency, but not construction frequency, than verbs produced by individuals with milder cases of aphasia.

3.4 Discussion of Phase 1

In Phase 1, typical participants and participants with acquired aphasia completed a verbal fluency task in which they named verbs in response to eight unique argument structure constructions. The correlation between the number of times typical participants generated verbs in the task was significantly related to verbs' lexical and construction frequency values. For most constructions, the relationship to construction frequency was stronger than to lexical frequency. The verbs that participants with aphasia produced more than once in the task were higher in lexical and construction frequency than verbs they produced only once in the task. After a brief consideration of the power of the statistical analyses in Phase 1, the findings from each participant group will be discussed in turn in this section.

3.4.1 Power of Phase 1 analyses

Statistical power can be defined technically as 'the probability of rejecting a false null hypothesis' (Wilson VanVoorhis & Morgan, 2007, p. 43). That is, it is the power of a statistical test to correctly identify a real effect. Statistical power is related to the alpha-level, effect size, sample size and distribution of a statistical test, and statistical power of 80% is conventionally taken to be adequate (Cohen, 1992). A power analysis allows researchers to evaluate whether the sample size of their study is sufficient for their selected statistical test to achieve this level of power. If a study is under-powered, their statistical analysis risks failing to identify a true effect.

In Phase 1, correlational analyses were used to explore data from typical participants in terms of the number of responses they produced in the verbal fluency task. VanVoorhis and Morgan (2007) recommended sample sizes of around 50 as reasonable to test for relationships using correlations with sufficient power. The minimum sample size of a correlation in Phase 1, where correlations tested the strength of the relationship between the number of times verbs were produced in the task and verbs' frequency measures, was 47 (see Table 3.10), and the maximum was 171 (see Table 3.9). Therefore, it is likely that Phase 1 included sufficient items in correlational analyses to avoid missing a real effect in the analysis of data from typical participants.

Differences in data from participants with aphasia were investigated using non-parametric Mann-Whitney tests. Shieh, Jan and Randles (2006) reported sample sizes between 20 and 28 for the Mann-Whitney test to achieve 90% power in detecting large differences in unequal group sizes. In Phase 1, some group sizes were as low as 12, indicating that the analysis was likely vulnerable to low power, especially in the cases of the analysis of grammatical verbs only and comparisons of LM and TP to typical participants. Given these circumstances, the direction of the difference in the dataset can be taken as the key finding from this analysis, because non-significant results from the statistical tests may be a result of their low power.

Apart from the power analyses of the statistical tests employed in Phase 1, the study derives some power - in a conceptual sense - from its description of the data under investigation, rather than statistical significance testing (Kern, 2013). The similarity of the distributions of data from typical participants across argument structure constructions, shown in Figure 3.1, indicates that the study successfully identified a consistent pattern in the dataset. The responses from participants with aphasia in Table 3.11 may be the best way to explicate the clinical data from Phase 1, given the relatively small number of verbs they produced as a result of their language impairments.

3.4.2 Construction frequency in typical participants

Ellis et al. (2014) reported a significant effect of both lexical frequency and construction frequency in their investigation of the intransitive motion construction. In their data, correlations between the number of times verbs were produced in response to a construction and verbs' construction frequency were generally stronger than verbs' lexical frequency. The results from typical participants in Phase 1 replicate the findings of Ellis et al. for the intransitive motion construction: the correlation between the number of times verbs were produced in response to the intransitive motion construction and verbs' construction frequency values was significantly stronger than the correlation between the number of times verbs were produced and their lexical frequency values.

Phase 1 demonstrated that the same pattern holds for five additional argument structure constructions, including the conative, ditransitive, removal, transitive and way constructions. The significant effect of lexical frequency on verb production may be attributed to the nature of the verbal fluency task, where the lexemes most often named in the task are high in lexical frequency (Henley, 1969). The significant effect of construction frequency indicates the strength of the association between verbs and argument structure constructions, or what Diessel (2015) termed lexical links in the network architecture of grammar. This finding demonstrates that knowledge of the context in which a verb is used makes up a crucial part of a speaker's linguistic knowledge.

Two constructions, the caused motion and the passive, differ from the others in that the relationship to construction frequency was not stronger than the relationship to lexical frequency. This following sections consider reasons why this may be so.

Caused motion

The identity of the preposition included in the stimulus sentence appeared to have influenced the verbs that typical participants produced in response to the caused motion construction. Table 3.10 showed that the caused motion stimulus sentence containing the preposition phrase *to us* showed a stronger correlation between the number of times participants generated verbs in the verbal fluency task and verbs' construction frequency than lexical frequency. In contrast, the sentence stimulus containing the preposition phrase *over there* showed a significantly stronger correlation for lexical than construction frequency. These findings suggest that the two stimulus sentences included in the verbal fluency task as instances of the caused motion construction may in fact have been two distinct constructions.

This conclusion raises the issue of how abstract argument structure constructions may be. In their study, Ellis et al. (2014) identified argument structure constructions as frames that contained a pronoun, verb and preposition. The twenty constructions they included differed in these prepositions; that is, the authors treated the stimuli sentences 'she ___ across the…' and 'she ___ towards the…' as unique constructions. This differs from Goldberg's (1995) approach, where she describes argument structure constructions at the phrase level. For example, Goldberg (1995) argued that both sentences *Sam helped him into the car* and *Sam* *guided him through the terrain* (p. 162) are instances of the caused motion construction. Goldberg's (2003) claim that 'it's constructions all the way down' (p. 223) goes some way to encompassing these two approaches, by recognising that prepositions, in addition to the phrase-level structure of an entire sentence, qualify as constructions and contribute to the meaning of a sentence as a whole. In this way, one could conceptualise an argument structure construction as an umbrella term for a number of related, lexically-specified constructions of the same syntactic form. However, since the identity of the preposition in such a family of sentences appears to influence speakers' responses to the construction, it remains to be seen whether argument structure constructions - conceived in this abstract way - are relevant to human language processing, or are purely an artefact of a linguist's endeavour to describe language in its entirety. Results from the current study suggest that typical speakers are influenced by lexical attributes of argument structure constructions.

Passive

Two factors may explain why construction frequency was not more strongly related to the number of times participants generated verbs in response to the passive construction than lexical frequency. First, the passive does not correspond to an argument structure construction that encodes event-level meaning in the same way as, for example, the conative or removal constructions refer to specific events. Rather, Goldberg (2003) described the passive as a discourse-level construction that allows a theme, instead of an agent as in non-passive sentences, to function as the topic of an utterance. Second, the passive can occur with a wider range of verbs than other argument structure constructions. The passive elicited more verb types than any other construction in Phase 1 (n = 171; see Table 3.9). Virtually any verb whose semantic representation contains an agent and patient can occur in the passive construction. The combination of these factors - a discourse-level construction that can occur with a great variety of verbs types - could explain why construction frequency did not have a greater effect than lexical frequency on typical participants in Phase 1 for the passive construction.

Additionally, the number of times participants generated verbs in response to the passive construction was significantly related to verbs' frequency measures for the sentence stimulus containing the subject pronoun *it*, but not *we*. This finding echoes those from Bock's (1986) study, where passive production in the structural priming paradigm was associated with non-human agents. That semantic properties of the subject influence sentence structure suggests a close relationship between the sentence subject and verb. Kamide, Altmann & Haywood (2003, Experiment 2) observed the effects of this relationship in an eye-tracking study using a visual world paradigm. Their participants looked more often at a picture of a motorbike upon hearing the sentence *the man will ride* than the sentence *the girl will ride* and more often at a picture of a carousel upon hearing the sentence *the girl will ride* than the sentence *the man will ride*. These results suggest that probabilistic measures associated with the subject affect sentence processing, in addition to verbs and argument structure. This effect may have been caused the observed differences in participants' responses to the two versions of the passive construction in Phase 1.

3.4.3 Participants with aphasia in Phase 1

The four participants with aphasia in Phase 1 were all able to produce verbs in response to argument structure constructions, and the majority of their verb responses were grammatical in the stimulus sentence. In response to the third aim of the study, findings demonstrated that the syntactic form of argument structure constructions, in the shape of stimulus sentences composed entirely of function words and devoid of lexical semantic content, was sufficient to elicit verbs from speakers with aphasia. This section will discuss findings from participants with aphasia.

3.4.3.1 Individual participants with aphasia

This section explores the performance of the four participants with aphasia, DS, EF, LM and TP, in more detail.

DS

DS was the least severely impaired participant with aphasia. He was the only participant with aphasia to score within the normal range in all three language assessments, including the action-naming, function word processing and grammaticality judgement tasks. He produced the greatest number of verb types in the verbal fluency task compared to other participants with aphasia, and the majority of his responses were grammatical. DS was able to produce verbs in response to seven of the eight argument structure constructions in Phase 1: he was only unable to produce verbs in response to the intransitive motion construction, though overall he produced fewer verb types in the task compared to typical participants.

DS performed similarly to typical participants in several ways. First, the lexical frequency and construction frequency of the verbs he produced in the verbal fluency task did not differ significantly from the frequency of verbs produced by typical participants. Second, many of DS's verb responses were similar to those produced by typical participants, as indexed by the number of times typical participants produced verb types in response to each construction (see Table 3.11). For example, DS produced the verb *buy* in response to the removal construction, and typical participants produced this verb 12 times in the task, making it one of the five most frequently supplied verbs in response to the removal construction. DS produced verbs that were often produced by typical participants in response to the caused motion, ditransitive and passive constructions, as well. However, DS was unable to produce an grammatical verbs in response to the conative construction, and two of his responses to the transitive construction were ungrammatical.

EF

EF performed within the normal range in the function word processing and grammaticality judgement tasks. The two errors she produced in the action-naming task were visual errors. These results suggest that EF's language impairment would not likely have impeded her ability to process materials included in the verbal fluency task.

EF produced grammatical verbs in response to seven of the eight argument structure constructions; she was unable to produce any verbs in response to the passive construction. In this respect, EF was similar to other individuals with aphasia who showed particular difficulty processing the passive construction (Faroqi-Shah & Thompson, 2003; Friederici & Graetz, 1987). When only her grammatical responses were considered, EF produced verbs that were significantly higher in construction frequency than verbs produced by typical participants. Like DS, EF often named verbs that were frequently produced by typical participants. These verbs were also high in construction frequency, such as *give* in the caused motion construction and *look* in the conative construction. EF produced similar verbs in response to the removal, transitive and way constructions.

LM

LM performed within the normal range in the function word processing and grammaticality judgement tasks. Most of his errors in the action-naming task were phonological paraphasias, and these errors differed from the target by only one phoneme. LM was the only participant with aphasia who exclusively produced grammatical verbs in response to each of the eight argument structure constructions. Like DS and EF, many of LM's verb productions were often generated by typical participants (see Table 3.11). LM appears to present with a language impairment that uniquely affects phonological output processing, but leaves intact semantic and grammatical processing.

The grammatical competence of speakers with aphasia who have similar language profiles to LM were explored in two previous studies. DK (Friedmann & Gvion, 2007) was diagnosed with output conduction aphasia, an impairment that results in the production of phonological errors in language output (Gvion & Friedmann, 2012). DK showed equivalent performance to a group of typical participants on the comprehension of garden path sentences, which contained complex syntactic phenomena such as embedding. EA

(Friedrich, Martin & Kemper, 1985) presented with spontaneous speech that was grammatical, but marked by phonemic paraphasias. The grammaticality of EA's spontaneous sentence productions was at ceiling. The case of LM in Phase 1, and of DK and EA in previous research, suggest that individuals with phonological coding deficits can have intact language processing at the verb and sentence level: a specific deficit in phoneme selection or production does not necessarily affect grammatical processing.

TP

TP was the most severely impaired participant with aphasia in Phase 1. He scored outside the normal range in the action-naming and grammaticality judgement tasks. His performance in the function word processing task suggests, however, he was able to process the words contained in the stimulus sentences in the verbal fluency task.

TP produced verbs with the highest lexical frequency compared to other participants with aphasia, and his verb productions were significantly higher in lexical frequency than verbs produced by typical participants (see Section 3.3.3.3). In this way, TP's performance corresponds to Huck et al.'s (2017) finding that participants with more severe aphasic impairments showed a greater effect of word frequency in sentence reading than participants with less severe impairments, as measured by overall performance and lexical-semantic processing abilities. These findings suggest that words with high lexical frequency remain more accessible than words with low lexical frequency in the linguistic systems of individuals with moderate aphasia severity.

TP appeared to show little knowledge of the constructions in which verbs can occur. Only 39% of his responses in the verbal fluency task were judged to be grammatical (see Table 3.8). His performance in this task was characterised by the repetition of the same verb types in response to different constructions. He produced the verb *say* in response to all eight argument structure constructions in the task, the verb *sell* in response to five of the constructions and the verb *have* in response to four of the constructions. Compared to other participants with aphasia, fewer of TP's verb productions were often produced by typical participants. TP produced the verb *say* in response to the transitive construction, which was produced 13 times by typical participants; the verbs *put* and *take* were each produced five times by typical participants in response to the caused motion and passive constructions (see Table 3.11). With these exceptions, TP's verb productions were rarely attested in data from typical participants. This is consistent with the finding that TP's verbs were lower in construction frequency than verbs produced by typical participants (see Section 3.3.3.3).

TP's performance in the verbal fluency task suggests that the associations between verbs and argument structure constructions can be weakened in individuals with moderate aphasia, though this claim warrants substantiation by a more detailed investigation of language in aphasia than was undertaken in Phase 1. Together with the effect of lexical frequency on TP's verb productions, as discussed above, the following conclusion can be tentatively postulated, subject to further evidence: individuals with more severe cases of aphasia primarily access word-level constructions rather than larger linguistic units, whereas individuals with milder forms of aphasia and typical adults have access to constructions greater than the single word.

3.4.3.2 Multi-word frequency effects in aphasia

Section 2.2.2 observed that there is a lack of research on multi-word frequency effects in aphasia. Findings from this study begin to address that gap in knowledge. DS, EF and LM all produced some verbs in the verbal fluency task that were high in construction frequency and often produced by typical participants. Their responses demonstrate the availability of verbs that are closely associated with particular argument structure constructions. These constructions were defined by certain lexical items, and so the effect of construction frequency demonstrated by these three participants with aphasia may arise from the frequency of the association between verbs and other lexical items. For example, DS, EF and LM all produced verbs with high construction frequency in response to the removal construction, which contains a preposition phrase headed by *from*. This preposition may have served as a successful cue to verbs with high construction frequency due to the strength of the collocation between those verbs and *from* in language use. In a similar way, the other constructions were also defined by particular lexical items, and those lexical items may have cued the production of verbs with high construction frequency. These results indicate that the frequency of multi-word n-grams can influence language processing in aphasia.

*

This chapter reported the verbal fluency task of Phase 1, where the construction frequency of verbs in argument structure constructions had a significant effect on language processing in typical participants and, to some extent, in participants with aphasia. The effects of lexical frequency and construction frequency will be further explored in the sentence processing tasks of Phase 2.

4 Phase 2 method: Sentence processing tasks

4.1 Introduction

The current chapter describes the method for Phase 2, which employed a grammaticality judgement task and a sentence completion task to further investigate the effects of lexical frequency and construction frequency.

4.1.1 Differences between Phase 1 and Phase 2 tasks

In the Phase 1 verbal fluency task, participants were given time to reflect on the relationship between verbs and argument structure constructions. In the tasks employed in Phase 2, participants were supplied with lexical verbs, to which they responded following brief exposure to argument structure constructions. The differences between tasks in Phase 1 and Phase 2 arose from two important motivations.

First, supplying verbs to participants in Phase 2 capitalises on the residual language abilities of participants with aphasia. Participants in Phase 1 were required to generate verbs entirely independently in the verbal fluency task, without any cues about verbs that could occur in the argument structure constructions under consideration. Individuals with aphasia can produce linguistic structures in a constrained context, such as picture naming, that they are not able to employ in a less constrained context, such as spontaneous speech or conversation (Boyle & Coelho, 1995; Carragher, Conroy, Sage & Wilkinson, 2012; Conroy, Sage & Lambon Ralph, 2009; Pashek, 1998). Therefore, the independent generation of lexical items in Phase 1 may have functioned to make the verbal fluency task more difficult for participants with aphasia than other types of language assessment.

Second, the tasks in Phase 2 were designed to allow less time for participants to contemplate the relationship between verbs and argument structure constructions than they had in Phase 1. In Phase 2, participants were exposed to argument structure constructions for a few seconds on each trial, rather than the 30 seconds they had to respond to sentences in Phase 1. In this way, the tasks employed in Phase 2 reduced participants' explicit consideration of argument structure constructions.

4.1.2 Preparation of data from Phase 1

The dataset of verbs that participants generated in Phase 1 served as the basis for the materials employed in Phase 2. This section describes the rationale and procedure for preparing raw data from Phase 1 for use in Phase 2.

The relationship between the form and function of constructions is one-to-many: the same syntactic form can serve multiple purposes in language. For example, the conative construction refers to an event in which an agent directs action towards another entity, in a sentence such as *the children waved at their teacher*. The entity at which action is directed is encoded as a preposition phrase headed by *at*. However, *at* can also denote a location, as in the sentence *the children ate at the picnic table*. This sentence has the same syntactic form as the conative construction, i.e. *subject* - *verb* - *oblique*$_{\mathrm{at}}$, but it is not an instance of the construction: no action is directed towards another entity. Rather, the preposition phrase marks the location of the action.

Because noun phrases were represented as pronouns in the sentence stimuli used in the verbal fluency task, participants in Phase 1 sometimes produced verbs that resulted in sentences that were not instances of the target constructions. That is, the conative sentence stimulus in Phase 1, *they* ___ *at it*, elicited both verbs *wave* and *eat*. These verbs resulted in sentences that were, respectively, consistent and inconsistent with the semantics of the conative construction. In order to ensure that only verbs consistent with the semantics of the constructions were considered for inclusion in materials for Phase 2, verbs were judged as to whether they 'fit' the argument structure constructions under consideration.

The author and two other independent raters, who were both native speakers of British English and held postgraduate qualifications in linguistics, provided ratings for verbs in the dataset from Phase 1. Raters were asked to judge whether each verb could be used in the sentence stimulus in response to which it was produced, and whether the resulting sentence was an instance of the target construction. The author and one other rater provided judgements for the entire dataset. Disagreements were arbitrated by the third independent rater. Only verbs that were judged as a good fit with the construction by two of the three raters were considered for inclusion in the materials for Phase 2.

This process had the secondary benefit of excluding verbs that were so unusual in certain constructions that participants in Phase 2 might interpret them as ungrammatical. For example, *you cycle at us* could be used as a legitimate instance of the conative construction to describe a situation in which a cyclist travels towards a group of people. However, *cycle* was not judged as a good fit with the conative construction by two of the three raters. Therefore, the rating process ensured that the verbs associated with constructions in Phase 2 all resulted in acceptably grammatical sentences.

4.1.3 Phase 2 research aims and hypotheses

Phase 2 explored the degree to which specific properties of language usage affect language processing. Specifically, the aims of Phase 2 were:

- to examine the effects of lexical frequency and construction frequency on participants' responses to verbs, given prior exposure to argument structure constructions;
- to investigate any differences and similarities in performance between typical participants and participants with acquired aphasia; and
- to analyse the effect of different patterns of residual language ability on the performance of participants with acquired aphasia.

Given results from Phase 1 and previous research, verbs high in lexical frequency and construction frequency were expected to be processed with greater ease by typical participants than verbs with low frequency, as evidenced by a higher number of target responses and shorter response times. Whether such a pattern holds for participants with aphasia - that is, whether the effects of linguistic experience affect language processing after brain damage - motivates the investigation.

4.2 Method

4.2.1 Participants

Participants in Phase 2 included typical adults and adults with acquired aphasia. All participants were native speakers of British English with sufficient vision and hearing to complete the language tasks. Participants had no history of speech or language difficulties or psychiatric impairment. Typical participants were aged between fifty and eighty years.

Participants with aphasia received a diagnosis of acquired aphasia from a qualified speech and language therapist. They had been living with aphasia for at least six months prior to taking part in the study, and they were neurologically and medically stable at the time of testing. Participants with aphasia took part in screening tasks in order to make an informed decision about whether their language abilities were well-matched to the demands of the study.

No typical participant who took part in Phase 1 also took part in Phase 2. Participants with aphasia from Phase 1 took part in the pilot stage of Phase 2; participants with aphasia who completed the main study of Phase 2 had not taken part in any prior activity in the project.

4.2.2 Ethics, recruitment and consent

This study received ethical approval from the University Research Ethics Committee in the Department of Human Communication Sciences at the University of Sheffield.

Typical participants received a written information sheet about the study. They were given a copy of this sheet to keep if they decided to take part in Phase 2. Participants had the opportunity to ask questions before they consented to participate. Consent was obtained via a written document, and participants were given a signed copy of this consent form to keep if they decided to take part.

Participants with aphasia received an information sheet designed for readers with aphasia, following the Stroke Association's Accessible Information Guidelines (Herbert et al., 2012). Participants with aphasia were be enabled via supportive communication techniques to ask any questions they had about the research. A written information sheet was also provided to participants' family members or carers. Participants with aphasia, and their family members or carers, were given copies of these information sheets to keep if they decided to take part in the screening tasks. The researcher answered any questions from participants and their family members or carers before they consented to participate in the screening tasks. Consent was obtained via a written document designed for readers with aphasia, following the Stroke Association's Accessible Information Guidelines (Herbert et al., 2012). Participants were given a signed copy of this consent form if they decided to take part. Ethical approval for Phase 2, and the information sheets and consent forms provided to typical participants and participants with aphasia, are provided in Appendix E.

Typical participants were recruited via distribution lists from the University of Sheffield, local charitable groups and organisations for retired persons, among communities in Yorkshire and the East Midlands. Participants with aphasia were recruited from voluntary organisations for adults who have had stroke.

4.2.3 Design

Phase 2 included a computerised grammaticality judgement task and sentence completion task. In both tasks, participants read a sentence stimulus silently and then pressed a button to reveal a written verb, presented as a single word. In the grammaticality judgement task, participants decided whether the verb could be used in the sentence stimulus and indicated their decision via a button press. The task contained an equal number of 'yes' and 'no' judgements. In the sentence completion task, participants produced the entire sentence aloud, and their vocal responses were recorded. Each task contained six argument structure constructions, and a different set of verbs was included in each of the two tasks.

All participants completed each task once. Typical participants attended a single testing session, and the order of the two tasks was counterbalanced across typical participants. Participants with aphasia took part in further language assessments, in addition to the two tasks. They completed each task over two testing sessions, and attended a total of at least four testing sessions.

Table 4.1 Argument structure constructions included in Phase 2

Construction	Form	Example stimulus
Caused motion	subject – verb – oblique	They ___ it to her
Conative	subject – verb – $oblique_{at}$	We ___ at you
Ditransitive	subject – verb – $object_1$ – $object_2$	I ___ him something
Intransitive motion	subject – oblique	We ___ through there
Removal	subject – verb – $oblique_{from}$	You ___ it from me
Transitive	subject – verb – object	They ___ us

Phase 2 began with a pilot phase in order to evaluate the suitability of the procedure of the two tasks for typical participants and participants with aphasia. The main study proceeded after changes to the method had been implemented as a result of the pilot study, which is described in Chapter 5.

4.2.4 Materials

Argument structure constructions were presented as written sentences, with a blank space in place of the verb. Verbs were presented as single words subsequent to the argument structure constructions.

4.2.4.1 Argument structure constructions

Six argument structure constructions were included in each of the two tasks. Goldberg (1995; Johnson & Goldberg, 2013) identified these constructions as meaningful, abstract linguistic units in which a variety of verbs can occur. The same set of constructions used in Phase 1 were used in Phase 2, except the passive and way constructions from Phase 1 were not included in Phase 2. The constructions employed in Phase 2, their syntactic forms and an example stimulus of each are shown in Table 4.1.

As in Phase 1, argument structure constructions were composed entirely of function words, such as pronouns and prepositions, so no lexical semantic content was available from the sentence stimuli. Subject pronouns included only the forms *I*, *you*, *we* and *they*, because these forms can refer to an antecedent of either gender and agree with the uninflected verb form.

4.2.4.2 Verb stimuli

Each of the six constructions in Table 4.1 was paired with a set of four monosyllabic verbs that differed in lexical and construction frequency. Each task therefore contained a total of 24 verbs. A different set of verbs was included in the grammaticality judgement task and the sentence completion task. Verbs were presented in their uninflected forms.

Construction frequency and lexical frequency of the verbs associated with each construction were in orthogonal variation such that each set contained two pairs of verbs, where one pair

Table 4.2 Set of verb materials for conative construction in grammaticality judgement task

CONATIVE: We ___ at you		Lexical frequency	
		High	Low
Construction frequency	High	look 1151, 25	laugh 98, 17
	Low	come 1512, 1	fly 90, 1

Note. In cells containing verbs, first value is lexical frequency and second value is construction frequency.

had high construction frequency and one pair had low construction frequency (see Section 4.2.4.3 below for information on construction frequency). In each pair, one member had high lexical frequency and one member had low lexical frequency. Verbs that differed in one type of frequency were matched for the other; for example, the two verbs with high construction frequency had similar values for construction frequency, but differed in their value of lexical frequency. Lexical frequency was based on verbs' lemma frequencies in instances per million (Leech et al., 2001), and construction frequency was based on the number of times typical participants generated verbs in response to constructions in the Phase 1 verbal fluency task.

The set of verbs associated with the conative construction in the grammaticality judgement task is provided in Table 4.2 as an example. The full set of materials for both tasks is included in Appendix F.

For ease of reference, verb groups, or conditions, will be designated throughout this thesis by their levels of construction frequency and lexical frequency in square brackets. For example, verbs with high construction frequency and high lexical frequency will be denoted as [high cx, high lex]; verbs with high construction frequency and low lexical frequency will be denoted as [high cx, low lex]; verbs with low construction frequency and high lexical frequency will be denoted as [low cx, high lex]; and verbs with low construction frequency and low lexical frequency will be denoted as [low cx, low lex].

The grammaticality judgement task contained equal numbers of 'yes' and 'no' judgements. 'No' judgements were elicited by verbs that were not produced in response to a target construction in Phase 1, matched as far as possible to the lexical frequency of the 24 verbs included as experimental items. For example, the four verbs included in the grammaticality judgement task to elicit 'yes' judgements for the conative construction are shown in Table 4.2. Each of these verbs was matched to a verb of similar lexical frequency that participants in Phase 1 did not produce in response to the conative construction, i.e. the verb had a construction frequency of zero. The lexical frequencies of the verbs included in the grammaticality judgement task to elicit a 'no' judgement are included in the description of materials provided in Appendix G.

4.2.4.3 Defining lexical and construction frequency

Lexical frequency referred to the lemma frequency of a verb in instances per million, that is - how often any form of the verb occurs per million words. Lexical frequency values were based on the British National Corpus and taken from Leech et al. (2001). Construction frequency referred to the number of times typical adults produced the verb in response to the argument structure construction in the Phase 1 verbal fluency task. The maximum value for construction frequency was 40, which reflects each of the twenty typical participants in Phase 1 naming the verb in response to each of the two versions of the construction included in the verbal fluency task.

Most research on frequency specifies numerical values that define 'high' and 'low' frequency. However, no established norms exist on what constitutes 'high' and 'low' lexical frequency. Keuleers, Brysbaert and New (2010) noted that researchers often define these terms differently, usually specifying low frequency words as occurring fewer than five or ten times per million. In research on aphasia, low frequency can be defined by as many as 20 instances per million (e.g. Druks and Masterson's (2000) *An Object and Action Naming Battery*).

Defining high and low lexical frequency as a single numerical value across was not possible across both tasks in Phase 2, because verbs were selected based on construction frequency as well as lexical frequency. Therefore it was not possible to select a numerical value above which all verbs included as high lexical frequency items occurred, and below which all verbs included as low lexical frequency items occurred. Rather, high and low lexical frequency were determined relatively within each construction. Of the verbs most frequently produced in response to each construction in Phase 1, i.e. those with high construction frequency, verbs with the highest and lowest lexical frequency values were selected. Verbs with high construction frequency were then matched with verbs of similar lexical frequency that were produced only once or twice in response to each construction in Phase 1, i.e. those with low construction frequency. This procedure resulted in a set of four verbs per construction and was repeated to create another set, for a total of two sets of four verbs per construction for use in the two sentence processing tasks.

The measure of lexical frequency used as the basis for selecting verbs for inclusion in the experiments was lemma frequency, or how often verbs occurred in English in any form, e.g. inflected with the morpheme -*s* or -*ing*. Three other lexical frequency measures of the verbs included in the final set of materials were checked, including the frequency of the

Table 4.3 Mean construction frequency and lexical frequency values for verbs in each condition of Phase 2 grammaticality judgement task

Verb Group	Construction Frequency	Lexical Frequency
high cx, high lex	17 (7)	533 (323)
high cx, low lex	14 (6)	49 (29)
low cx, high lex	1 (0)	514 (505)
low cx, low lex	1 (0)	51 (29)

Note. Means, with standard deviations in brackets.

Table 4.4 Results of independent samples *t*-tests evaluating differences in construction frequency and lexical frequency between verbs in each condition of Phase 2 grammaticality judgement task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
Construction frequency				
high cx, high lex	-			
high cx, low lex	$t(10) = 0.81$ $p = 0.439$	-		
low cx, high lex	$t(5) = 5.34$ $p = 0.003$*	$t(5) = 5.19$ $p = 0.003$*	-	
low cx, low lex	$t(5) = 5.34$ $p = 0.003$*	$t(5) = 5.19$ $p = 0.003$*	All verbs' cx freq = 1	-
Lexical frequency				
high cx, high lex	-			
high cx, low lex	$t(10) = 6.52$ $p < 0.001$*	-		
low cx, high lex	$t(10) = 0.53$ $p = 0.607$	$t(10) = -5.14$ $p < 0.001$*	-	
low cx, low lex	$t(10) = 5.77$ $p < 0.001$*	$t(10) = 0.06$ $p = 0.956$	$t(10) = 4.70$ $p = 0.001$*	-

Note. * indicates significant *p*-value, at level of Bonferroni correction ($p < 0.01$).

verb form included in the study (i.e. the frequency of the uninflected verb form, in instances per million from Leech et al. (2001)), raw frequency (from Davies (2004-)) and the frequency of verbs as letter strings (raw frequency from Davies (2004-)). Correlations were strong, positive and significant for these three additional lexical frequency measures, including the frequency of the uninflected verb form in instances per million ($r = 0.97$, $p < 0.001$), the raw frequency of the verb form ($r = 0.97$, $p < 0.001$) and the frequency of the letter string ($r = 0.97$, $p < 0.001$).

4.2.4.4 Confirmation of frequency differences between verb groups

T-tests were used to confirm intended differences in construction frequency and lexical frequency between verbs in each condition in the Phase 2 tasks. Because multiple comparisons were performed on each set of verbs, a Bonferroni-corrected *p*-value of 0.01

was used ($\alpha = 0.05 \div 5$). Lexical frequency values were log transformed, and statistical tests were carried out in SPSS.

Verbs in the grammaticality judgement task

The means and standard deviations of the construction frequency and lexical frequency values of verbs in each condition in the grammaticality judgement task are shown in Table 4.3. Results of the independent samples *t*-tests evaluating the differences in construction frequency and lexical frequency are shown in Table 4.4.

All intended differences in construction frequency and lexical frequency were evident in the set of verbs in the grammaticality judgement task. In addition, conditions that were controlled for one type of frequency were not significantly different from one another.

Verbs in the sentence completion task

The means and standard deviations of the construction frequency and lexical frequency values of verbs in each condition in the sentence completion task are shown in Table 4.5. Results of the independent samples *t*-tests evaluating the differences in construction frequency and lexical frequency are shown in Table 4.6.

All intended differences in construction frequency and lexical frequency were evident in the set of verbs in the sentence completion task. In addition, conditions that were controlled for one type of frequency were not significantly different from one another.

4.2.5 Sentence stimuli

Each verb was paired with a sentence stimulus that took the form of an argument structure construction. Sentences were composed of function words, such as pronouns and prepositions. Sentences did not contain any lexical semantic information, so the sentence stimuli encoded only the meaning of the event referred to by the argument structure construction.

As in Phase 1, sentences began with the subject pronouns *I*, *you*, *we* and *they*, because these subject pronouns occur with the uninflected form of a verb. The pronouns *he, she* and *it* were not included in the materials, because these pronouns occur with a verb form marked with an *–s* for third person singular agreement. Thus, participants were not required to read verb inflections in the grammaticality judgement task, or to produce inflections in the sentence completion task. Stimuli sentences contained the object pronouns *me*, *you*, *him*, *her*, *it*, *us* and *them*.

Table 4.5 Mean construction frequency and lexical frequency values for verbs in each condition of Phase 2 sentence completion task

Verb Group	Construction Frequency	Lexical Frequency
high cx, high lex	20 (10)	1339 (726)
high cx, low lex	13 (7)	155 (90)
low cx, high lex	1 (0.4)	909 (688)
low cx, low lex	1 (0.4)	140 (71)

Note. Means, with standard deviations in brackets.

Table 4.6 Results of independent samples *t*-tests evaluating differences in construction frequency and lexical frequency between verbs in each condition of Phase 2 sentence completion task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
Construction frequency				
high cx, high lex	-			
high cx, low lex	$t(10) = 1.44$ $p = 0.181$	-		
low cx, high lex	$t(5) = 4.79$ $p = 0.005$*	$t(5) = 4.50$ $p = 0.006$*	-	
low cx, low lex	$t(5) = 4.79$ $p = 0.001$*	$t(5) = 4.50$ $p = 0.006$*	All verbs' cx freq = 1 or 2	-
Lexical frequency				
high cx, high lex	-			
high cx, low lex	$t(10) = 4.86$ $p = 0.001$*	-		
low cx, high lex	$t(10) = 0.96$ $p = 0.359$	$t(10) = -4.22$ $p = 0.002$*	-	
low cx, low lex	$t(10) = 5.53$ $p < 0.001$*	$t(10) = 0.06$ $p = 0.952$	$t(10) = 4.90$ $p = 0.001$*	-

Note. * indicates significant *p*-value, at level of Bonferroni correction ($p < 0.01$).

Each stimulus sentence in both tasks was a unique combination of subject and object pronouns. Combinations of subject and object pronouns were determined pseudo-randomly with the constraint that object pronouns made reference to a separate entity than the subject pronoun; this constraint prevented the formation of nonsensical sentences such as *I ___ at me*.

Participants saw an equal number of sentences that began with the same subject pronoun. Each subject pronoun was used once in each task per construction. Within each construction, object pronouns were not repeated. For example, the sentence stimuli for the conative construction in the sentence completion task included the stimuli sentences *we ___ at you*, *you ___ at us*, *they ___ at me* and *I ___ at her*. Each sentence contained a different subject and object pronoun, and the pairing of subject and object pronouns in each sentence stimulus was unique.

The same subject pronoun was not assigned to verbs in the same frequency group of the same construction across tasks. To illustrate, [high cx, high lex] verbs for the conative construction included *talk* in the sentence completion task, *look* in the grammaticality judgement task to elicit a 'yes' judgement and *say* in the grammaticality judgement task to elicit a 'no' judgement. These three verbs had comparable lexical and - for two of the verbs - construction frequencies, so they were each paired with sentence stimuli that contained different subject pronouns.

In three instances, sentence stimuli for the removal construction contained the word *there*, to match form of the construction that elicited the verb in Phase 1. This permitted the grammatical inclusion of some verbs in the removal construction, such as *lift* in *I lift it from there* and *reach* in *they reach it from there*.

4.2.6 Stimuli order

The order of stimuli within each task was determined pseudo-randomly following a number of conditions: (1) adjacent trials did not contain verbs that shared an initial phoneme; (2) no more than three verbs from the same verb group (i.e. [high cx, high lex], [high cx, low lex], [low cx, high lex] or [low cx, low lex]) appeared consecutively; and (3) no more than three consecutive 'yes' or 'no' judgements were required in the grammaticality judgement task.

4.3 Procedure

The grammaticality judgement task and the sentence completion task were presented on a laptop computer via the software PsychoPy (Peirce, 2007).

4.3.1 Grammaticality judgment task

In the grammaticality judgement task, participants decided whether verbs could be used in sentence contexts. Each trial involved the sequential presentation of a written sentence and a written verb. A blank space stood in place of the verb in the written sentence. After the verb appeared on screen, participants pressed a button to indicate whether or not the verb could be used in the sentence stimulus. Button presses were performed by a participant's preferred hand via a horizontal button box, because participant handedness has been shown to influence the speed of motor responses (Goodin, Aminoff, Ortiz & Chequer, 1996). The 'yes' button was coloured green and the 'no' button was coloured red. Participants decided on the orientation of the button box, i.e. whether the 'yes' button was placed on the right or left.

The button box contained three bottom-activated buttons. A white button was located in the centre of the button box, and the red and green buttons were positioned at the edges of the button box. The box operated by means of an Arduino Leonardo microcontroller, and it connected to the laptop via a USB cable.

For example, a trial could take the following form:

We _____ at you

LOOK

Sentence text was black in colour, centred on a white background. Trials began with the appearance of a sentence stimulus. Participants read the sentence silently, with no time limit. They then pressed the white button to reveal a verb, presented as a single word in red. The sentence stimulus remained on screen, so no recall was involved. Participants decided whether the verb could be used in the sentence stimulus and recorded their decision via a button press. They pressed the green button if the verb could be used in the sentence, or the red button if it could not be used in the sentence. There was no time limit for the decision. The button press terminated the trial. The sentence of the next trial appeared after 500 milliseconds.

The trial design is shown in Figure 4.1. Activity that occurred on the computer is shown above the timeline, using the example cited above, and the responses that participants made are shown below the timeline.

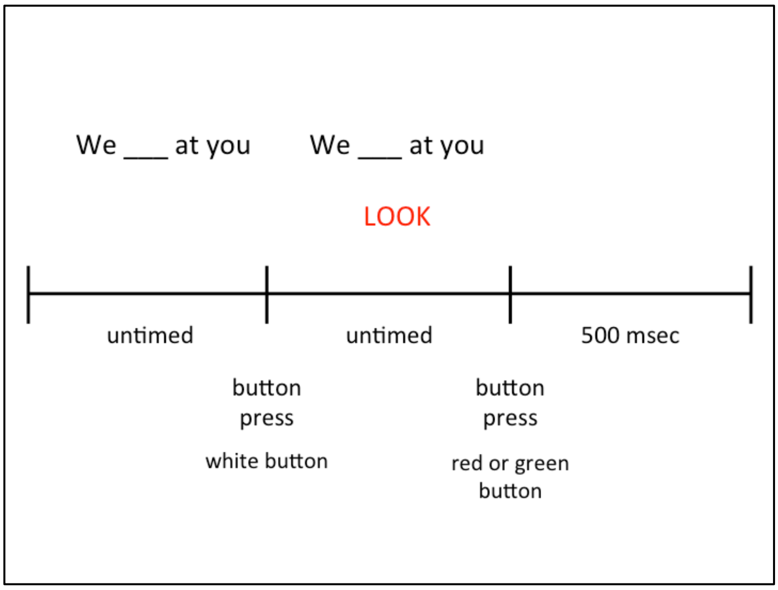


Figure 4.1 Trial design of Phase 2 grammaticality judgement task

The grammaticality judgement task began with twelve practice items, including a 'yes' and 'no' judgement for each of the six argument structure constructions. The task contained 24 trials designed to elicit 'yes' judgements and 24 trials designed to elicit 'no' judgements.

Participants were given the following instructions in the grammaticality judgement task:

> You will see a sentence with a blank space in place of one word. Press the white button to show a word. I would like you to decide whether the word can be used in the sentence. When you're making your decision, it's not whether the word is correct or incorrect like you learned in school, but whether it's something you would hear in the language around you. Press the green button if the word can be used in the sentence and the red button if the word cannot be used in the sentence. Do you have any questions?

Participants' questions were answered before proceeding. Typical participants then completed the practice items and the 48 trials without interruption.

4.3.2 Sentence completion task

In the sentence completion task, participants produced verbs in a sentence context. The format of each trial was the same as for the grammaticality judgement task, but participants were not required to make a yes-no judgement on the grammaticality of the verb in the sentence stimulus. Participants pressed a white button to reveal a verb, and this button press also activated a microphone. Participants produced the entire sentence aloud, replacing the blank space in the sentence with the verb. Both the sentence and the verb remained on screen during production, so no recall was involved. Participants pressed the white button when they were finished speaking, and this button press terminated the audio recording. The design of these trials is shown in Figure 4.2, where activity on the computer screen is shown above the timeline and participants' responses are shown below the timeline.

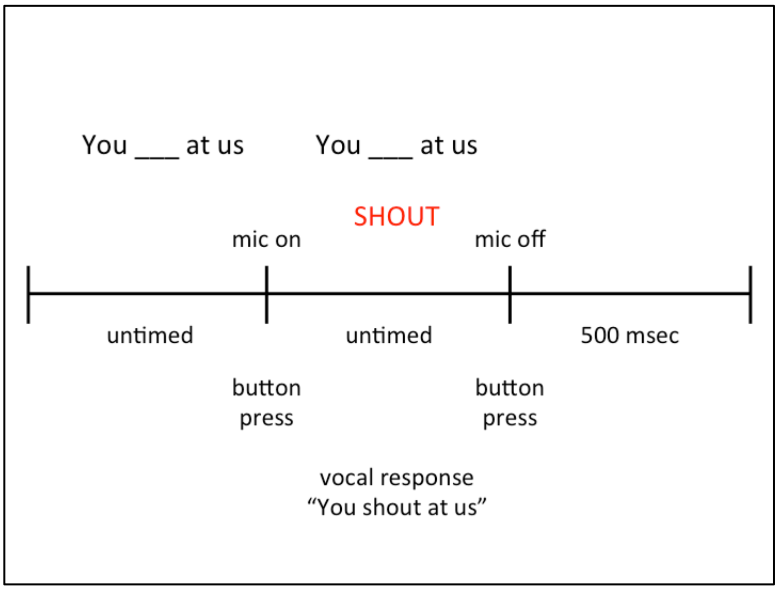


Figure 4.2 Trial design of Phase 2 sentence completion task

The sentence completion task began with six practice items, including one instance of each of the six argument structure constructions. The task contained a total of 24 items.

Participants were given the following instructions in the sentence completion task:

> You will see a sentence with a blank space in place of one word. Press the white button to show a word. Put the word into the blank space and say the entire sentence out loud. Press the white button when you are finished speaking to move on. Do you have any questions?

Participants' questions were answered before proceeding. Typical participants then completed the practice items and the 24 trials without interruption.

4.3.3 Data recording

Participants' audio responses in the sentence completion task were captured via a Samson C01U Pro condenser microphone, connected to the laptop via a USB. The microphone was positioned to the left of the laptop computer, approximately 18 inches away from participants' mouths.

4.4 Data analysis

This section describes how participants' responses in the grammaticality judgement and sentence completion tasks were captured and processed, and provides a background to the statistical analysis performed on the datasets.

4.4.1 Dependent variables for analysis

The number of target responses and response times in the grammaticality judgement and sentence completion tasks were recorded and analysed, as described below.

4.4.1.1 Dependent variables in the grammaticality judgement task

The experimental presentation software, PsychoPy, created a log file each time the grammaticality judgement experiment was run. This file contained information on whether the participant's judgement matched the target, i.e. the 'yes' or 'no' response, and the response time for each trial in the task. These data were extracted from the log files and entered into the analysis.

4.4.1.2 Dependent variables in the sentence completion task

In addition to a log file, the PsychoPy software captured the microphone's input as an audio file for each trial in the sentence completion task. The author manually checked audio files from all participants for the presence of speech errors. Participants' productions were included in the analysis if they represented the correct reading of the target sentence. Errors in pronoun production were accepted, e.g. the production of a subject pronoun other than the one contained in a sentence stimulus, provided the error did not impact the fluency of sentence production. For example, a trial was discarded if participants produced an inaccurate subject pronoun but immediately corrected it, because this mistake affected the measurement of response times. Productions were excluded from the analysis if they contained a non-target verb, a non-target construction or any material prior to the production of the sentence that affected the measurement of response times, such as commentary, conversation with the researcher or coughing.

The response time of participants' spoken productions in the sentence completion task was investigated as verb production latencies. This measure represented the time from the presentation of the verb on screen to participants' production of the verb in a sentence context. All verbs were preceded by monosyllabic subject pronouns. Verb production latencies were extracted from audio files via a custom-made computer script that ran on a freely available software programme for the analysis of audio files (Cunningham, 2017), as described below.

Automatic extraction of verb production latencies

Each participant response was stored in a separate audio file. The researcher produced an orthographic transcription, and it was associated with each audio file. A custom script created a TextGrid file for each audio file, given the associated transcription, in the software Praat (Boersma & Weenink, 2017). A TextGrid can be used to store different levels of transcription, such as orthographic and phonetic. The PraatAlign tool (Lubbers & Torreira, 2016) was then used to produce a phonetic transcription. The PraatAlign tool takes data from Praat, the audio file and the orthographic transcription and produces a phonetic alignment based on speech recognition models. The speech recognition alignment used the hidden Markov model toolkit (HTK; Young et al., 2006). The alignment process used the British English models supplied with PraatAlign (Schiel, 1999).

Quality control of automatically extracted verb production latencies

After the initial alignment, the researcher inspected results. Any speech onsets that were more than two standard deviations away from the mean onset were visually inspected. Most of these outliers resulted from extraneous noise at the start of the audio recording, due to the release of the button press, causing the TextGrids to align to this acoustic artefact rather than participants' start of speech. In these cases, alignments were repeated with the initial 500 milliseconds of audio removed from the file that was aligned. Any remaining files with start times of less than 400 milliseconds were excluded from analysis as experimental error.

Participants with aphasia produced more variable sentence productions than typical participants. They sometimes paused between words, whereas typical participants produced all sentences fluently. The researcher visually inspected the 264 audio files from participants with aphasia and their associated TextGrids to ensure that the value the script extracted from the audio file as the verb production latency did in fact correspond to the measurement at that point in time in the audio file. This ensured the value did not correspond to, for example, a period of silence preceding the production of the verb. In cases where the script extracted a value that did not correspond to the verb production latency, the value that actually did was used in the analysis.

Reliability of verb production latencies

The reliability of verb production latencies was investigated by comparing script-generated measurements to manually-derived measurements using intraclass correlation coefficients (ICCs). ICCs assess agreement between two raters in situations where neither rater is treated as a reference, based on the statistical test ANOVA (Barnhart, Haber & Lin, 2007). ICCs were calculated in SPSS.

The researcher measured verb production latencies by hand for three randomly-selected trials from every typical participant. These 270 trials comprised 12.5% of the dataset from typical participants. There was a high level of agreement between the verb production latencies as measured by the script and by the researcher ($ICC = 0.93$).

The researcher measured verb production latencies by hand for five randomly-selected trials from each participant with aphasia who took part in the sentence completion task. These 55 trials comprised 20.8% of the dataset from participants with aphasia. There was a high level of agreement between verb production latencies as measured by the script and by the researcher ($ICC = 0.91$).

4.4.2 Planned analyses

4.4.2.1 Planned analyses of data from typical participants

Response time data from typical participants were analysed as a two-way, factorial analysis of variance (ANOVA). Construction frequency and lexical frequency were included as the independent variables. ANOVAs were carried out in SPSS. Effect sizes are reported as r (Field, 2009). Significant interactions were explored with paired-samples t-tests.

Simultaneous random effects in ANOVA

Statistical analysis in linguistics has historically suffered from what Clark (1973) termed the ‘language-as-fixed-effect fallacy’ (Clark, 1973, p. 336). He investigated studies in which researchers performed statistical analyses that treated participants as random effects but the linguistic items under consideration as fixed effects. Random effects are those which are randomly sampled from a population, and therefore a significant result indicates the effect generalises to that population. In contrast, fixed effects are those which are selected by the researcher, and therefore no claims can be made about how the effect generalises to a population (Tabachnick & Fidell, 2007). Clark (1973) observed that language researchers performed analyses that treated linguistic phenomena as fixed effects, yet made claims about their findings as if they were random effects. Clark (1973) provided some procedures for correcting for this assumption by calculating the *F*-statistic in ANOVA twice - once on participant means and once on item means. The two *F*-statistics can be compared using a formula to compute the statistic minF', which reflects whether results are significant when both participants and items are treated as random effects, and thus whether findings generalise to populations of both participants and items. This procedure was adopted in the analysis of data from typical participants. Repeated-measures ANOVAs were used to analyse by-participant data, and independent ANOVAS were used to analyse by-item data. The minF' statistic was calculated manually following the equation in Clark (1973).

Additional analyses evaluating the effect of single-word processing

Response times in the grammaticality judgement and sentence completion tasks reflected participants' response to a verb, given prior exposure to a sentence context. Participants' response times in the tasks therefore reflected multiple related psycholinguistic processes, including, at minimum:

(1) single-word processing – participants spent some time recognising the verb as a single word;

(2) verb-construction integration – participants spent some time comparing or integrating the verb and sentence stimulus; and

(3) response production – participants produced a response, by either making a grammaticality judgement or producing a sentence.

This study predicted that responses to high frequency verbs would be faster than to low frequency verbs. However, a number of factors are known to affect (1), above, single-word processing.

Response times in single-word processing tasks are influenced by a number of properties of the single word, including a lexical item's frequency, length and lexical neighbourhood. Regression models that test the effect of these variables on response times to single words identify frequency as the most important predictor, accounting for 40.5% of the variance (Brysbaert et al., 2011). Recent research has identified age-of-acquisition as another explanatory factor in single word processing tasks (Kuperman et al., 2012; see Section

2.2.1.3), and semantic variables such as imageability can also affect processing times (Cortese & Fugett, 2004). It was not possible to match the verbs included in Phase 2 on all variables known to affect single-word processing, such as length, neighbourhood size and imageability. Additional analyses were undertaken in order to ensure that any effects observed in the main analysis survived other effects related to language processing. These additional analyses also investigated the effect of age, gender, education and - in the grammaticality judgement task - handedness on responses from typical participants.

Another attribute of the verbs in Phase 2 included in additional analyses was verbs' response times in a lexical decision task. Lexical decision times were taken to reflect the relative ease or difficulty of processing verbs in isolation, simultaneously accounting for the characteristics of single words described above. Lexical decision times were taken from the British Lexicon Project, which provides published data on lexical decision times for several thousand English words, based on the performance of undergraduate and graduate students at the University of London (Keuleers et al., 2012).

If response times in the Phase 2 tasks were related to any of the measures described above, results from Phase 2 could be argued simply to reflect single-word processing. In order to ensure any effects identified in Phase 2 were due to processes over and above single-word processing, i.e. the hypothesised integration of verbs and constructions, significant correlates to response times were included as covariates in an ANCOVA to test whether the effect of construction frequency or lexical frequency survived the inclusion of the related effect in the statistical model.

4.4.2.2 Planned analyses of data from participants with aphasia

Participants with aphasia were analysed as a group and as individuals. Data from participants with aphasia were analysed using non-parametric statistical tests, including the Wilcoxon signed-ranks test for paired samples and the Mann-Whitney test for independent samples. Non-parametric tests were used instead of parametric tests due to the small number of data points in samples from participants with aphasia and lack of reliable method for replacing extreme response times that would affect the accuracy of the mean. These tests were carried out in SPSS.

Comparisons between the performance of participants with aphasia and typical participants were made following methods advanced by Crawford and his colleagues. Historically, the standard method in neuropsychology research for comparing the performance of a clinical case to a group of typical participants, or controls, was the use of z-scores. A z-score refers to the point on the normal curve corresponding to a case's performance, and statistical significance can be inferred from this point, with a case performing significantly worse than controls if the score corresponds to a *z*-score of -1.64 or lower (Crawford & Howell, 1998). Crawford and Howell (1998) observed that the use of *z*-scores overestimates the abnormality of a case's score, because it treats the parameters of the sample statistics from controls as if they were population parameters. The authors argue that the use of *z*-scores in neuropsychological hypothesis testing should in fact be avoided.

In response to the need for more accurate methods of analysis in neuropsychology, Crawford, Howell, Garthwaite and their colleagues devised various tests to evaluate the performance of cases in comparison to typical participants. Their methods are based on a modified *t*-distribution, rather than *z* (Crawford & Howell, 1998). The methods produce the *p*-value associated with this test statistic, as well as a measure of the rarity of a case's score, in terms of the estimated percentage of the typical population that would demonstrate more extreme performance than the case, and confidence intervals for that estimate (Crawford & Garthwaite, 2002). The method of analysis relevant to the current investigation is the standardised difference test, which compares the difference between a case's performance in two tasks to the difference the typical population shows in the tasks. The tasks are defined by the means and standard deviations of typical participants. In the present study, the two 'tasks' were defined as responses to high and low frequency verbs. The standardised difference test was applied once to compare each participant with aphasia to typical participants in terms of the difference shown to verbs with high and low construction frequency and again to make the comparison for the difference shown to verbs with high and low lexical frequency. Both dependent variables, the number of target responses and response time, were investigated in this way.

The most recent version of the standardised difference test is based on a Bayesian approach to statistics. This approach differs from the classic, or frequentist, approach to statistics in the treatment of parameters and the interpretation of interval estimates. In the frequentist approach, parameters are treated as fixed but unknown, whereas in the Bayesian approach parameters are treated as random variables with associated probability distributions (Crawford & Garthwaite, 2007). In the frequentist approach, interval estimates are known as confidence intervals, but in the Bayesian approach they are known as credible intervals. The interpretation of the two types of interval vary slightly. Specifically, a frequentist interpretation of a 95% confidence interval of X can be stated as:

> If we could compute a confidence interval for a large number of control samples collected in the same way as the present control sample, about 95% of them would contain the true [value of X] (Crawford & Garthwaite, 2007, p. 348).

In contrast, a Bayesian interpretation of a 95% credible interval can be stated as:

> There is a 95% probability that the true [value of X] lies within the stated limits (Crawford & Garthwaite, 2007, p. 348).

Crawford and Garthwaite (2007; Crawford, Garthwaite & Porter, 2010) argue that the Bayesian interpretation is more consistent with the way most psychologists understand interval estimates.

The Bayesian Standardised Difference Test (BSDT) is preferable to the standardised difference test based on the frequentist approach because it can discriminate between situations in which a case's scores are very extreme from situations in which a case's scores are less extreme, even if the difference between the standardised scores is the same. Additionally, unlike the frequentist-based test, it produces interval estimates of the abnormality of the difference between a case's standardised scores (Crawford & Garthwaite, 2007). The equations for the BSDT and its associated effect sizes can be found in Crawford and Garthwaite (2007) and Crawford et al. (2010).

Bayesian Standardised Difference Tests were carried out on the freely-available computer programme DiffBayes_ES.exe. The standards of reporting follow those recommended by Crawford et al. (2010).

The effects of construction frequency and lexical frequency on responses in the Phase 2 tasks were represented as effect scores, following the method employed by Huck et al. (2017). These authors used effect scores to represent differences between response times to items with different levels of frequency as proportions (response time to low frequency items ÷ response time to high frequency items). From this proportion, 1 was subtracted to gain an effect score, so positive values represent frequency effects in the predicted direction and negative values represent frequency effects in the reverse direction. To illustrate, an effect score of 0.5 represents response times that are 50% longer to low frequency verbs than to high frequency verbs. Importantly, both Crawford and Garthwaite's (2007) Bayesian Standardised Difference Test and Huck et al.'s (2017) method of calculating effect scores account for differences between response times with reference to their overall length; that is, small differences between long response times are treated as less impactful than small differences between shorter response times.

4.4.2.3 Power analyses of Phase 2 study

'Power' refers to the likelihood of a statistical test to produce a statistically significant outcome (Cohen, 1988). Power is related to the sample size of a study, the size of the effect under investigation and the selected *p*-value (Cohen, 1992). Power analyses were produced after the completion of Phase 2 data analysis in order to evaluate whether the study was sufficiently powered to detect the effects under investigation, and to highlight directions for further research. Analyses were conducted on the Power and Precision software.

Planned analyses of data from adults with aphasia in Phase 2 involved non-parametric statistical tests, for the reasons described in Section 4.4.2.2. There is a recognised lack of accessible methods for evaluating the power of non-parametric tests:

> Virtually all treatments of power analysis…focus on parametric tests […] Whilst these tests are more powerful than their non-parametric counterparts, there has been scant attention paid to the power of non-parametric methods. Even the 'bible' of power analysis (Cohen, 1988) does not describe how to assess non-parametric power (Mumby, 2002, p. 85).

In light of this, power estimates were reported based on the assumptions of performing a parametric analysis, and commentary on the impact of the non-parametric approach on these estimates is provided in the following results sections.

4.5 Language assessments for participants with aphasia

In addition to completing the grammaticality judgement and sentence completion tasks described in this chapter, participants with aphasia took part in additional screening and assessment. Appendix H lists the items included in novel assessments in Phase 2.

4.5.1 Screening

Participants with aphasia completed three screening tasks before agreeing to take part in Phase 2. These tasks were designed with the twin purposes of (1) giving participants the opportunity to experience of the types of tasks they would be asked to complete in the project, and (2) indicating whether participants with aphasia retained sufficient residual language processing to meet the demands of the novel experimental tasks in Phase 2. The screening tasks included a sentence reading task, a function word processing task and a grammaticality judgement task. These tasks were administered in this order for all participants with aphasia.

Sentence reading screening

The sentence reading task contained five items from PALPA 37 (Kay, Lesser & Coltheart, 1992). This screening task was included to indicate whether participants could process written sentences. No verb that appeared in the experimental tasks was included in this screening task. Items in this task were scored as correct if participants were able to read aloud the verb and each argument in a sentence. Semantic or phonological errors were accepted. Each sentence was presented individually on an A5 sheet of paper as a single line of black, size 36, bold text in Helvetica font.

Function word processing screening

The function word processing task was included to indicate whether participants could process the lexemes included in the sentence stimuli of the experimental tasks. This screening task contained five function words and five non-words, formed by changing one letter of the function words. Real word items were pronouns and prepositions from Gilner and Morales (2005). The task did not contain any word that appeared in the experimental tasks. Participants were asked to read each word silently and respond 'yes' or 'no' to whether the word was a real word of English, in the manner of a lexical decision task, like that employed in Phase 1 and by Herbert et al. (2014). Words were presented as a list centred on an A4 sheet of paper in black, size 22, bold Helvetica font. The task required no more than two consecutive 'yes' or 'no' responses, and non-words were not adjacent to the real words from which they were derived.

Grammaticality judgement screening

A grammaticality judgement task was included to indicate whether participants could recognise violations in argument structure, as this skill was necessary in the experimental grammaticality judgement task. Five items from Kim and Thompson (2000) were included. These items contained sentences with verbs presented in grammatical argument structures, and with verbs presented in sentences that contained an extra argument. No verb that appeared in the experimental tasks was included in this screening task. Each sentence was presented individually in black, size 36, bold Helvetica type on an A5 sheet of paper.

4.5.2 Assessment

Participants with aphasia completed three main assessments in Phase 2 in order to allow their performance in the novel experimental tasks to be related to their cognitive and language function. Assessments permitted a profile of their aphasia to be created, including a diagnosis of their aphasia type. Assessments included (1) tests of processing the experimental stimuli; (2) cognitive and language subtests from the *Comprehensive Aphasia Test* (CAT) (Swinburn, Porter & Howard, 2004); and (3) additional assessments of verb and sentence processing, including subtests from *The Verb and Sentence Test* (VAST) (Bastiaanse, Edwards & Rispens, 2003), items from *An Object and Action Naming Battery* (OANB) (Druks & Masterson, 2000) and novel assessments created for this study.

Typical participants in Phase 2 contributed normal data for the novel assessments used in the study. Each typical participant completed three tasks in an assessment session, including the experimental grammaticality judgement and sentence completion tasks, followed by one of the five tasks for which normal data was needed: a novel anagram task, sentence production task, lexical decision of function words, action-naming task or object-naming task.

4.5.2.1 Processing of experimental stimuli

Participants' processing of the experimental stimuli was assessed via a lexical decision and a grammaticality judgement task.

Lexical decision of function words

The lexical decision task was included to evaluate whether participants could process the words contained in the sentence stimuli in Phase 2. The task contained all the function words included in the sentence stimuli in the grammaticality judgement and sentence completion tasks. The lexical decision task contained these pronouns and prepositions as real words and non-words formed by changing one letter of each of the real words. Non-words were orthographically and phonologically illegal letter strings in English. The task contained a total of 30 items. Items were presented as single words in black, size 22 bold Helvetica type, centred on an A4 sheet of paper.

Grammaticality judgement task

This grammaticality judgement task was included as an assessment in Phase 2 to determine whether participants recognised the sentences in the experimental tasks as grammatical structures in English. The task contained the items from the grammaticality judgement task in Phase 1 that corresponded to the six constructions included in Phase 2, but only written sentence forms were presented as stimuli in Phase 2. Each construction corresponded to a grammatical sentence and an ungrammatical sentence, formed by changing the order of the words in the verb phrase. Sentences that instantiated the same construction contained different words. This task contained a total of 12 sentences.

4.5.2.2 Comprehensive Aphasia Test

The CAT Cognitive Screen was administered in order to identify any cognitive difficulties of participants with aphasia. The Cognitive Screen included a line bisection task and tasks of semantic memory, word fluency, recognition memory, gesture use and arithmetic.

Subtests of the CAT relating to language comprehension and production were also administered. The results of these tasks allowed participants' type of aphasia to be classified. Tests of language comprehension included the comprehension of spoken and written words, spoken and written sentences, and spoken paragraphs. Tests of expressive language included the repetition of simple and complex words, non-words, digit strings and sentences; the naming of objects and actions; and the reading aloud of complex words, function words and non-words. Assessment of spontaneous speech took the form of a picture description task in the CAT.

The CAT writing subtest and disability questionnaire were not administered.

4.5.2.3 Further assessment of verb and sentence processing

Participants' knowledge of verbs and sentences was further assessed via subtests from the VAST, bespoke naming tasks containing items from the OANB and novel assessments for this study, including a sentence production task and an anagram task. These assessments were included in Phase 2 to supplement the assessment of verbs and sentences contained in the CAT. For example, the CAT naming subtests include the assessment of 24 objects but only five actions. These further assessments allowed a more thorough assessment of participants' verb and sentence processing abilities that that afforded by the CAT.

The Verb and Sentence Test

Four subtests from *The Verb and Sentence Test* were administered in Phase 2. In a verb comprehension task, participants were asked to select one of four images that matched a spoken verb. In two verb production tasks, participants were asked to supply a missing verb in a sentence context. Participants were shown images and written sentences with missing verbs and had to complete the sentence with a finite or infinite verb, as appropriate. In a sentence construction task, participants were asked to describe a picture in one sentence. Responses were accepted if they consisted of any plausible description of the target picture.

An Object and Action Naming Battery

Materials for action- and object-naming tasks were taken from *An Object and Action Naming Battery*. Each task contained thirty items, and participants were asked to describe images in one word. The action-naming task contained ten verbs from each of the three frequency ranges identified in the OANB: low frequency verbs occurred 20 times per million or fewer, medium frequency verbs occurred between 21 and 100 times per million, and high frequency verbs occurred more than 100 times per million (frequency counts from Francis & Kučera (1982)). Five transitive and five intransitive verbs were included from each frequency range. Nouns were matched to the frequency of the verbs according to the manual provided by Druks and Masterson (2000). Items were presented to all participants in the same pseudo-random order, following the constraints that the list contained no more than three consecutive transitive or intransitive verbs, and no adjacent verbs began with the same phoneme. Items in the object-naming task were also ordered such that no adjacent objects began with the same phoneme. Each task began with two practice items. All images were presented individually on A4 sheets of paper.

Sentence production

In a novel sentence production task, participants were asked to produce a sentence that contained a written verb. Because this task did not involve any images, verbs that were not easily represented pictorially could be included in this assessment. In order to maximize the range of potential argument structures participants might produce, the task included two verbs produced in response to each of the eight constructions in Phase 1, including one verb each of high and low lexical frequency. Low frequency verbs ranged from one to 12 instances per million, and high frequency verbs ranged from 142 to 460 instances per million (Leech et al., 2001). In addition to these 16 items, four additional items classed as three-place verbs by Kim and Thompson (2000) were included, because three-place verbs were not included in the OANB.

Verbs were presented individually in lower case letters in bold, size 40, Helvetica font on size A5 sheets of paper. Verbs were presented in the present participle form, i.e. *hiking*. All participants saw the same pseudo-random list order, with the conditions that no adjacent verbs began with the same phoneme and no more than three consecutive verbs were high or low frequency. The task began with two practice items. Sentences were scored as incorrect if the structure was not produced by any typical participant who contributed normal data to the task. For example, all typical participants produced the verb *deserting* in a sentence that contained a direct object; therefore, productions from participants with aphasia that did not contain a direct object were counted as incorrect.

Anagram task

A novel anagram task was based on the method for the anagram task included in the VAST but contained materials uniquely tailored to the constructions in Phase 2. Participants were given a set of cards and asked to create a sentence. Each card contained one constituent of a sentence. For example, the sentence *an artist needed the paintbrush* was presented on three cards, each one containing a phrase *an artist*, *needed* or *the paintbrush*. Sentences were scored on the order in which participants placed the cards, regardless of any sentence that participants may have produced by reading aloud. To illustrate, for the target sentence *a professor taught her class the information*, the production *a professor taught the information in her class* was scored as incorrect: despite being grammatically acceptable with the insertion of the word *in*, the production did not match the constituent order of the target sentence. Conversely, the inaccurate reading aloud of correctly placed cards was not penalised.

The task included two sentences corresponding to each of the six argument structure constructions that were included in the experimental tasks in Phase 2, for a total of 12 items. One sentence contained a verb with high construction frequency and one sentence contained a verb with low construction frequency. There was no overlap between verbs included in the anagram task and the experimental grammaticality judgement or sentence completion tasks. All words were presented in lower case, bold, size 40 Helvetica font, on A5-sized cards, with no capitalisation or punctuation to cue card order.

4.5.2.4 Aphasia types

The aphasia type that best characterised the language impairment of each participant with aphasia was determined following the description of aphasia sub-types provided in Davis (1993), who proposed a system to evaluate aphasia types based on participants' fluency, comprehension and repetition abilities.

For the purpose of classifying aphasia types, participants' comprehension ability was taken as their composite score in the subtests of the *Comprehensive Aphasia Test* relating to the auditory comprehension of single words and sentences. Participants' repetition ability was taken as their score in the subtest of the *Comprehensive Aphasia Test* relating to the repetition of single words.

Note that in the present study, aphasia types reflect the profile of participants' linguistic capabilities. No claim is made about the site of their neuroanatomical lesions, as no imaging was undertaken in the current work.

4.5.2.5 Assessment of participants with aphasia

Participants with aphasia completed assessments over four to six sessions, each of which lasted no more than two hours. Each session took place at least one week after the previous session. Sessions began with the administration of half the items from one experimental task, followed by additional language assessments. Participants met with the researcher individually in a quiet room of their own home, or in the Department of Human Communication Sciences at the University of Sheffield.

*

The next chapter reports the results of the pilot investigation of the grammaticality judgement and sentence completion tasks with typical participants and participants with aphasia. Participants in Phase 2 are described in Chapter 6, and results of the main investigation follow in Chapters 7 and 8.

5 Pilot study for Phase 2

5.1 Introduction

This chapter reports an important stage of research practice, the pilot study. The chapter opens by examining the role of pilot studies in successful research and continues by recounting the pilot investigation of the grammaticality judgement and sentence completion tasks for Phase 2.

Porta (2014) described a pilot study as 'a small-scale test of the methods and procedures to be used on a larger scale if the pilot study demonstrates that these methods and procedures can work' (para. 1). The main purpose of a pilot study is to assess the feasibility of and identify necessary modifications to the main study, with regard to matters such as recruitment, documentation, informed consent procedures, methods of assessment and data collection tools (Leon, Davis & Kraemer, 2010). Pilot testing is an important component of the research process: pilot studies 'can…help to avoid doomed main studies' (Thabane et al., 2010, p. 9).

The objective of the current pilot study was to assess the delivery of the experimental grammaticality judgement task and sentence completion task to typical participants and participants with aphasia. In addition, the three screening tasks were piloted for participants with aphasia. Specifically, this pilot study aimed (1) to evaluate the use of the equipment, clarity of instructions and accuracy of recording that was planned for the delivery of the main study; (2) to investigate whether the screening tasks effectively indicated which participants with aphasia were suitable for the study; and (3) to examine several versions of stimuli presentation.

5.2 Method of pilot study

5.2.1 Participants

Ten typical participants and three participants with aphasia took part in the pilot study. Typical participants were recruited following the procedure described in Section 4.2.2. Participants with aphasia had all taken part in Phase 1 and included participants DS, EF and TP. The time between their participation in Phase 1 and the pilot study of Phase 2 was at least 14 months. This sample size was selected based on the feasibility of completing the pilot study within the timeframe of the project.

The group of typical participants contained seven women and three men. All reported normal or corrected-to-normal vision and hearing. Characteristics of this group are shown in Table 5.1 below. All participants were native speakers of British English, and some reported knowledge of a second language.

Table 5.1 Background of typical participants in Phase 2 pilot study

	M	*SD*	minimum	maximum
Age in years	67	6.2	61	81
Years in education	18	2.9	13	22

Note. Years in full-time education includes primary, secondary and higher education.

Table 5.2 Background of participants with aphasia at time of testing for Phase 2 pilot study

	DS	EF	TP
Age in years	71	62	71
Years in education	14	13	10
Time post-onset	11;2	4;0	11;4

Note. Years in full-time education includes primary, secondary and higher education; time post-onset shown in years;months format.

For comparison, the characteristics of the three participants with aphasia at the time of assessment are shown in Table 5.2. The three individuals presented with chronic mild-to-moderate non-fluent aphasia.

5.2.2 Versions of stimuli presentation

The grammaticality judgement task and the sentence completion task involved the presentation of a written sentence stimulus and the capture of participants' responses to subsequently presented verbs. One of the aims of this pilot study was to examine different versions of presenting the sentence and verb stimuli.

Three versions of stimuli presentation were piloted for each task. All three versions contained the six constructions and 24 verbs of interest in each task, described in Chapter 4. The versions differed in the number of unique stimulus sentences included in the tasks, and whether stimuli were blocked by argument structure construction. Piloting the three versions of stimuli presentation was motivated by the need to avoid creating tasks that were too difficult for participants with aphasia to complete, without sacrificing the difficulty of the tasks for typical participants - and thereby introducing floor and ceiling effects to the groups.

Two main considerations regarding participants with aphasia informed the decision to pilot three versions of stimuli presentation. First, participants with aphasia may have found it difficult to cope with sentence stimuli that changed on each trial, especially given that sentence stimuli were composed entirely of function words. It was hypothesised that participants with aphasia might incorrectly produce lexemes from previous sentence stimuli on a trial of the sentence completion task. If materials were not blocked by construction, this would lead to uncertainty as to whether a sentence production corresponded to the target construction. Second, a single presentation of a construction may not have been sufficient to activate its representation for participants with aphasia. As sentence processing in aphasia may be impaired in the temporal dimension (Berndt, 1998), participants with aphasia may require repeated exposure to the same construction in order to fully process it.

The three versions of stimuli presentation are described in the following sections.

5.2.2.1 Unblocked version of stimuli presentation

In the unblocked version of stimuli presentation, trials were not blocked by construction. Each verb was paired with a unique sentence stimulus. In addition to the conditions of trial order listed in Section 4.2.6, two other conditions were included for the unblocked version of stimuli presentation. First, at least two constructions intervened between stimulus sentences that instantiated the same construction. Second, adjacent trials did not contain stimulus sentences that began with the same subject pronoun.

In this version, the grammaticality judgement task contained 48 items, and the sentence completion task contained 24 items. The grammaticality judgement task began with 12 practice items, containing two instances of each of the six constructions, and the sentence completion task began with six practice items, containing one instance of each of the six constructions. This version of stimuli presentation is referred to as ‘unblocked’ throughout this chapter.

5.2.2.2 Multiple blocked version of stimuli presentation

In the multiple blocked version of stimuli presentation, trials were blocked by construction, and multiple forms of sentence stimuli were included in each block. Each verb was paired with a unique sentence stimulus, and all sentences that instantiated the same construction were presented consecutively. So, sentence stimuli were blocked by construction, but the form of each sentence stimulus changed on each trial. The order of verbs from the four verb groups differed within each block; that is, each block contained a different order of verbs from the groups [high cx, high lex], [high cx, low lex], [low cx, high lex] and [low cx, low lex]. Adjacent trials did not contain sentence stimuli that began with the same subject pronoun.

In the grammaticality judgement task, each block began with two trials that were not included in the analysis, one which elicited a ‘yes’ judgement and one which elicited a ‘no’ judgement, for a total of 60 trials in the task. In the sentence completion task, each block began with four trials that were not included in the analysis, for a total of 48 trials in the task. The grammaticality judgement task began with 12 practice items, including two verbs for each of the six constructions, and the sentence completion task began with six practice items, including one instance of each of the six constructions. This version of stimuli presentation is referred to as 'multiple blocked' throughout this chapter.

5.2.2.3 Single blocked version of stimuli presentation

In the single blocked version of stimuli presentation, trials were blocked by construction, and each block contained a single stimulus sentence. So, participants saw only six unique stimulus sentences over the course of either task.

The order of presentation of verbs was the same as for the multiple blocked version. The grammaticality judgement task contained 60 trials, and the sentence completion task contained 48 trials. Practice items for the grammaticality judgement task and sentence completion task included five verbs for each of two constructions, for a total of 10 practice items. This version of stimuli presentation is referred to as 'single blocked' throughout this chapter.

5.2.3 Pilot testing

5.2.3.1 Assessment of typical participants

Typical participants took part in the grammaticality judgement task and the sentence completion task. They also took part in the sentence production task described in Section 4.5.2.3, which was created as a novel assessment for Phase 2. Typical participants contributed normal data for the sentence production task during the pilot study.

The order of the grammaticality judgement task and sentence completion task was counterbalanced across typical participants. Each session concluded with the sentence production task. Most participants experienced a different version of stimuli presentation in each task, in order to allow the collection of qualitative feedback regarding how participants experienced different versions of stimuli presentation in comparison to one another. Typical participants were tested at the Department of Human Communication Sciences at the University of Sheffield, or in a quiet room of their own home.

5.2.3.2 Assessment of participants with aphasia

Participants with aphasia took part in the three screening tasks, the grammaticality judgement task, the sentence completion task and the anagram task, which was described in Section 4.5.2.3 as a novel assessment for Phase 2 designed to probe participants' comprehension of the argument structure constructions included in the study.

Assessment was completed in a single session that lasted no more than two hours. Each session began with the three screening tasks and concluded with the anagram task. The order of the grammaticality judgement task and sentence completion task was counterbalanced across participants. Participants experienced a different version of stimuli presentation in each task, in order to allow the collection of qualitative feedback on participants' experience with different versions of stimuli presentation. The researcher met with participants with aphasia in a quiet room of their own home.

The first participant with aphasia in the pilot study, DS, took part in the full version of each task. However, DS reported that the tasks were very tiring. Because participants with aphasia in the main study would not be asked to complete the grammaticality judgement, sentence completion and anagram tasks in the same assessment session, subsequent sessions with EF and TP contained only half the items from the grammaticality judgement and sentence completion tasks. This served to reduce participant fatigue during assessment in the pilot investigation.

5.3 Results from the pilot study

This section reports results from the pilot study. Results from typical participants and participants with aphasia are described as a case series, as it is inappropriate to use inferential statistics to evaluate a pilot study (Leon et al., 2011). Because the purpose of the pilot study was to assess the planned delivery of the tasks, and not investigate the research hypotheses, this section provides only a description of participants' performance.

5.3.1 Results from typical participants

This section presents results from typical participants in the pilot study. Table 5.3 shows the proportion of target responses[1] that each typical participant produced in the grammaticality judgement task, as well as the conditions to which their non-target responses were produced. Response time data for both tasks are shown in Figures 5.1 and 5.2. Participants are grouped by the version of stimuli presentation they experienced, either unblocked, multiple blocked or single blocked. Each bar in Figures 5.1 and 5.2 represents participants' mean response time to verbs in the same condition, i.e. [high cx, high lex], [high cx, low lex], [low cx, high lex] or [low cx, low lex] verbs. Bars corresponding to these verb groups are ordered from left to right. Response times to non-target responses were removed from the presentation of response time data.

Response times in Figure 5.2 for the sentence completion task represent the time between the presentation of the verb on screen and participants' production of the verb, or verb

---

[1] The 'proportion of target responses' may be described as 'accuracy'; however, the term 'proportion of target responses' was preferred to the term 'accuracy' for consistency with the usage-based approach to language adopted in this research, where 'grammaticality' can be said to arise from language usage rather than researchers' a priori assumptions.

Table 5.3 Proportion of target responses and number of non-target responses per condition produced by typical participants in grammaticality judgement task of pilot study

Stimuli presentation	Participant	Proportion of target responses	*N*	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
Unblocked	303	0.96	48	0	0	1	0
	305	0.90	48	0	0	0	0
	307	0.92	48	0	0	1	1
	310	0.98	48	0	0	1	0
	mean	0.94					
Multiple blocked	301	0.97	60	0	0	1	0
	306	0.90	60	0	0	0	0
	308	0.93	60	0	0	0	0
	mean	0.93					
Single blocked	302	0.92	60	0	0	0	0
	304	0.92	60	0	0	1	1
	309	1.00	60	0	0	0	0
	mean	0.95					

Note. Number of non-target responses per verb group shown in right-hand columns.

production latencies. These measurements were taken manually based on visual inspection of the waveform and spectrogram in Praat (Boersma & Weenink, 2017), because the procedure for the automatic recognition of verb production latencies described in Section 4.4.1.2 and used in the main study was in development at the time of the pilot study.

Grammaticality judgement task

As shown in Table 5.3, typical participants generally produced target responses in the grammaticality judgement task. All participants produced at least 90% of the target responses in the entire task. The majority of non-target responses were produced in response to verbs that were included in the task to elicit a 'no' judgement; that is, non-target responses generally represented the acceptance of a verb included in the task as ungrammatical. Typical participants produced at most two non-target responses in the set of 24 verbs of interest in the task. All non-target responses were produced in response to verbs with low construction frequency.

Figure 5.1 shows that all ten typical participants responded more quickly to verbs with high than low construction frequency. The unblocked version of stimuli presentation elicited the most diverse pattern of responses from typical participants, as 303 and 305 responded most quickly to verbs with high construction frequency but showed little difference in responses to verbs that differed in lexical frequency. 307 responded more quickly to verbs with high construction and lexical frequency, while 310 responded most slowly to the verbs in the

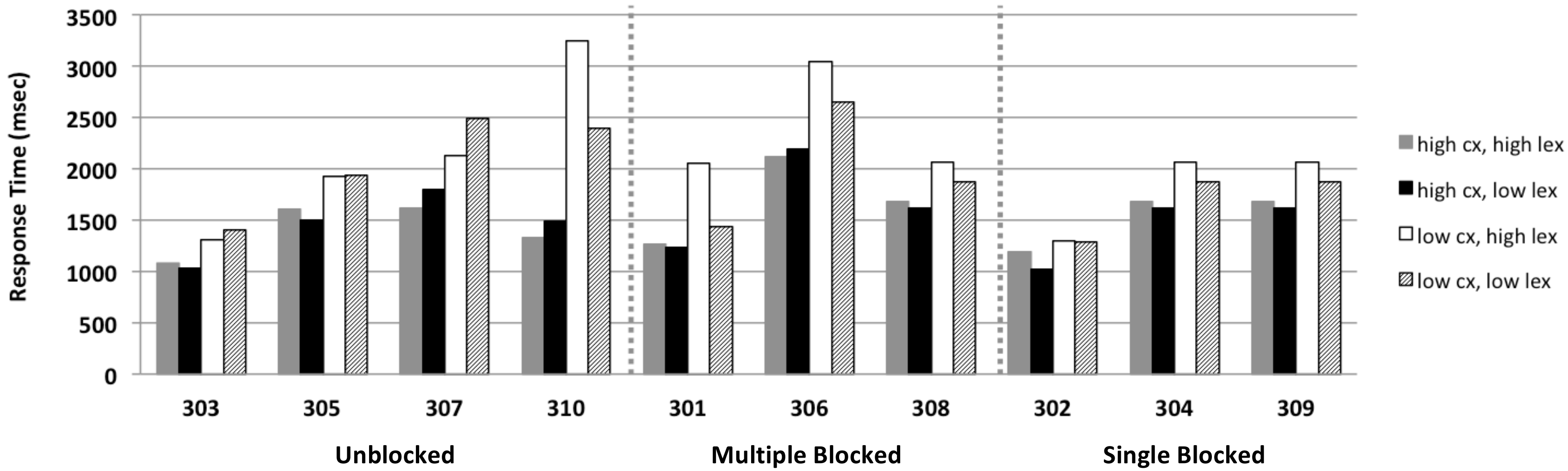
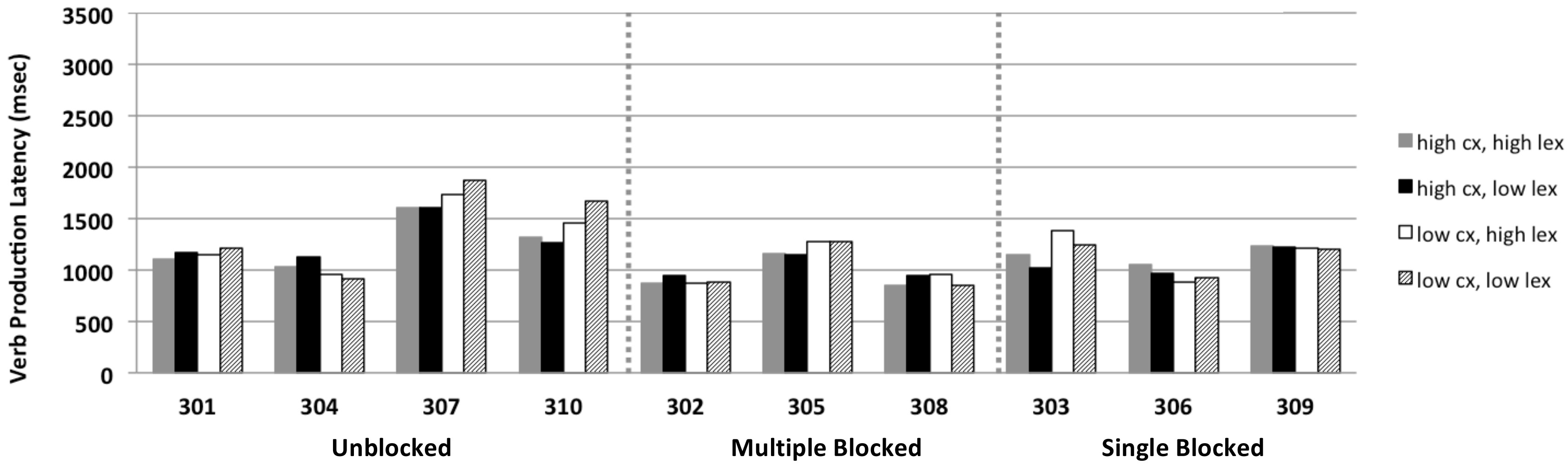


Figure 5.1 (above) Mean response times from typical participants to verbs in each condition in pilot grammaticality judgment task

Figure 5.2 (below) Mean verb production latencies from typical participants of verbs in each condition in pilot sentence completion task



[low cx, high lex] verb group. The blocked versions of stimuli presentation elicited one main pattern of responses from typical participants: response times showed little difference between high and low lexical frequency for verbs with high construction frequency and the longest response times to verbs in the [low cx, high lex] verb group.

Sentence completion task

Typical participants produced so few non-target responses in the sentence completion task that this data was uninformative. Only one typical participant produced a non-target response to a verb of interest in the study, mistaking the word 'there' for 'here' on the removal construction and producing a self-correction in the response 'they reach it from /h/ from there'.

As shown in Figure 5.2, response times to verbs in the sentence completion task were more variable than in the grammaticality judgement task. In the unblocked version of stimuli presentation, 301 and 307 responded most quickly to verbs with high construction and high lexical frequency; 304 appeared to show a reverse effect of construction frequency; and 310 appeared to show an effect of construction frequency, and an effect of lexical frequency at the level of low construction frequency. There were fewer clear patterns of responses for the blocked versions of stimuli presentation: only 303 and 305 responded more quickly to verbs with high construction frequency at both levels of lexical frequency.

Power

The power calculations reported in this section were performed after Phase 2 data collection, and these estimates did not inform the decision to recruit samples in Phase 2. Using the estimates of means and standard deviations from the 10 typical participants in the pilot study, a sample size of 27 participants would be necessary to reach 80% power to detect an effect of construction frequency in the grammaticality judgment task, at an alpha-level of 0.05. A sample size of approximately 1700 participants would be necessary to detect an effect of lexical frequency in the grammaticality judgement task.

Using estimates from the pilot study, a sample size of approximately 1500 typical participants would be necessary to reach 80% power to detect an effect of construction frequency in the sentence completion task, at an alpha-level of 0.05. A sample size of over 4000 participants would be necessary to detect an effect of lexical frequency.

As described in Section 5.1, the main purpose of a pilot study is to trial the procedures associated with data collection (Leon et al., 2010), and that was the first aim of the current pilot study. For this reason, the data generated in the main study of Phase 2 can be considered more appropriate to serve as the basis for a power analysis to inform future research, and these power analyses are reported in Chapters 7 and 8.

5.3.2 Results from participants with aphasia

This section reports results from participants with aphasia. The proportion of target responses that participants produced in each task is shown in Table 5.4, along with the number of non-target responses in each verb group. Response time data are shown in Figure 5.3 (over page). Times to non-target responses were removed from the presentation of

Table 5.4 Proportion of target responses and number of non-target responses per verb group from participants with aphasia in Phase 2 pilot study

Stimuli presentation	Participant	*N*	Proportion of target responses	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
Grammaticality judgement task							
Unblocked	EF	32	0.94	0	0	1	0
Multiple blocked	DS	60	0.82	0	1	1	1
Single blocked	TP	30	0.67	0	1	0	0
Sentence completion task							
Unblocked	DS	24	0.39	4	4	2	2
Multiple blocked	TP	18	0.47	2	2	2	2
Single blocked	EF	24	0.83	0	0	0	0

Note. Version of stimuli presentation; participant; total number of trials in task; proportion of target responses; number of non-target responses per verb group shown in right-hand columns.

response time data. Note that participants experienced a different number of items in each task, based on the version of stimuli presentation they received. Results will be discussed by participant.

DS

DS took part in the full version of both tasks. In the grammaticality judgement task, he produced three errors to the verbs of interest and was slowest to respond to verbs from the [low cx, low lex] verb group. His errors in the sentence completion task included production of non-target pronouns, non-target verbs and non-target constructions. At the beginning of the sentence completion task, DS perseverated on the verb *want*, which was included in the practice trials. His verb production latencies were longest to verbs from the [low cx, low lex] verb group.

EF

Of the three participants with aphasia, EF achieved the highest proportion of target responses in both tasks. In the grammaticality judgement task, she scored within the normal range in the production of target responses. In the sentence completion task, EF's verb production latencies were the shortest of the three participants with aphasia and comparable to those of typical participants. Her errors included the repetition of a subject pronoun and the production of non-target pronouns, but she made no errors on the verbs of interest in the task. Her verb production latencies were longest to verbs from the [low cx, low lex] verb group.

TP

TP produced the lowest number of target responses in the grammaticality judgement task; however, he produced a non-target response to only one verb of interest in the pilot study. His performance was characterised by difficulty in rejecting ungrammatical sentences. He was the fastest participant with aphasia to respond in the grammaticality judgement task, and his response times were similar to those of typical participants. In the sentence completion task, TP produced the target sentence in response to only three of the twelve verbs of interest that he saw in the task. His errors included the production of non-target pronouns, such as his response 'I shout at you' for the target sentence 'you shout at us'. These errors were made in response to every verb group. TP was unable to produce a response on three trials, and these trials all contained verbs with low construction frequency. TP responded most slowly to [low cx, low lex] verbs in both tasks.

Power

Because only three participants with aphasia took part in the pilot study for Phase 2, a power analysis of this data would be uninformative. Analysis and commentary on the power of the Phase 2 study on data from adults with aphasia are provided in the reporting of Phase 2 results in Chapters 7 and 8.

5.3.3 Summary of results from pilot study

Results from the pilot study demonstrated that typical participants and participants with aphasia were able to cope with the demands of the novel grammaticality judgement and sentence completion tasks.

The three participants with aphasia produced few non-target responses to the verbs of interest in the grammaticality judgement task, and the types of errors they produced in the sentence completion task informed the selection of the version of stimuli presentation used in the main study. Results on participants' performance in the pilot study should be interpreted with caution due to the low number of items that participants experienced, but the response time data show that participants with aphasia generally responded most slowly to verbs from the [low cx, low lex] verb group in each task, providing some preliminary evidence that participants with aphasia may be sensitive to effect of construction frequency during sentence processing.

5.4 Changes to Phase 2 method

As a result of the pilot study, a number of revisions was made to the materials and method for the grammaticality judgement task and the sentence completion task, and the assessment of participants with aphasia in Phase 2. This section reports those changes and the rationale behind them.

5.4.1 Stimuli presentation

One of the objectives of this pilot study was to examine different versions of stimuli presentation. Based on the results described above, it was decided that typical participants would experience the unblocked version of stimuli presentation for the grammaticality judgement and sentence completion tasks. This version of stimuli presentation elicited the most variable patterns of participant responses in the grammaticality judgement task and the most consistent differences with respect to construction frequency in the sentence completion task.

The blocked versions of stimuli presentation may have influenced language processing in typical participants in ways problematic for the study. In blocked versions, participants could anticipate the form of each construction before it appeared, because it was the same as the preceding sentence stimulus. Therefore, sentence stimuli did not serve as new contexts for verbs on each trial. Participants could also easily compare verbs within each block to one another. In the blocked versions, the four verbs of interest were presented consecutively within each block, so participants may have based their responses to verbs in the grammaticality judgement task in comparison to one another, rather than as a function of the sentence stimulus. Indeed, many participants reported that some verbs felt more natural in the sentences than others. This feeling may have been exacerbated in the blocked versions.

It was also decided that participants with aphasia would experience the single blocked version of stimuli presentation. In the sentence completion task, all three participants with aphasia at times produced a subject pronoun different from the pronoun shown on screen. Though this type of response was not counted as an error, some participants self-corrected, and this affected the measurement of the verb production latency. For example, DS responded ‘I no we owe them something’ in the unblocked version of the sentence completion task. The correction of the initial pronoun made his verb production latency data unusable.

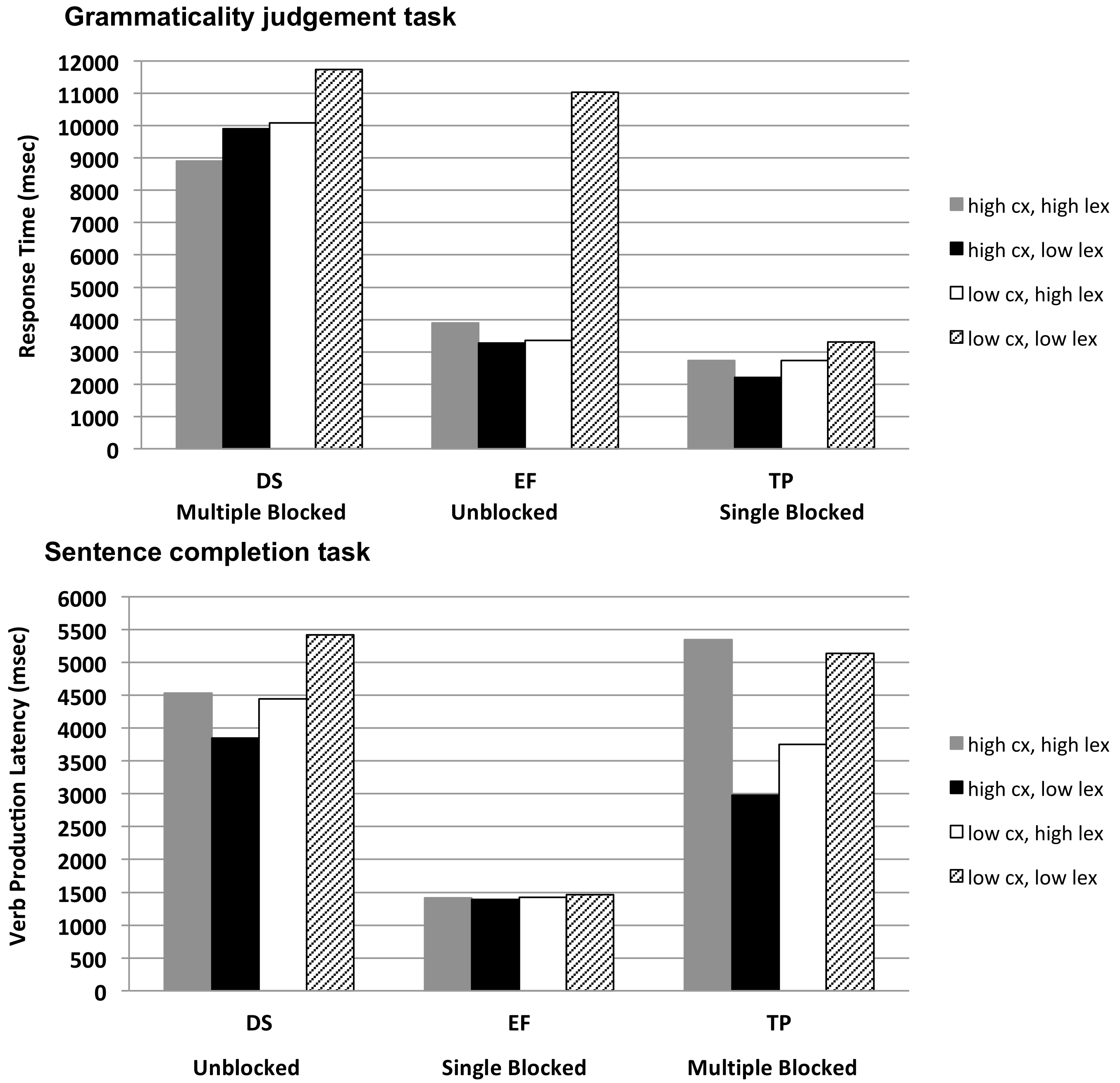


Figure 5.3 Median response times from participants with aphasia to tasks in Phase 2 pilot study

Note. Results from grammaticality judgement task shown in upper pane and results from sentence completion task in lower pane; scales of y-axes differ between the upper and lower panes.

Additionally, the single blocked version allowed participants with aphasia to practise each sentence stimulus before experiencing the verbs of interest in the study. The single blocked version of the sentence completion task in the pilot study contained eight verbs per construction: the first four verbs were not included in the analysis, and the final four were the verbs of interest. EF experienced the single blocked version of the sentence completion task and did not make any errors on the four verbs of interest. For the main study of Phase 2, the number of practice items in each block was reduced to two, in order to reduce practice effects. So, within each block, participants with aphasia experienced six verbs in the main study: the first two were practice items, and the final four were the verbs of interest in the study.

5.4.2 Materials

Several minor changes were made to the materials for Phase 2. The combination of sentence stimuli and verbs was changed if the subject pronoun in the sentence stimulus and the verb began with the same phoneme, in order to reduce phonological similarity between the sentence stimulus and verb. This affected the verb *win* in the grammaticality judgement task, which was originally paired with the sentence stimulus *we ___ it from him*, and the item *walk* in the sentence completion task, which was originally paired with the sentence stimulus *we ___ through there*. The change was implemented by exchanging the sentence stimulus with one originally paired with another verb included for the same construction, resulting in the sentences *you win it from her* and *you walk through there* for participants with aphasia.

Lists were also checked to ensure that adjacent verbs did not share a phonological or orthographic rhyme. This resulted in changing the order of *give* and *drive* in the sentence completion task so the verbs were not presented subsequently to one another.

5.4.3 List orders

Several unique list orders were created for Phase 2, in order to control for practice effects during participants' completion of the tasks. For typical participants, four unique list orders were generated for each task. Practice items were checked to ensure that their order did not violate a condition of the formation of the lists, as described in Section 4.2.6. The order of the grammaticality judgement task and the sentence completion task was counterbalanced for typical participants.

In the main study, participants with aphasia experienced half the items from the grammaticality judgement task or sentence completion task per assessment session. In order to ensure that participants' first experience with the verbs of interest was in the grammaticality judgement task or the sentence completion task, rather than on another language assessment included in the project, all participants with aphasia experienced the same order of tasks; that is, the order of the grammaticality judgement task and sentence completion task was not counterbalanced for participants with aphasia. However, the order in which constructions were presented within each assessment session varied for participants with aphasia. To illustrate, all participants with aphasia saw the conative, ditransitive and intransitive motion constructions in the grammaticality judgement task in the first assessment session, but the order in which participants saw the three constructions varied by participant, in order to reduce a practice effect in group-level data.

5.4.4 Grammaticality judgement task

In the pilot grammaticality judgement task, participants chose which hand to use to perform button press responses, because handedness can affect the timings of motor responses (Goodin et al., 1996). The participants with aphasia all responded with their premorbid non-dominant hand, due to right-side hemiplegia as a result of stroke. In order to ensure the performance between typical participants and participants with aphasia was comparable in the main study, typical participants were asked to respond with their non-dominant hand. This necessitated positioning the button box in the same way for all participants, rather than letting participants choose the orientation of the box, as they had in the pilot study. The green 'yes' button was positioned on the right and the red 'no' button on the left for all participants in the main study.

For participants with aphasia, a visual reminder was added to the computer screen when 'yes' or 'no' responses were required in the grammaticality judgement task. DS mentioned that the task was confusing with three buttons, although all three participants with aphasia in the pilot study were able to meet the demands of the task after the practice trials. In order to ease the memory demands of the task for participants with aphasia, green- and red-coloured boxes were added to the computer screen as a reminder to press a green or red button when required. The coloured boxes provided a visual cue that the necessary response at that point in time was the green or red button. To avoid visual similarity between the verb, which was presented in red, and the red button reminder, the colour of the verb was changed to blue.

5.4.5 Sentence completion task

Instructions for typical participants were added to the sentence completion task to request participants to read each sentence aloud as naturally as possible. In the pilot study, some typical participants read sentences aloud with a measured prosody, almost as if reading each word aloud in isolation. This type of production could indicate automatic or sublexical processing. Therefore, an explicit instruction regarding how to read sentences aloud was added to the task.

A modification to the trial design was implemented to rectify the high number of technical failures of the audio recordings that occurred during the pilot study. The success rate for capturing complete audio recordings from typical participants ranged from 29% to 96% per participant (mean 72%). That is, the microphone failed to record, or terminated the recording of a response early, on an average of 28% of trials per participant. This fail rate was deemed unacceptably high.

The fault in capturing audio recordings was isolated as the button press used to terminate the microphone at the end of each trial. Recordings proved to be complete if the microphone was programmed for a set duration of time. Therefore, the sentence completion task for typical participants was changed so that participants were given three seconds to produce vocal responses. This time was sufficient to capture most responses produced in the pilot study, and it seemed feasible that typical participants could manage to produce responses within that timeframe. After the microphone recorded for three seconds, the experiment moved on to the next trial automatically.

For participants with aphasia, the programming of the microphone had a larger impact. TP required up to ten seconds to produce complete sentences aloud. This contrasted with the performance of EF, who performed similarly to typical participants. Therefore, participants with aphasia took part in a short block of practice trials during the assessment session prior to the sentence completion task, and the microphone was programmed for the maximum duration of their responses in the practice trials.

In addition, fixation crosses were added to the tasks for typical participants and participants with aphasia, in order to direct participants' attention to upcoming stimuli.

5.4.6 Assessment of participants with aphasia

5.4.6.1 Screening tasks

One of the objectives of this pilot study was to evaluate the success of the screening tasks in indicating whether participants with aphasia would be suitable for inclusion in the main study. DS, EF and TP all responded to one item in the grammaticality judgement screening task inaccurately. All three participants with aphasia accepted the sentence 'the boy was carrying' as 'correct'. However, Kim and Thompson's (2000) target response for this item was 'incorrect', because the verb *carry* usually occurs with a direct object. In response, this item in the grammaticality judgement screening task was replaced with an item from Kim and Thompson (2000) that contained an extra argument, rather than a missing argument.

5.4.6.2 Assessment schedule

In response to feedback from participants with aphasia that the grammaticality judgement and sentence completion tasks were very difficult, the assessment schedule for Phase 2 was modified as a result of the pilot study. The original plan involved participants with aphasia taking part in the full version of the grammaticality judgement task in one assessment session and full version of the sentence completion task in the next. In order to reduce participant fatigue, the assessment schedule was modified so that participants with aphasia experienced half of the constructions from one task in each session; that is, participants with aphasia took part in three constructions from the grammaticality judgement or sentence completion task per assessment session.

5.4.6.3 Testing environment

Finally, the conditions for testing participants in their own homes was clarified for all participants. In one instance during the pilot study in a participant's home, a dog continually interrupted the meeting, and a phone rang during one of the computerised tasks, which affected response latencies. In another instance, a participant with aphasia was sitting in an armchair and had the television on when the researcher arrived. As a result, participants who wished to participate in the project in their own homes were advised that the researcher would need to meet with them in a quiet environment, free from noise like pets or the television. Participants were also told that they would need to sit with the researcher at a large table.

*

This pilot study resulted in improvements to the design of Phase 2. Some improvements pertained to data collection procedures, such as increased accuracy in the recording of participants' responses in the sentence completion task, and others resulted in a better experience for participants, such as advance notice of the necessary testing environment.

The participants who took part in Phase 2 are described next in Chapter 6. Results from the Phase 2 grammaticality judgement task follow in Chapter 7 and results from the sentence completion task in Chapter 8.

6 Participants in Phase 2

This chapter describes the participants who took part in Phase 2. Typical participants are introduced first, and the remaining sections report results from participants with aphasia in the Comprehensive Aphasia Test, other assessments of verb and sentence processing and screening tasks.

6.1 Typical participants in Phase 2

All typical participants in Phase 2 were native speakers of British English. Participants reported normal or corrected-to-normal vision and hearing. Some reported knowledge of a language in addition to English. No participant reported any history of speech or language difficulties or psychiatric impairment. A total of 90 typical participants were assessed in Phase 2. Data from three participants were excluded from the study because they revealed a history of stammering during assessment. Therefore data from 87 typical participants were entered into analyses in Phase 2. Information on the background of typical participants is shown in Table 6.1.

Table 6.1 Background information on typical participants in Phase 2

	M	*SD*	Minimum	Maximum
Age in years	66	6.7	50	80
Years in education	16	3.2	10	23

Note. Years in education includes full-time study in primary, secondary and higher education.

The group of typical participants included 43 women and 44 men. Most typical participants reported professional backgrounds, such as nursing, teaching, accountancy and managerial or administrative positions.

6.2 Background information on participants with aphasia in Phase 2

A total of 14 adults with acquired aphasia took part in Phase 2. Background information about these participants is shown in Table 6.2. Results were obtained from self-report measures based on biographical interviews between the researcher and participants. All participants received a diagnosis of acquired aphasia from a qualified speech and language therapist and incurred aphasia as a result of a CVA.

Information about ethical approval, informed consent and recruitment for participants in Phase 2 was discussed in Section 4.2.2.

Table 6.2 Background information on participants with aphasia in Phase 2

Participant	Age (years)	Gender	Pre-morbid handedness	Years in education	Previous employment
BF	56	male	right	21	Professional
CJ	60	male	right	10	Manual
GW	84	male	right	16	Education
HM	35	female	right	18	Education
IC	70	male	left	20	Academic
JF	73	male	right	11	Professional
KT	67	female	right	11	Clerical
MJ	58	male	right	16	Manual
PD	64	female	right	10	Clerical
RE	62	female	right	10	Retail
SP	85	male	right	9	Manual
UT	55	male	right	15	Professional
VH	61	male	right	16	Finance
WD	75	male	right	13	Retail

Note. Years in education includes full-time study in primary, secondary and higher education.

6.3 Results from the Comprehensive Aphasia Test

Three components of the *Comprehensive Aphasia Test* (CAT) (Swinburn et al., 2004) were administered in Phase 2, including (1) a cognitive screen, (2) assessments of language comprehension and (3) assessments of expressive language. Participants VH and WD had severe apraxia and did not take part in any assessments of expressive language. Scoring followed the guidance on the administration of the *Comprehensive Aphasia Test* provided in the test manual (Swinburn et al., 2004). This section describes the tasks and scoring procedures for the subtests of the CAT, including how impaired performance was identified. Participants' scores follow in Table 6.3 (over page). Throughout this and the following chapters, participants with aphasia in Phase 2 are listed in order of their score in the written sentence comprehension subtest of the CAT.

6.3.1 Cognitive screen

The cognitive screen contained six subtests. The line bisection subtest screened for visual neglect, and scores represent the degree of deviation in participants' identification of the centre of printed lines from the actual centre. The semantic memory subtest is a nonverbal test of semantic memory, and participants were asked to identify which of four images was related to a target image, for example, selecting *mitten* as related to *hand*. In the recognition memory subtest, participants were asked to identify which of four images they had previously witnessed in the cognitive screen. These images had been included in the semantic memory subtest. In the word fluency subtest, participants named items in response to the categories *animals* and the letter *S*. The response time to each category was one minute. The subtest on gesture object use screened for apraxia and required participants to produce the gesture that was associated with six objects, such as pinching fingers on an imaginary clothes line in response to an image of a clothes pin. Finally, the arithmetic subtest screened for acalculia and contained six sums, covering the operations of addition, subtraction and multiplication.

For the semantic memory, recognition memory and arithmetic subtests in the cognitive screen, the scoring procedure awarded one point to each correct response. Scores in the word fluency subtest represent the total number of unique words that participants produced in response to both categories. In the gesture object use subtest, two points were awarded to a correct response.

6.3.2 Language comprehension

The assessment of language comprehension encompassed both auditory and visual comprehension. Comprehension subtests at the word and sentence levels involved participants selecting one of four images that matched the target. The scoring procedure awarded two points to correct and timely responses, and one point for responses delayed over five seconds, self-corrections by the participant or repetitions by the researcher. No points were awarded for incorrect responses. For this reason, the number of items in each subtest is distinguished from the maximum score in Table 6.3 (over page).

In the assessment of auditory comprehension of paragraphs, the researcher read aloud two short paragraphs and asked participants four yes-no questions about each. Questions were paired such that two questions referred to the same information, and participants had to respond to each part of the question correctly to gain a point.

6.3.3 Expressive language

Assessments of expressive language included repetition, picture naming, reading aloud and picture description. Repetition subtests included the repetition of single words, morphologically complex words and non-words. A maximum of two points was awarded to each correct response, in the manner described above, and phonemic or dyspraxic errors were not accepted. Repetition of digit strings assessed short-term memory, and scores in this subtest represent the maximum number of digits that participants could correctly repeat. Similarly, scores in the sentence repetition subtest indicated participants' sentence span in terms of the number of content words contained in the longest sentence that participants correctly repeated. Assessments of picture naming included subtests on object naming and action naming. Assessments of reading aloud included subtests on single words, morphologically complex words, function words and non-words. As before, a maximum of two points was awarded for each correct response.

Single words	16	32	30-32	32	**23**	**29**	32	**18**	**24**	32	**19**	**25**	30	**22**	30	-	-
Complex words	3	6	6	6	**0**	6	6	**0**	6	6	**4**	**3**	6	6	6	-	-
Non-words	5	10	4-10	10	**1**	8	6	**0**	6	10	6	8	7	**5**	8	-	-
Digit span		7	5-7	5	**3**	6	5	**3**	6	5	5	5	**3**	5	**2**	-	-
Sentence span		6	6	**5**	**4**	6	6	**3**	6	**5**	**5**	6	**3**	**4**	**0**	-	-
Picture naming																	
Object naming	24	48	42-48	46	**30**	44	48	**23**	44	45	**31**	**41**	46	**41**	**35**	-	-
Action naming	5	10	8-10	9	**6**	9	**8**	**2**	**8**	10	**8**	10	**8**	**2**	**2**	-	-
Reading aloud																	
Single words	24	48	44-48	46	**31**	47	**44**	**31**	**44**	46	**36**	**32**	46	48	**25**	-	-
Complex words	3	6	5-6	6	**2**	5	6	**0**	6	6	**2**	**3**	6	6	**0**	-	-
Function words	3	6	4-6	6	4	4	6	6	6	6	6	6	6	6	**2**	-	-
Non-words	5	10	6-10	**2**	**0**	**4**	**6**	**4**	9	10	**2**	10	10	7	**0**	-	-
Picture description																	
Appropriate items			21-58+	29	41	38	**17**	29	21	32	25	**12**	27	**17**	22	-	-
Inappropriate items			0-1	**10**	1	**3**	0	**7**	1	**5**	**4**	0	**3**	**7**	1	-	-
Syntactic variety			4-6	6	5	5	4	6	6	6	4	4	4	**2.5**	5	-	-
Well-formedness			5-6	6	6	5	5	6	**4**	5	**3**	**2**	5	**4**	**4**	-	-
Speed			3	**2**	**2**	**2**	**1**	3	**1.5**	**2.5**	**1.5**	**2**	**2**	3	**1.5**	-	-

Note. Number of items in each task (*N*); maximum possible score (Max score);

Table 6.3 Scores from participants with aphasia in the Comprehensive Aphasia Test

	N	Max score	Normal range	BF	CJ	GW	HM	IC	JF	KT	MJ	PD	RE	SP	UT	VH	WD
Cognitive screen																	
Line bisection	3	±6	-3.5 - 0.5	-2	**3.5**	-1	**-2.5**	-1	-1	0.5	**-2.5**	2	-2	-1.5	-1.5	-1.5	**-5**
Semantic memory	10	10	8-10	10	9	10	10	10	9	10	10	9	10	9	10	**8**	**5**
Recognition memory	10	10	8-10	9	**5**	10	10	10	10	10	9	**8**	9	10	10	10	10
Word fluency	2		12-56	20	29	**11**	25	20	**7**	**9**	14	20	**8**	**1**	**8**	**0**	**0**
Gesture object use	6	12	8-12	11	**9**	12	10	12	12	12	11	10	11	11	**9**	12	**4**
Arithmetic	6	6	1-6	3	6	6	6	6	6	6	4	6	3	3	5	**0**	2
Language comprehension																	
Auditory comprehension																	
Words	15	30	25-30	29	29	29	30	27	29	**25**	**25**	**25**	27	**18**	27	**21**	**15**
Sentences	16	32	26-32	29	**22**	28	28	28	30	**0**	**23**	28	**25**	**22**	**21**	**16**	**8**
Paragraphs	8	4	3-4	4	**1**	4	4	4	4	**1**	3	4	4	**2**	3	**2**	3
Visual comprehension																	
Words	15	30	27-30	28	28	30	30	30	**27**	30	**23**	**22**	**26**	30	**24**	**19**	**10**
Sentences	16	32	24-32	30	28	28	27	26	26	24	**22**	**22**	**21**	**20**	**18**	**12**	**5**
Expressive language																	
Repetition																	

scores in bold indicate impaired performance; (-) indicates not assessed.

6.3.4 Picture description

The final expressive language subtest reported in Table 6.3 is a spoken picture description task. Scoring of this task followed the procedure set out in Swinburn et al. (2004) and covered five dimensions of participants' responses. Appropriate lexical items refers to the number of items that were produced in the correct context, with no phonological errors. Inappropriate lexical items refers to the number of productions that were semantically inappropriate, or incorrectly produced in terms of phoneme selection.

Syntactic variety is a score on a scale from zero to six that reflects the diversity of syntactic structures included in a participant's response, such as noun phrases, verb phrases and embedded clauses. A score of zero indicates a total lack of syntactic structure, and a score of six indicates use of the full range of syntactic structures. Grammatical well-formedness is a score on a scale from zero to six that reflects the accuracy of grammatical realisations, such as inflectional morphemes and arguments. A score of zero indicates that no phrases were well-formed, and a score of six indicates that all phrases were well-formed and none, such as auxiliaries, tense marking or verb arguments, was missing. Finally, speed is a score on a scale from zero to three that reflects the rate of speech production. A score of zero indicates significant delay, and a score of three indicates a normal speech rate.

6.3.5 Identification of impaired performance

Scores in Table 6.3 that indicate impaired performance are shown in bold. Swinburn et al.'s (2004) published cut-off scores for impaired performance equate to the score that at least 95% of the typical population exceeded. For this reason, the lowest score in the normal range is the same or lower than the cut-off score indicating impaired performance for some tasks, including subtests of line bisection, semantic memory, recognition memory, gesture object use, auditory comprehension of words, visual comprehension of words, repetition of non-words, action naming, reading aloud of single words and reading aloud of non-words.

6.4 Aphasia diagnoses

Details relating to participants' diagnoses of aphasia are shown in Table 6.4, including the time post-onset at the first point of assessment in Phase 2, severity of expressive language, fluency of spontaneous language production, aphasia syndrome and details of participants' experiences with speech and language therapy.

This section continues with an explanation of how aphasia severity, fluency and aphasia syndromes were determined, and describes the participants in Phase 2 whose language abilities can be classified as within the various aphasia syndromes.

Table 6.4 Characteristics of aphasia for participants in Phase 2

	Time post-onset	Severity of expressive language	Fluency	Aphasia syndrome	Previous speech and language therapy
BF	1;7	4	Fluent	Anomic	Therapy in hospital and 9 months following stroke
CJ	5;6	5	Fluent	Wernicke's	Four months following hospital discharge
GW	2;6	5	Fluent	Anomic	Therapy in hospital and six months after discharge
HM	6;11	4	Fluent	Anomic	Therapy in hospital and one year following discharge
IC	2;3	4	Fluent	Conduction	2-3 days per week for three months following CVA
JF	7;10	4	Fluent	Anomic	Therapy in hospital and three months after discharge
KT	3;2	4	Fluent	Transcortical sensory	No health service therapy
MJ	15;4	3	Non-fluent	Broca's	Therapy for one year following CVA
PD	6;8	4	Fluent	Conduction	Therapy in hospital and one year following discharge
RE	2;8	4	Fluent	Transcortical sensory	One year following hospital discharge
SP	3;6	3	Fluent	Wernicke's	Some therapy after hospital discharge
UT	10;0	3	Fluent	Transcortical sensory	Therapy in hospital and 18 months in community
VH	5;10	1	Non-fluent	Global	Several months following CVA
WD	1;0	1	Fluent	Wernicke's	Therapy following hospital discharge

Note. Time post-onset shown in years;months format; severity of expressive language and fluency based on Boston Diagnostic Aphasia Test (Goodglass & Kaplan, 1983); syndrome classifications based on Davis (1993). Alternating rows in grey for ease of reference.

6.4.1 Severity

Table 6.4 reports the severity of participants' expressive language impairment. The Aphasia Severity Rating Scale from the Boston Diagnostic Aphasia Test (Goodglass & Kaplan, 1983) was used to rate the communication of participants with aphasia in spontaneous conversation with the researcher. The scale ranges from 1 to 5, with 1 representing most severely impaired individuals and 5 representing least severely impaired individuals. A summary of the scale is shown in Table 6.5 (over page).

Table 6.5 Summary of Aphasia Severity Rating Scale (Goodglass & Kaplan, 1983)

Rating	Summary
1	Communication through fragmentary expression; great need for inferencing by listener
2	Conversation about familiar topics possible, with assistance from listener
3	Conversation about almost all topics possible, with little assistance from listener
4	Some obvious loss of ability, but with no significant limitation on conversation
5	Difficulties not obvious to listener

The aphasia severity for most participants in Phase 2 was rated as 4, indicating a mild form of aphasia that did not form a significant barrier to communication. This rating scale, however, is based on a judgement of participants' general performance in conversation and is not informative regarding their psycholinguistic capabilities. For example, both CJ and GW were rated as a 5, indicating their language difficulties were not obvious in a conversation setting, but CJ showed impairments in auditory comprehension, naming and reading aloud not shared by GW. Participants will be discussed in more detail regarding their performance in specific linguistic processing abilities relevant to the Phase 2 experimental tasks in Section 6.4.3 below.

6.4.2 Fluency

The distinction between fluent and non-fluent aphasia was derived using the Profile of Speech Characteristics included in the Boston Diagnostic Aphasia Test manual (Goodglass & Kaplan, 1983). This profile is based on six dimensions of speech output, including intonation, phrase length, articulation, grammatical variety, verbal paraphasias and word finding. Each dimension in the profile is associated with a seven-point scale on which the researcher rates the performance of the participant from minimum to maximum abnormality. Ratings were based on recorded samples of participants' conversational speech from assessment sessions.

6.4.3 Aphasia syndromes and participant profiles

The classification of aphasia syndromes was based on participants' performance in the *Comprehensive Aphasia Test*, shown in Table 6.3. Classifications followed the system provided in Davis (1993), which is based on participants' fluency, auditory comprehension and repetition ability, in association with descriptions of typical presentations of each syndrome provided by Goodglass and Kaplan (1983) and Damasio (1998). The following sections provide a summary of the aphasia syndromes as detailed by those authors and described in Section 2.1.1.1. The following sections also contain profiles of the participants in Phase 2 to which the syndrome label applies. Each participant will be discussed with particular reference to impairments of semantic processing, written word comprehension, written sentence comprehension and spoken production, as these abilities underpin performance in the Phase 2 grammaticality judgement and sentence completion tasks.

Anomic aphasia

Anomic aphasia refers to a language impairment characterised by fluent language production in the context of little impairment to auditory comprehension. Individuals with anomic aphasia produce grammatically well-formed language, and word-finding difficulties tend to be the most noticeable attribute of this type of aphasia. Four participants in Phase 2 were diagnosed with anomic aphasia, and all participants in this group obtained an aphasia severity rating of 4 or 5, indicating the presence of a mild expressive language impairment. These participants showed little evidence of semantic processing deficits and performed within the typical range in CAT subtests of auditory and visual comprehension.

BF demonstrated the most proficient language capabilities of all participants with aphasia in Phase 2. He made very few errors overall in the *Comprehensive Aphasia Test*, and the only scores that indicated impairment included subtests of sentence repetition and non-word reading aloud. He was able to produce a typical amount of spoken language output in the picture description task.

GW also presented with very mild anomic aphasia. His only scores in the *Comprehensive Aphasia Test* that were classified as impaired included single-word repetition and non-word reading aloud. Like BF, he produced a typical amount of spoken output in the picture description task.

HM’s speech output was characterised by a slow rate of speech with pauses between words and clauses, and she demonstrated relatively little language in the picture description task. She scored in the impaired range in the CAT action-naming subtest, because she described the target ‘winding’ as ‘telling the time’.

JF’s speech output showed atypical articulation of complex onsets, in words like *crab*, and of the approximants /r/ and /w/. He performed outside the normal range in the subtest of written word comprehension due to three instances of a delayed response.

Conduction aphasia

Conduction aphasia is characterised primarily by phonemic paraphasias in language production, which are especially prominent in repetition tasks, in the context of well-preserved comprehension of spoken language. Two participants in Phase 2 were diagnosed with conduction aphasia, IC and PD.

IC presented as a typical case of conduction aphasia. His spoken output was fluent but punctuated by frequent phonemic paraphasias and instances of *conduit d'approche*. He had no deficits in language comprehension. His low score in the object-naming subtest was attributable to phonemic errors consistent with conduction aphasia, but IC was unable to approximate the target in the action-naming subtest for three of the five items. In these instances, he described target actions, for example, labelling the target *sawing* as 'cutting a bit of wood'.

PD had a mild form of conduction aphasia, as her spoken output was marked by the repetition of phonemes, especially at the beginning of words. She also showed occasional word-finding difficulties in spontaneous speech. PD showed impaired comprehension of written words and written sentences, and she demonstrated a mixed profile of errors in these tasks. Notably, PD's responses in subtests of reading aloud were indicative of surface dyslexia. The target *yacht* first elicited the response /jætʃt/, to which PD commented 'no, we don't say that' and then corrected her response to /jɑt/. Her reading aloud of non-words was more fluent than the reading aloud of real words, indicating preserved reading ability via letter-to-sound correspondences.

Transcortical sensory aphasia

Transcortical sensory aphasia describes fluent language output, with the relative preservation of repetition skills in comparison to auditory comprehension. Though transcortical sensory aphasia usually refers to the preservation of repetition in the context of 'a *severe* Wernicke's aphasia' (Goodglass & Kaplan, 1983, p. 91; emphasis added), all three participants in Phase 2 who showed this profile of performance performed within the normal range in some language comprehension subtests.

KT demonstrated differential performance in the comprehension of auditory and written language. Her scores in the *Comprehensive Aphasia Test* indicated a deficit in auditory comprehension, but typical ability in the comprehension of written materials. KT showed no evidence of the 'severe anomia' that can occur in speakers with transcortical sensory aphasia (Goodglass, 1993).

RE showed impaired comprehension of written words and sentences, and some evidence for impaired semantic processing. Most of her errors in subtests of single-word comprehension were semantic, and most of her errors in subtests of sentence comprehension involved the selection of responses with reversed roles from the target, such as selecting *the soldier hits the singer* in response to the target *the singer hits the soldier*. Semantic errors were also present in her responses to naming tasks, as RE named *saxophone* as 'ukulele' and described the target *winding* as ‘turning up the time’. RE was able to produce a typical amount of language in the picture description task, and no semantic errors were evident in her spontaneous language output.

UT showed impaired comprehension of written words and sentences, as well as impaired naming abilities. His errors in the *Comprehensive Aphasia Test* indicated impaired semantic processing. All UT’s errors in single-word comprehension subtests involved the selection of a semantic distractor. Most of his errors in subtests of naming objects, naming actions and reading words aloud involved the production of a response that was semantically related to the target, and his errors in sentence comprehension tasks included the selection of responses with reversed roles. UT also showed a particular difficulty with the production of verbs. When attempting to name *typing* in the action-naming subtest, he described the target as ‘the girl is on a typewriter’. His score in the sentence repetition subtest was zero: his responses contained all the nouns in the target sentences, but none of the verbs. For example, he repeated the sentence *the girl eats the apple* as ‘the girl has the apple’.

Wernicke’s aphasia

Wernicke’s aphasia is a fluent aphasia characterised by significant impairments to auditory comprehension. Participants in Phase 2 with this diagnosis included CJ, SP and WD.

The language impairment of CJ was nearly undetectable to a casual observer; however, upon testing, CJ revealed deficits in the comprehension of spoken sentences and paragraphs, achieving one of the lowest scores in the subtest of spoken paragraph comprehension of all participants in Phase 2. His comprehension of written language was unimpaired, though his reading aloud of single words was characterised by the production of words orthographically related to the target, such as reading the target *trout* as ‘trot’. Most of CJ’s errors in naming subtests were semantically acceptable alternatives to the target.

SP showed some perceptual difficulties when taking part in Phase 2, as he was hard-of-hearing but did not wear hearing aids and professed to find the accent of the researcher difficult to understand at times. SP was easily frustrated by his language difficulties and distressed by his occasional perseverations in spontaneous speech. However, he showed no evidence of cognitive impairment based on his scores in the CAT cognitive screen. SP’s score in the written sentence comprehension subtest indicated impairment, but this was due to delayed rather than inaccurate responses.

WD had severe aphasia coupled with severe apraxia of speech, to the extent that he did not participate in any language output tasks. He demonstrated severe comprehension impairments for spoken and written language. His responses in subtests of single-word comprehension indicated semantic impairment, as his errors were the selection of semantic distractors.

Broca's aphasia

MJ was diagnosed with Broca's aphasia, a syndrome that refers to non-fluent language production with relative sparing of auditory comprehension. He produced only one- and two-word utterances in spontaneous speech but responded accurately and appropriately in discourse-level comprehension tasks. He showed impaired comprehension of written words and sentences, and his errors in subtests of single-word comprehension were the selection of semantic distractors. His errors in subtests of sentence-level comprehension were the selection of responses with reversed roles from the target.

As a syndrome, Broca's aphasia is associated with agrammatism, a pattern of language production described in Section 2.1.2 as lacking in function words and inflectional morphology, showing limited use of syntactic structures and overuse of infinitival verb forms (Goodglass, 1997). Agrammatic output may or may not occur with concomitant difficulties in comprehending complex syntactic sentences, or those in which the assignment of thematic roles to syntactic structure deviates from the order of subject, action and theme (Bastiaanse & van Zonneveld, 2005; Caplan & Futter, 1986; Grodzinsky, 1986).

Based on his performance in the CAT, MJ's pattern of expressive and receptive language was consistent with a diagnosis of agrammatism. MJ scored outside the normal range in the assessment of well-formedness, which reflects the accuracy of grammatical elements such as inflectional morphemes and verb arguments. All the lexical verbs that he produced in the picture description task were the present participle or infinitival forms, such as 'looking' and 'see'. MJ scored in the impaired range in subtests of auditory and visual sentence comprehension. As sentences increased in complexity across the tasks, the number of MJ's errors and delayed responses increased. Each task included two reversible passive sentences and two sentences that contained embedded clauses. Of the eight possible points from reversible passive sentences, MJ scored three. Of the eight possible points from embedded sentences, MJ scored one. Though an assessment of agrammatism may ideally be based on an investigation of a larger sample of spoken output than available from this project, MJ shows signs consistent with agrammatism in his performance in the CAT.

Global aphasia

The remaining participant in Phase 2, VH, was diagnosed with severe, global aphasia. Apraxia affected his language to the extent that he did not take part in the assessment of expressive language in Phase 2. His language output was restricted to short, stereotyped utterances, and production was characterised by phonemic distortion. His errors in subtests of language comprehension included delayed responses and the selection of semantic distractors.

6.5 Results from assessments of verb and sentence processing

Table 6.6 (over page) shows results from the verb and sentence processing tasks described in Section 4.5.2.3. Action and object naming were assessed with items from *An Object and Action Naming Battery* (Druks & Masterson, 2000). Four subtests from *The Verb and Sentence Test* (Bastiaanse et al., 2002) were employed to investigate verb comprehension, the production of finite and infinitive verb forms and sentence production. The remaining assessments of sentence and stimuli processing were novel assessments designed for Phase 2, including anagram, sentence production, function word processing and grammaticality judgement tasks. For all assessments in Table 6.6, participants' scores are shown as the number of target responses they produced in each task. VH and WD did not take part in expressive language tasks due to severe apraxia.

Results from these assessments will be discussed with regard to the participants in Phase 2 who showed verb-specific deficits, impairments in argument structure and reading difficulties that could impact their processing of the stimuli in the grammaticality judgement and sentence completion tasks in Phase 2. First, however, a note of caution is warranted about the interpretation of scores from *The Verb and Sentence Test*.

6.5.1 Interpretation of scores from The Verb and Sentence Test

There is some reason to be cautious about interpreting participants' scores in the first three subtests of *The Verb and Sentence Test* (VAST) as indicative of general impairment to verb processing. The VAST is an adaptation of a Dutch assessment, and its development began with the translation into English from Dutch (Bastiaanse et al., 2002). Peculiarly, the target response for the sentence *the boy* ___ *the soup* is *blows*. Though the verb *blow* can occur in the transitive construction when referring to musical instruments, such as *the boy blows the trumpet*, it may be more unusual in English when used to refer to an action that does not require contact with the object; *the boy blows <u>on</u> the soup* could be argued to be more typical in this instance.

The verb comprehension test required participants to match a spoken verb to one of four pictures. Images included the target action, a semantically-related action, a noun related to the target verb and a noun unrelated to the target verb. This range of distractor images meant there was a high degree of visual similarity between images of the target verb and the related noun. To illustrate, the target image for the item *cycling* showed a person riding a bicycle, and the related noun image showed only a bicycle. The researcher produced target verbs in their present participle form, but many participants in Phase 2 initially expressed uncertainty between the target image and the image of the related noun. High scores in this task may therefore reflect participants' knowledge of verb morphology to a greater degree than verb semantics.

The VAST subtests on finite and infinitive verbs required participants to produce a verb to complete a sentence that described a picture. In the test of finite verbs, targets were the present tense, third-person verb form. For example, participants saw a picture of a woman approaching a vehicle and the sentence *the woman* ___ *to her car*. The target verb was *walks* or *runs*. In the test of infinitive verbs, targets included the infinitive form of verbs. For example, participants saw a picture of a woman making an egg and the sentence *mother wants to* ___ *an egg*. The target verb was *fry*. The scoring procedure counted as correct productions that matched the target verb form. However, there was variability in the range of semantically acceptable lexical items that could be used to describe the pictures in the tasks. To illustrate, *goes* is an acceptable response to the sentence *the woman* ___ *to her car*, and *cook*, *make* or *eat* are acceptable responses to the sentence *mother wants to* ___ *an egg*. Despite the semantic and syntactic acceptability of such responses, they were not counted as correct in the task. Therefore, scores on these VAST subtests may underestimate participants' abilities in verb production, because they do not discriminate between ability in lexical selection and morphological agreement.

To illustrate, CJ demonstrated a better score in the VAST subtest of finite verbs than infinitive verbs, whereas the production of infinitive verb forms can be expected to more accurate than the production of finite verb forms (Bastiaanse et al., 2002). His scores were in fact because he produced acceptable alternatives to target verbs in the infinitive task, for example, by naming the action *drill* as *bore* and *sit* as *rest*, but these were not counted as correct in the scoring procedure.

Table 6.6 Scores from participants with aphasia in assessments of verb, sentence and stimuli processing in Phase 2

	N	Typical parts.	Normal range	BF	CJ	GW	HM	IC	JF	KT	MJ	PD	RE	SP	UT	VH	WD
An Object and Action Naming Battery																	
Action naming	30	20	28-30	**27**	**27**	**27**	**27**	**25**	29	30	**24**	**26**	**22**	**15**	**18**	-	-
Object naming	30	20	29-30	29	**27**	30	29	29	29	30	**28**	**28**	30	**25**	**25**	-	-
The Verb and Sentence Test																	
Verb comprehension	40		38-40	40	40	39	**35**	40	39	**37**	38	**37**	**34**	**32**	**31**	**28**	**3**
Filling in finite verbs	10		8-10	**1**	10	8	**1**	**4**	**0**	**5**	**0**	**7**	**4**	**1**	**0**	-	-
Filling in infinitive verbs	10		8-10	8	**5**	8	10	**7**	**7**	**7**	8	9	**7**	**4**	**5**	-	-
Sentence construction	20		16-20	19	**15**	17	19	18	19	18	19	19	20	**12**	18	-	-
Sentence processing																	
Anagrams	12	12	12	12	**9**	**11**	**11**	**10**	12	12	**10**	**4**	**9**	**11**	**2**	**8**	**3**
Sentence production	20	20	20	**14**	**17**	**19**	**19**	**15**	**18**	**18**	**12**	**16**	**16**	**13**	**6**	-	-
Stimuli processing																	
Function word processing	30	26	28-30	30	**25**	30	30	30	30	30	29	**27**	30	28	**27**	30	**21**
Grammaticality judgement	12	20	10-12	**9**	**9**	12	**9**	12	11	10	10	10	**7**	11	**4**	**9**	**8**

Note. Number of items in each task (*N*); number of typical participants contributing normal data to novel assessments (Typical parts.); scores outside normal range shown in bold; (-) indicates not assessed.

6.5.2 Performance in verb and sentence processing tasks

Object and action naming

JF and KT performed within the normal range in both the action and object naming tasks. Five participants performed outside the normal range in action naming but not object naming. The scores of BF, GW and HM in action naming were outside the normal range by just one item; IC and RE showed more impaired performance in action naming. Five other participants performed outside the normal range in both action and object naming tasks: CJ did so by just one item in each task. MJ, PD, SP and UT showed impaired action and object naming, and their scores in action naming were worse than their scores in object naming.

The Verb and Sentence Test

Scores from *The Verb and Sentence Test* (Bastiaanse et al., 2002) should be interpreted with some caution, for the reasons outlined in Section 6.5.1. Recall that participants in Table 6.6 are listed in order of their written sentence comprehension scores in the CAT. Participants' performance in the VAST auditory verb comprehension task follows this general pattern, with most participants demonstrating impairment in both or neither of these tasks. Most participants achieved higher scores in the production of infinitive verbs compared to finite verbs. CJ performed outside the normal range in the sentence construction task because he had stopped engaging in the task by its conclusion.

Sentence processing

Novel anagram and sentence production tasks were included to evaluate participants' sentence processing abilities. In the anagram task, participants arranged written constituents in the correct order to make a sentence. Only BF, JF and KT achieved a typical level of performance in the anagram task. Across participants, non-target responses in this task included the creation of unusual but acceptable sentences, such as *through a park the runners jogged*, and the insertion of words to make a grammatical sentence, especially in response to the ditransitive construction, such as *a professor taught the information in her class* for the target *the professor taught her class the information*.

In the sentence production task, participants produced a sentence that contained a target verb. Sentences were accepted if the structure was produced by the group of typical participants. Some participants demonstrated consistent patterns of errors in this task. IC's sentence productions contained missing arguments; he included no locations in sentences containing the verbs *placing*, *putting* and *handing*, and a lack of direct object in the sentence containing *deserting*. MJ produced missing arguments and made phonological errors in verb reading, for example, producing 'plaiting' for the target *placing*. RE produced some target verbs as nouns, for example, producing a sentence containing the phrase *make a donation* in response to the target *donating*. SP struggled with the task instructions and rarely produced a complete sentence, despite showing some knowledge of arguments that should occur. UT declined to respond to most items and produced errors in verb reading, reading *hiking* as ‘kicking’ and *snatching* as ‘scratching’.

Stimuli processing

Novel tasks of function word processing and grammaticality judgement assessed participants’ processing of the materials included in the grammaticality judgement and sentence completion tasks in Phase 2. Most participants performed within the normal range in the function word processing task. Performance in the grammaticality judgement task indicated the positive bias observed by Bastiaanse et al. (2002), where most errors reflected the acceptance of ungrammatical sentences.

The following sections identify participants who show difficulties processing verbs, argument structures and written words based on their performance across tasks, because these abilities form the basis of processing materials and producing responses in the experimental tasks in Phase 2.

6.5.3 Verb-specific deficits

Two participants showed impaired processing of verbs compared to nouns at the single-word level. HM and RE both scored within the normal range in assessments of object-naming (OANB) and the comprehension of spoken nouns (CAT), but were impaired in assessments of action-naming (OANB) and the comprehension of spoken verbs (VAST).

RE’s verb deficits may have been related to difficulties with verb morphology. Half of her errors in action-naming (OANB) contained a nominal form of the target verb, as she named the target *skiing* as ‘he’s going down the skis’ and the target *shaving* as ‘having a shave’. In the sentence production task, RE produced an uninflected verb form for ten of the 20 items, despite the target consisting of a verb in the present participle form, i.e. *hiking*, and she produced the target as a noun in two instances. Despite this, RE conjugated verbs appropriately in the sentence production task, as her responses showed accurate subject-verb agreement, such as ‘she hands out pens’ and ‘it sticks’.

UT appeared to demonstrate a specific deficit in verb reading. He performed very poorly in all tasks that required the comprehension of written verbs, including the anagram task, sentence production task and grammaticality judgement task. In the sentence production task, UT produced acceptable responses for only three of the 20 items. On seven occasions he failed to produce any response. His errors on other items included inaccurate reading of the target verb, as he read aloud target verbs as orthographically related verbs (reading aloud the target *hopping* as 'hoping') as well as unrelated verbs (reading aloud the target *leaning* as 'thinking'). This contrasts with his ability to produce sentences to describe pictures, as his performance in the VAST sentence construction subtest was in the normal range.

6.5.4 Argument structure impairments

Errors from two participants indicated problems producing argument structures that accurately reflected sentence semantics. Though BF performed within the normal range in the VAST sentence construction task, he achieved a score of only 14 out of 20 in the novel sentence production task. His errors in these tasks indicated occasional abnormalities in using prepositions and marking agency. In response to the image in the VAST sentence construction task, the target *the child is scratching the man*, BF responded 'she is wiping his eyes of muck'. In the sentence production task, he produced the sentences 'I'm placing this table to sit' and 'the rain is spitting'. However, he was one of only three participants to perform at ceiling in the anagram task, suggesting his difficulties mainly affected spontaneous spoken production.

MJ's atypical sentences in the novel sentence production task were due to missing arguments and some difficulty in assigning agency. For example, he produced the sentences 'it's donating to her' and 'they following over there'. His responses often lacked necessary auxiliary verbs, and all his errors in the VAST sentence construction task were agreement errors. This pattern of language output supports the conclusion that MJ's language output shows signs of agrammatism (see Section 6.4.3). Further evidence to this end is MJ's complete inability to produce finite verbs in the VAST.

In addition, WD showed evidence for difficulties in comprehending sentences. He obtained the worst score in the CAT subtest of written sentence comprehension and identified the correct order of written constituents in the anagram task for only three of 12 items. WD demonstrated a severe deficit in the processing of written sentences.

6.5.5 Reading difficulties

Several participants in Phase 2 had difficulty reading written words, including MJ, CJ and PD. MJ's errors in the sentence production task included the production of verbs orthographically related to the target. This type of error accounted for all of CJ's errors in the CAT single-word reading subtest. PD also made this type of error in the function word processing task, as she accepted the non-words *qo* and *lhem* as real words, because she read them as 'go' and 'them', indicating a lexical deficit.

WD obtained the lowest score in the function word processing task in the group of participants with aphasia. Most of his errors in the task were the acceptance of non-words.

6.6 Results of Phase 2 screening tasks

This final section shows the results of the three screening tasks in Phase 2 for participants with aphasia. Table 6.7 shows the number of number of target responses in each screening from participants with aphasia.

Table 6.7 Scores from participants with aphasia in screening tasks in Phase 2

Participant	Sentence reading *N* = 5	Function word processing *N* = 10	Grammaticality judgement *N* = 5
BF	5	10	5
CJ	5	10	4
GW	5	10	5
HM	5	10	5
IC	5	10	5
JF	5	10	5
KT	5	10	5
MJ	3	10	3
PD	5	9	4
RE	5	10	5
SP	5	10	5
UT	3	10	4
VH	0	10	5
WD	0	10	2

Results from the sentence reading and function word processing screening tasks were consistent with results from the more comprehensive language assessments reported in Table 6.6. Most errors in the grammaticality judgement tasks resulted from participants' acceptance of ungrammatical sentences.

*

This chapter described the typical participants and participants with aphasia who took part in Phase 2. Most participants with aphasia presented with mild, fluent aphasia. MJ was diagnosed with non-fluent aphasia and showed characteristics of agrammatism in his expressive and receptive language. Some participants showed difficulties in the processing of verbs, argument structure or written language, and WD produced the lowest scores in most language assessments. Overall, participants retained sufficient language processing abilities to take part in the grammaticality judgement task and sentence completion task in Phase 2. The following chapters report results from the experimental grammaticality judgment task and sentence completion task in Phase 2.

7 Results from Phase 2 grammaticality judgement task

This chapter reports results from the experimental grammaticality judgement task in Phase 2. The dependent variables included participants' number of target responses and response times. Section 7.1 reports results from typical participants, Section 7.2 reports results from participants with aphasia.

7.1 Results from typical participants

The dataset for the grammaticality judgement task contained responses from 87 typical participants to 48 verbs. The method of verb selection was detailed in Section 4.2.4.2. To summarise, 24 verbs were included in the task to elicit 'yes' judgements, and 24 verbs were included to elicit 'no' judgements. Verbs that were included to elicit 'yes' judgements were members of one of four conditions based on their values of construction frequency and lexical frequency: [high cx, high lex], [high cx, low lex], [low cx, high lex] or [low cx, low lex]. Each of the six constructions was associated with a set of four verbs, including one pair of verbs that was high in construction frequency and one pair of verbs that was low in construction frequency. Within each pair, one verb was high in lexical frequency and one verb was low in lexical frequency. Verbs that differed in one type of frequency were matched for the other. The lexical frequency of verbs that were included to elicit 'no' judgements were matched as far as possible to the lexical frequency of verbs that were included to elicit 'yes' judgements.

7.1.1 Responses to verbs in grammaticality judgement task from typical participants

This section reports participants' number of target responses and mean response times to all verbs included in the grammaticality judgement task.

7.1.1.1 Number of target responses to verbs in grammaticality judgement task

Table 7.1 (over page) shows the proportion of typical participants in Phase 2 who produced the target response to each of the verbs in the grammaticality judgement task. Verbs are shown in their frequency conditions by construction. Participants produced a total of 68 non-target responses, or 3.26% of data, to verbs included in the grammaticality judgement task to elicit a 'yes' judgement.

Typical participants produced the lowest number of target responses to verbs in the [low cx, high lex] frequency condition for four of the six constructions. The number of target responses to verbs with high construction frequency was at ceiling. Participants tended to produce a lower number of target responses to verb included in the task to elicit a 'no'

Table 7.1 Proportion of typical participants producing target response to verbs in grammaticality judgement task

	high cx, high lex		high cx, low lex		low cx, high lex		low cx, low lex	
Verbs included to elicit 'yes' judgements								
Caused motion	SEND	0.99	POST	0.99	SELL	1.00	CHUCK	0.99
Conative	LOOK	0.99	LAUGH	1.00	COME	0.87	FLY	0.97
Ditransitive	SHOW	1.00	COOK	0.99	BUILD	0.97	PACK	0.93
Intransitive motion	RUN	1.00	JUMP	0.99	LEAD	0.82	ROLL	0.97
Removal	HEAR	0.98	STEAL	0.97	WIN	0.93	LIFT	0.95
Transitive	LIKE	0.99	HATE	1.00	TRY	0.95	FEAR	1.00
Mean (SD)		0.99 (0.01)		0.99 (0.01)		0.92 (0.07)		0.97 (0.03)
Verbs included to elicit 'no' judgements								
Caused motion	GROW	0.99	TRIP	1.00	CAUSE	0.98	BOIL	0.98
Conative	SAY	0.89	FIT	0.99	THINK	1.00	SLEEP	0.99
Ditransitive	SEEM	0.99	SMELL	0.97	SPEAK	0.98	GLANCE	0.98
Intransitive motion	PUT	0.82	HAND	0.83	PLACE	0.83	FETCH	0.72
Removal	MEET	0.99	CLIMB	0.95	BREAK	0.46	STARE	1.00
Transitive	LIVE	0.98	NOD	0.98	FALL	1.00	SIGH	0.99
Mean (SD)		0.94 (0.07)		0.95 (0.06)		0.88 (0.21)		0.94 (0.11)

Note. *N* = 87.

judgement than a 'yes' judgement.

7.1.1.2 Response times to verbs in grammaticality judgement task

Table 7.2 shows typical participants' mean response times to verbs in the grammaticality judgement task. Times to non-target responses were excluded from reporting, and values represent original times prior to the application of the cleaning procedures described in Section 7.1.2.

Mean response times show that typical participants exhibited the longest response times to verbs in the [low cx, high lex] frequency condition. Response times to verbs included to elicit 'no' judgements were longer than response times to verbs included to elicit 'yes' judgements in the grammaticality judgement task for three out of four conditions.

7.1.2 Screening and cleaning procedures for response time data in grammaticality judgement task

Analysis of variance (ANOVA) was employed to explore the effects of construction frequency and lexical frequency on the response times of typical participants in the Phase 2 grammaticality judgement task. ANOVA is a parametric test, and certain assumptions about

Table 7.2 Mean response times to verbs in grammaticality judgement task from typical participants, in milliseconds

	high cx, high lex		high cx, low lex		low cx, high lex		low cx, low lex	
Verbs included to elicit 'yes' judgements								
Caused motion	SEND	1700 (703)	POST	1667 (721)	SELL	1724 (644)	CHUCK	2214 (1679)
Conative	LOOK	1465 (464)	LAUGH	1490 (459)	COME	2173 (977)	FLY	2338 (1410)
Ditransitive	SHOW	1567 (1072)	COOK	1756 (949)	BUILD	1872 (1121)	PACK	2391 (2012)
Intransitive motion	RUN	1665 (651)	JUMP	1662 (676)	LEAD	2855 (2031)	ROLL	2223 (1190)
Removal	HEAR	1967 (1389)	STEAL	2236 (3583)	WIN	2665 (2145)	LIFT	2239 (893)
Transitive	LIKE	1340 (348)	HATE	1573 (430)	TRY	2184 (1332)	FEAR	1625 (726)
Mean (SD)		1617 (216)		1731 (264)		2246 (440)		2172 (277)
Verbs included to elicit 'no' judgements								
Caused motion	GROW	2071 (796)	TRIP	2346 (1114)	CAUSE	2063 (1128)	BOIL	1908 (660)
Conative	SAY	2066 (998)	FIT	2007 (1043)	THINK	1960 (897)	SLEEP	1860 (1789)
Ditransitive	SEEM	1933 (686)	SMELL	2317 (1042)	SPEAK	1996 (961)	GLANCE	2183 (1021)
Intransitive motion	PUT	2944 (1778)	HAND	3091 (3651)	PLACE	2576 (1423)	FETCH	2839 (1655)
Removal	MEET	2207 (1178)	CLIMB	1858 (614)	BREAK	3372 (3164)	STARE	1962 (746)
Transitive	LIVE	2227 (971)	NOD	1921 (851)	FALL	1799 (623)	SIGH	1951 (964)
Mean (SD)		2241 (361)		2257 (456)		2324 (564)		2117 (371)

Note. Standard deviation of means in brackets. Only response times to target responses included. N = 87.

the data should be met in order for the test to be reliable. These include the assumptions that (1) data are drawn from a normally-distributed population; (2) the variances in each condition are similar (i.e. homogeneity of variance); (3) observations are independent; and (4) the dependent variable is measured on at least an interval scale (Field, 2009). The research design of Phase 2 meets assumptions (3) and (4), as participants and items were independent from each other, and the dependent variable - participants' response times - was measured in milliseconds.

In assessing assumption (2), variances among conditions can be evaluated using F_{max}, which is the ratio of largest condition variance to the smallest. When sample sizes across conditions are equal, as they are in the design of Phase 2, a value of F_{max} up to 10 is acceptable (Tabachnick & Fidell, 2007, p. 86). Response time data were not normally distributed ($W(86) = 0.87$, $p < 0.001$) and so violate assumption (1); therefore, data were transformed (see Section 7.1.2.1).

7.1.2.1 Transformation of the data

Response time distributions can be described as ex-Gaussian, with a peak in the distribution around the mean and a tail to the right, representing long response times. Transformations are widely recommended prior to analysis when distributions are non-normal (e.g. Field, 2009; Tabachnick & Fidell, 2007). The benefits of data transformations are multiple: as well as correcting for non-normality of the distribution, they reduce the impact of extreme values and result in homogeneity of variance between groups. All these benefits were observed when the current data were subject to an inverse transformation.

Ratcliff (1993) demonstrated that the inverse transformation retained a high degree of power in ANOVA when outliers were present in the dataset, compared to other methods of correcting response time outliers, such as the use of cut-offs of certain times or the replacement of extreme values with values equivalent to a certain number of standard deviations from participants' means. He recommended the use of the inverse transformation in circumstances like those of the present study.

The result of the inverse transformation in the current dataset is shown in Figure 7.1. The histograms show participants' mean response times to all items in the grammaticality judgement task. The graph on the left plots the raw dataset, and the graph on the right plots the transformed dataset. The distribution of the raw data was significantly different to normal by the Shapiro-Wilk test ($W(86) = 0.87$, $p < 0.001$), but the distribution of the transformed data was not ($W(86) = 0.99$, $p = 0.799$). Note that the right tail in the histogram of the raw data becomes the left tail in the histogram of the transformed data due to the inverse transformation.

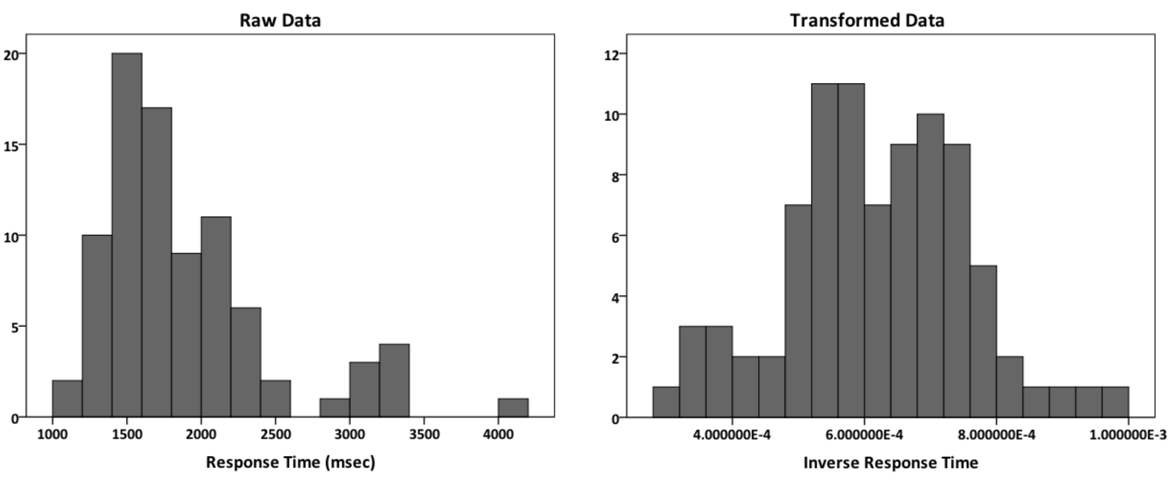


Figure 7.1 Effect of inverse transformation on distribution of raw response time data from typical participants in grammaticality judgement task

In the analyses that follow, ANOVAs were conducted on both raw data and transformed data. Summary statistics of the raw dataset are reported in milliseconds, in addition to back transformed values, following the reporting recommendations in Manikandan (2010). Standard deviations cannot be back transformed into the original scale (Bland & Altman, 1996) and so are represented in tables as (-). The condition means for participants and items that were entered into ANOVAs were calculated with transformed values, following the procedure described by Ratcliff (1993); that is, values in the dataset were transformed and then averaged.

7.1.2.2 Outliers and extreme values

Extreme values were defined as response times of 10 seconds and longer. This cut-off was selected because response times of 10 seconds and longer were extreme even when the dataset was transformed. A response time of 10 seconds may represent a different psychological process to the one that participants invoked on trials with shorter response times. With ten seconds of deliberation on a particular item, participants may have judged a dimension of the verb other than whether or not it was grammatical in the sentence stimulus, such as its naturalness. Some extreme values, such as the trial with a response time of 33 seconds, likely represent anomalies in testing conditions, such as an interruption to participants' taking part in the task.

Response times of 10 seconds and longer were removed from the dataset and replaced with values calculated using Lachaud and Renaud's (2011, p. 401) equation for the extrapolation of missing data, based on responses from all 87 participants. In this approach, missing values are replaced by the mean response times of all values from the participant and item corresponding to the missing piece of data. Though it is common practice to replace missing values with the grand mean of the entire distribution or a condition mean, this should in fact be avoided, because it can increase both Type I and Type II error rates, and reduce variance (Lachaud & Renaud, 2011). Lachaud and Renaud's method does not distort the dataset to the extent that replacement with a grand mean or condition mean does. Lachaud and Renaud (2011) provided a test of their procedure and showed that it performed well in terms of Type I error in circumstances similar to those of the current study, i.e. calculation of minF' and ex-Gaussian distribution. A total of 11 values, from seven individual participants, underwent this treatment.

7.1.2.3 Missing data

Only the times of participants' target responses were entered into statistical analyses. Times to non-target responses to verbs were excluded, resulting in missing values in the dataset. Most missing data points were found in response to [low cx, high lex] items (see Table 7.1).

This type of missing data can be described as 'missing not at random' (MNAR), or non-ignorable (Tabachnik & Fidell, 2007, p. 62). One solution to this situation is to analyse the subset of participants with complete data and then to repeat the analysis on data from all participants. If the results are similar between the analyses, confidence in them increases (Tabachnik & Fidell, 2007). This solution was adopted in the present research.

Section 7.1.3 reports results from two versions of the dataset. The first contained data only from participants who produced target responses to all 24 verbs included in the grammaticality judgement task to elicit a 'yes' judgement. That is, there were no missing data due to the production of non-target responses in this version of the dataset. The dataset included 51 of the 87 participants. A second version of the dataset contained data from all 87 participants, and response times to non-target responses were excluded.

7.1.3 ANOVAs on data from grammaticality judgement task

This section reports the results of ANOVAs that were performed on response time data from the grammaticality judgement task. ANOVAs were carried out separately over participant and item means. A repeated-measures, by-participant ANOVA was performed on response times to verbs in each of the four frequency conditions. A by-item ANOVA was performed on response times to verbs in each of the four conditions. Clark's (1973) minF' was then calculated to compare the *F*-ratios from the by-participant and by-item analyses in order to examine the effect of construction frequency, lexical frequency and their interaction when both participants and items were treated as random variables.

Throughout this section, effect sizes are reported as Pearson's correlation coefficient *r*, which Field (2009) recommends when an *F*-ratio refers to a comparison of only two means. Following Cohen (1988), values of *r* greater than or equal to 0.10 were interpreted as small effects, values greater than or equal to 0.30 as medium effects and values greater than or equal to 0.50 as large effects.

7.1.3.1 ANOVAs on data from participants who produced only target responses

Analyses in this section were performed on responses from the 51 participants who produced only target responses to all 24 verbs included in the grammaticality judgement task to elicit a 'yes' judgement. Analyses of this subset of participants were performed in response to missing data distributed not at random, as there were no missing data due to non-target responses in this version of the dataset (see Section 7.1.2.3). ANOVAs on both raw and transformed data are reported for by-participant and by-item analyses.

One participant was excluded from the analysis because her response times were very short: her mean response time to verbs in each of the four conditions was less than 1100 milliseconds. When this participant was included in the analysis, the distributions for [high cx, high lex] and [high cx, low lex] were significantly different from normal by the Shapiro-Wilk test (p = 0.027 for [high cx, high lex]; p = 0.008 for [high cx, low lex]). However, when she was removed from the analysis, the distributions for all four conditions were not significantly different from normal.

This dataset contained a total of 1200 data points. Seven of these data points were longer than 10 seconds and were therefore replaced using the method described in Section 7.1.2.2. These seven values were produced by four participants in response to six items.

Figure 7.2 shows results of the by-participant analysis of raw data in the form of the interaction graph from the ANOVA.

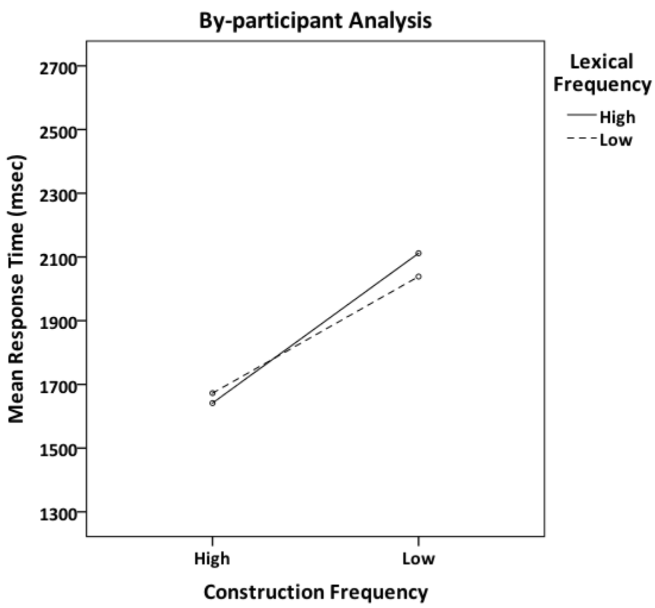


Figure 7.2 Interaction graph for by-participant analysis of raw data from subset of typical participants who produced only target responses in grammaticality judgement task

By-participant ANOVAs

Pearson's correlation coefficients for participants' response times between conditions in the grammaticality judgement task are shown in Table 7.3 (over page), from the 50 participants included in the analysis. Response times were significantly highly correlated among all conditions in the sample, underpinning the repeated-measures design for both raw and transformed data.

Table 7.3 Pearson's correlation coefficients for response times from subset of typical participants in each condition of the grammaticality judgement task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
Raw values				
high cx, high lex	-			
high cx, low lex	0.90**	-		
low cx, high lex	0.86**	0.85**	-	
low cx, low lex	0.79**	0.82**	0.87**	-
Inverse values				
high cx, high lex	-			
high cx, low lex	0.92**	-		
low cx, high lex	0.85**	0.88**	-	
low cx, low lex	0.84**	0.87**	0.91**	-

Note. ** $p < 0.01$ (one-tailed). $n = 50$.

Table 7.4 Summary statistics for by-participant analyses of dataset containing target responses from typical participants in grammaticality judgement task

	M	95% CI		*SD*	min	max
		Lower bound	Upper bound			
Raw data						
high cx, high lex	1641	1516	1767	442	1026	2882
high cx, low lex	1673	1553	1792	421	956	2954
low cx, high lex	2112	1906	2317	723	992	4319
low cx, low lex	2039	1848	2229	670	1108	4466
Back transformations						
high cx, high lex	1479	1394	1577	-	1017	2786
high cx, low lex	1524	1435	1625	-	951	2710
low cx, high lex	1788	1658	1942	-	984	3817
low cx, low lex	1755	1631	1898	-	1092	4115

Note. Mean (*M*), lower and upper bound of 95% confidence interval (CI), standard deviation (*SD*), minimum value (min) and maximum value (max).

Table 7.5 Post-hoc paired-samples *t*-tests on raw data from participants who produced only target responses in the grammaticality judgement task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	-			
high cx, low lex	$t(49) = -1.13$ $p = 0.264$ $r = 0.16$	-		
low cx, high lex	$t(49) = -8.03$ $p < 0.001$ $r = 0.75$	$t(49) = -7.23$ $p < 0.001$ $r = 0.72$	-	
low cx, low lex	$t(49) = -6.70$ $p < 0.001$ $r = 0.69$	$t(49) = -6.44$ $p < 0.001$ $r = 0.68$	$t(49) = 1.43$ $p = 0.158$ $r = 0.20$	-

Note. Significance at $p = 0.008$ with Bonferroni correction ($0.05 \div 6 = 0.008$).

Table 7.4 shows the summary statistics for the by-participant analyses of the dataset containing data from the 50 participants who produced no non-target responses. Statistics for the raw dataset are shown in the upper panel, in milliseconds, and the values from the back transformations are shown below.

A repeated-measures ANOVA on the raw data revealed a significant main effect of construction frequency on participants' response times ($F(1, 49) = 66.54$, $p < 0.001$), indicating that participants responded significantly more quickly to verbs with high construction frequency than low construction frequency. This represented a large effect ($r = 0.76$). The main effect of lexical frequency was not significant ($F(1, 49) = 0.45$, $p = 0.505$), indicating no difference in participants' response times to verbs of high and low lexical frequency. This represented a small effect ($r = 0.10$). The interaction between construction frequency and lexical frequency was approaching significance ($F(1, 49) = 3.88$, $p = 0.055$) and represented a medium-sized effect ($r = 0.27$). The interaction reflected the fact that for verbs with high construction frequency there was a small, non-significant effect of lexical frequency on participants' response times in the predicted direction, i.e. participants responded more quickly to verbs with high lexical frequency than low lexical frequency ($t(49) = -1.13$, $p = 0.264$, $r = 0.16$); however, for verbs with low construction frequency there was a small, non-significant effect of lexical frequency on participants' response times in the reverse direction, i.e. participants responded more slowly to verbs with high lexical frequency than low lexical frequency ($t(49) = 1.43$, $p = 0.159$, $r = 0.20$). Results of the full set of post-hoc *t*-tests are shown in Table 7.5.

A repeated-measures ANOVA on the transformed data revealed a significant main effect of construction frequency on participants' response times ($F(1, 49) = 117.82$, $p < 0.001$), indicating that participants responded significantly more quickly to verbs with high construction frequency than low construction frequency. This represented a large effect ($r = 0.84$). The main effect of lexical frequency was not significant ($F(1, 49) = 0.53$, $p = 0.469$), indicating no difference in participants' response times to verbs of high and low lexical frequency. This represented a small effect ($r = 0.10$). The interaction between construction frequency and lexical frequency was significant ($F(1, 49) = 5.87$, $p = 0.019$) and represented a medium-sized effect ($r = 0.33$). The interaction reflected the fact that for verbs with high construction frequency, there was a medium, non-significant effect of lexical frequency on participants' response times in the predicted direction, i.e. participants responded more quickly to verbs with high lexical frequency than low lexical frequency ($t(49) = 2.39$, $p = 0.021$, $r = 0.32$); however, for verbs with low construction frequency there

Table 7.6 Post-hoc paired-samples *t*-tests on transformed data from participants who produced only target responses in grammaticality judgement task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	-			
high cx, low lex	$t(49) = 2.39$ $p = 0.021$ $r = 0.32$	-		
low cx, high lex	$t(49) = 9.81$ $p < 0.001$ $r = 0.81$	$t(49) = 9.36$ $p < 0.001$ $r = 0.80$	-	
low cx, low lex	$t(49) = 8.79$ $p < 0.001$ $r = 0.78$	$t(49) = 8.16$ $p < 0.001$ $r = 0.76$	$t(49) = -1.13$ $p = 0.264$ $r = 0.16$	-

Note. Significance at $p = 0.008$ with Bonferroni correction (0.05 ÷ 6 = 0.008).

was a small, non-significant effect of lexical frequency on participants' response times in the reverse direction, i.e. participants responded more slowly to verbs with high lexical frequency than low lexical frequency ($t(49) = -1.13$, $p = 0.264$, $r = 0.16$). Results of the full set of post-hoc *t*-tests are shown in Table 7.6.

By-item ANOVAs

Summary statistics for the by-item analyses of the dataset containing data from the 50 participants who produced no non-target responses are available in Appendix I. A by-item ANOVA on the raw data revealed a significant main effect of construction frequency ($F(1, 20) = 18.49$, $p < 0.001$), indicating that verbs with high construction frequency elicited significantly shorter response times than verbs with low construction frequency. This represented a large effect ($r = 0.69$). The main effect of lexical frequency was not significant ($F(1, 20) = 0.05$, $p = 0.830$, $r = 0.05$), indicating that there was no difference between response times elicited by verbs with high and low lexical frequency. The interaction between construction frequency and lexical frequency was not significant ($F(1, 20) = 0.29$, $p = 0.598$) and represented a small effect ($r = 0.12$).

A by-item ANOVA on the transformed data revealed a significant main effect of construction frequency ($F(1, 20) = 18.76$, $p < 0.001$), indicating that verbs with high construction frequency elicited significantly shorter response times than verbs with low construction frequency. This represented a large effect ($r = 0.70$). The main effect of lexical frequency was not significant ($F(1, 20) = 0.04$, $p = 0.848$, $r = 0.04$), indicating that there was no difference between response times elicited by verbs with high and low lexical frequency. The interaction between construction frequency and lexical frequency was not significant ($F(1, 20) = 0.42$, $p = 0.524$) and represented a small effect ($r = 0.14$).

minF'

On the raw data, Clark's (1973) minF' indicated that the main effect of construction frequency was significant when both participants and items were treated as random effects (minF'(1, 46.03) = 14.47, $p < 0.001$), and this represented a medium effect ($r = 0.49$). The main effect of lexical frequency was not significant (minF'(1, 29.63) = 0.05, p = ns, r = 0.04), nor was the interaction between construction frequency and lexical frequency (minF'(1, 27.72) = 0.27, p = ns, $r = 0.10$).

On the transformed data, Clark's (1973) minF' indicated that the main effect of construction frequency was significant when both participants and items were treated as random effects (minF'(1, 26.60) = 16.18, $p < 0.001$), and this represented a large effect ($r = 0.61$). The main effect of lexical frequency was not significant (minF'(1, 23.04) = 0.04, p = ns, r = 0.04), nor was the interaction between construction frequency and lexical frequency (minF'(1, 22.92) = 0.39, p = ns, $r = 0.13$).

Interim summary

All analyses of data from the subset of typical participants who produced only target responses in the grammaticality judgment task revealed a large main effect of construction frequency, indicating shorter response times for verbs with high construction frequency. By-participant analyses identified a medium-sized interaction, indicating a non-significant effect of lexical frequency in the predicted direction on participants' responses to verbs with high construction frequency, but a non-significant effect of lexical frequency in the reverse direction on participants' responses to verbs with low construction frequency. Lexical frequency was not a significant main effect in any analysis.

7.1.3.2 ANOVAs on data from all participants

Analyses reported in this section included responses from a total of 86 typical participants apart from the single participant excluded for very short response times in the previous analysis who was also excluded from this analysis. The dataset contained a total of 1996 data points. Response times to non-target responses were excluded from the analysis, and these values were not replaced in the dataset. A total of 68 response times, from 36 participants, were excluded because they were times to non-target responses. A total of 11 values were extreme values greater than 10 seconds and were replaced by the procedure described in Section 7.1.2.2. These responses were produced by seven participants in response to eight items.

Figure 7.3 shows results of the by-participant analysis of raw data in the form of the interaction graphs from the ANOVA.

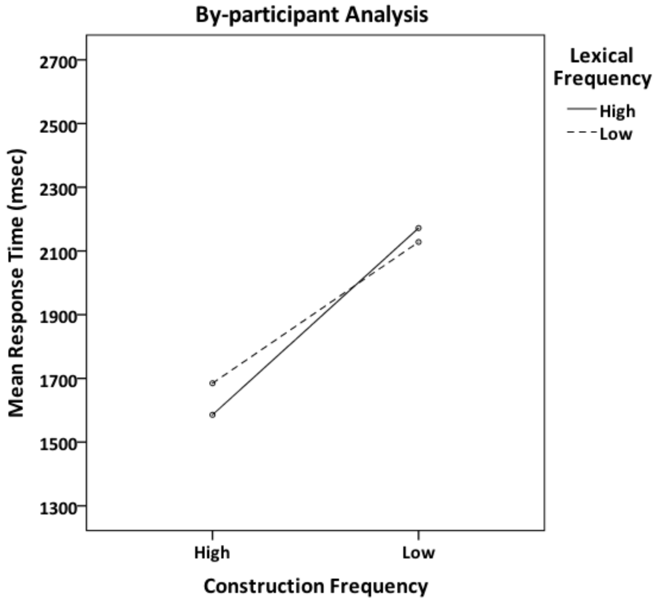


Figure 7.3 Interaction graph for by-participant analysis of raw data from all typical participants included in grammaticality judgement task

By-participant ANOVAs

Pearson's correlation coefficients for participants' response times between conditions are shown in Table 7.7, from the 86 participants included in the analysis. Significant correlations of high magnitude were evident among all conditions in the sample, supporting the repeated-measures design for both raw and transformed data.

Table 7.8 shows the summary statistics for the by-participant analyses of the dataset containing data from the 86 participants whose data were included in the analysis of the grammaticality judgement task. Statistics for the raw dataset are shown in the upper panel, in milliseconds, and the values from the back transformations are shown below.

A repeated-measures ANOVA on the raw data revealed a significant main effect of construction frequency on participants' response times ($F(1, 85) = 125.07$, $p < 0.001$), indicating that participants responded significantly more quickly to verbs with high construction frequency than low construction frequency. This represented a large effect ($r = 0.77$). The main effect of lexical frequency was not significant ($F(1, 85) = 0.92$, $p = 0.340$), indicating that participants showed no significant difference in response times to verbs with high and low lexical frequency. This represented a small effect ($r = 0.10$). The interaction between construction frequency and lexical frequency was significant ($F(1, 85) = 7.22$, $p = 0.009$) and represented a medium effect ($r = 0.28$). The interaction revealed that for verbs

Table 7.7 Pearson's correlation coefficients for response times from all typical participants in each condition of the grammaticality judgement task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
Raw values				
high cx, high lex	-			
high cx, low lex	0.84**	-		
low cx, high lex	0.82**	0.85**	-	
low cx, low lex	0.73**	0.83**	0.86**	-
Inverse values				
high cx, high lex	-			
high cx, low lex	0.88**	-		
low cx, high lex	0.81**	0.86**	-	
low cx, low lex	0.78**	0.87**	0.90**	-

Note. ** $p < 0.01$. $N = 86$.

Table 7.8 Summary statistics for by-participant analyses of dataset containing data from all typical participants

	M	95% CI		*SD*	min	max
		Lower bound	Upper bound			
Raw data						
high cx, high lex	1586	1496	1676	419	984	2882
high cx, low lex	1685	1571	1798	528	956	4112
low cx, high lex	2172	1995	2349	826	992	5237
low cx, low lex	2128	1964	2292	766	1108	4466
Back transformations						
high cx, high lex	1441	1379	1508	-	917	2786
high cx, low lex	1511	1440	1586	-	951	3484
low cx, high lex	1816	1718	1926	-	984	4566
low cx, low lex	1800	1702	1910	-	1092	4115

Note. Mean (*M*), lower and upper bound of 95% confidence interval (CI), standard deviation (*SD*), minimum value (min) and maximum value (max).

Table 7.9 Post-hoc paired-samples *t*-tests on raw data from all typical participants included in the grammaticality judgement task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	-			
high cx, low lex	$t(85) = -3.20$ $p = 0.002$ $r = 0.33$	-		
low cx, high lex	$t(85) = -10.08$ $p < 0.001$ $r = 0.74$	$t(85) = -9.72$ $p < 0.001$ $r = 0.73$	-	
low cx, low lex	$t(85) = -9.33$ $p < 0.001$ $r = 0.71$	$t(85) = -9.31$ $p < 0.001$ $r = 0.71$	$t(85) = 0.96$ $p = 0.342$ $r = 0.10$	-

Note. Significance at $p = 0.008$ with Bonferroni correction (0.05 ÷ 6 = 0.008).

with high construction frequency, lexical frequency had a medium, significant effect on participants' response times in the predicted direction, i.e. participants produced shorter response times to verbs with high lexical frequency than verbs with low lexical frequency ($t(85) = -3.20$, $p = 0.002$, $r = 0.33$); however, for verbs with low construction frequency, there was a small, non-significant effect of lexical frequency on the response times in the reverse direction, i.e. participants produced longer response times to verbs with high lexical frequency than verbs with low lexical frequency ($t(85) = 0.96$, $p = 0.342$, $r = 0.10$). Results of the full set of post-hoc *t*-tests are shown in Table 7.9 (back page).

A repeated-measures ANOVA on the transformed data revealed a significant main effect of construction frequency on participants' response times ($F(1, 85) = 272.03$, $p < 0.001$), indicating that participants responded significantly more quickly to verbs with high construction frequency than low construction frequency. This represented a large effect ($r = 0.87$). The main effect of lexical frequency was significant ($F(1, 85) = 5.26$, $p = 0.024$), indicating that participants responded significantly more quickly to verbs with high than low lexical frequency, and it represented a small effect ($r = 0.24$). The interaction between construction frequency and lexical frequency was significant ($F(1, 85) = 15.26$, $p < 0.001$) and represented a medium effect ($r = 0.39$). The interaction revealed that for verbs with high construction frequency there was a medium, significant effect of lexical frequency on participants' response times in the predicted direction, i.e. participants produced shorter response times to verbs with high lexical frequency than low lexical frequency ($t(85) = 4.06$, $p < 0.001$, $r = 0.40$); however, for verbs with low construction frequency, there was a non-significant effect of lexical frequency on response times in the reverse direction, i.e. participants produced longer response times to verbs with high lexical frequency than low lexical frequency ($t(85) = -0.65$, $p = 0.518$, $r = 0.07$). Results of the full set of post-hoc *t*-tests are shown in Table 7.10.

By-item ANOVAs

Summary statistics for the by-item analyses of the dataset containing data from the 86 participants whose data were included in the analysis of the grammaticality judgement task are available in Appendix I. A by-item ANOVA on the raw data revealed a significant main effect of construction frequency ($F(1, 20) = 26.79$, $p < 0.001$), indicating that verbs with high construction frequency elicited significantly shorter response times than verbs with low construction frequency. This represented a large effect ($r = 0.76$). The main effect of lexical frequency was not significant ($F(1, 20) = 0.05$, $p = 0.828$, $r = 0.05$), indicating that there was no significant difference in the response times elicited by verbs of high and low

Table 7.10 Post-hoc paired-samples *t*-tests on transformed data from all typical participants included in the grammaticality judgement task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	-			
high cx, low lex	$t(85) = 4.06$ $p < 0.001$ $r = 0.40$	-		
low cx, high lex	$t(85) = 14.92$ $p < 0.001$ $r = 0.85$	$t(85) = 12.96$ $p < 0.001$ $r = 0.81$	-	
low cx, low lex	$t(85) = 13.15$ $p < 0.001$ $r = 0.82$	$t(85) = 13.09$ $p < 0.001$ $r = 0.82$	$t(85) = -0.65$ $p = 0.516$ $r = 0.07$	-

Note. Significance at $p = 0.008$ with Bonferroni correction (0.05 ÷ 6 = 0.008).

lexical frequency. The interaction between construction frequency and lexical frequency was not significant ($F(1, 20) = 0.48$, $p = 0.499$, $r = 0.15$).

A by-item ANOVA on the transformed data revealed a significant main effect of construction frequency ($F(1, 20) = 27.01$, $p < 0.001$), indicating that verbs with high construction frequency elicited significantly shorter response times than verbs with low construction frequency. This represented a large effect ($r = 0.76$). The main effect of lexical frequency was not significant ($F(1, 20) = 0.23$, $p = 0.639$, $r = 0.11$), indicating that there was no significant difference in the response times elicited by verbs of high and low lexical frequency. The interaction between construction frequency and lexical frequency was not significant ($F(1, 20) = 0.61$, $p = 0.443$, $r = 0.17$).

minF'

On the raw data, Clark's (1973) minF' indicated that the main effect of construction frequency was significant when both participants and items were treated as random effects (minF'(1, 43.73) = 22.06, $p < 0.001$), and this represented a large effect ($r = 0.58$). The main effect of lexical frequency was not significant (minF'(1, 26.68) = 0.05, p = ns, $r = 0.04$), nor was the interaction between construction frequency and lexical frequency (minF'(1, 27.30) = 0.45, p = ns, $r = 0.13$).

On the transformed data, Clark's (1973) minF' indicated that the main effect of construction frequency was significant when both participants and items were treated as random effects (minF'(1, 24.11) = 24.57, $p < 0.001$), and this represented a large effect ($r = 0.71$). The main effect of lexical frequency was not significant (minF'(1, 21.78) = 0.22, p = ns) and represented a small effect ($r = .10$). The interaction between construction frequency and lexical frequency was also not significant (minF'(1, 21.62) = 0.59, p = ns) and represented a small effect (r = 0.16).

Interim summary

All analyses of data from the 86 typical participants included in the grammaticality judgement task revealed a large, significant main effect of construction frequency, demonstrating shorter response times for verbs with high construction frequency. By-participant analyses revealed a medium-sized interaction, indicating a medium, significant effect of lexical frequency in the predicted direction on participants' responses to verbs with high construction frequency but a small, non-significant effect of lexical frequency in the reverse direction on participants' responses to verbs with low construction frequency. These findings were similar to those from analyses of the subset of participants who produced only target responses in the grammaticality judgement task. A small, significant main effect of lexical frequency was revealed in the by-participant analysis of transformed data only.

Power

The study was sufficiently powered to detect medium to large effects. Power analyses revealed that the by-participant analysis was fully powered to detect the main effect of construction frequency (100%), which represented a large effect size (r = 0.77). Given the observed means and standard deviations, a sample size of approximately 1500 participants would be necessary to reach 80% power to detect the small (r = 0.10) main effect of lexical frequency, at an alpha-level of 0.05. The comparison which revealed a significant interaction between responses to verbs in the [high cx, high lex] and [high cx, low lex] conditions was fully powered (88%), and this was associated with a medium effect size (r = 0.33). However, the comparison between responses to verbs in the [low cx, high lex] and [low cx, low lex] conditions, which corresponded to a small effect size (r = 0.10), would require a sample size of approximately 750 participants to reach 80% power at an alpha-level of 0.05.

7.1.4 Construction analyses

Repeated-measures ANOVAs were performed in order to examine the effects of construction frequency, lexical frequency and their interaction in more detail for each of the six individual argument structure constructions included in the grammaticality judgement task. These analyses were conducted in order to investigate whether the significant main effect of construction frequency was robust for all constructions in the task, and whether particular constructions were driving the interaction identified in the by-participant analyses reported in Section 7.1.3. A 2x2x6 ANOVA was not performed at the outset because there was little reason to predict different participant response patterns for different constructions. In addition, only one verb was included in each condition per construction, so results from these analyses relate more specifically to the individual lexical items in the task than results from the ANOVAs reported in Section 7.1.3.

Responses from the 86 typical participants in the grammaticality judgement task were included in the analyses of individual constructions. Response times to non-target responses were excluded, and these values were not replaced in the dataset. Participants with missing data were excluded from the analysis of the repeated-measures design, resulting in data from different numbers of participants being included in the analysis of each construction. ANOVAs were performed on both raw and transformed data.

Figure 7.4 (over page) shows the interaction graphs for the six argument structure constructions included in the grammaticality judgement task, based on raw data. To preview the results, construction frequency was identified as a significant main effect in the analysis of all six constructions. The interaction indicating a reverse effect of lexical frequency on responses to verbs with low construction frequency was observed for the intransitive motion and transitive constructions. ANOVAs on data from each construction are reported in the following sections, and the full set of post-hoc *t*-tests are available in Appendix J.

Caused motion

Table 7.11 (over page) shows the summary statistics for the caused motion construction. A repeated-measures ANOVA on the raw data revealed a significant main effect of construction frequency ($F(1, 82) = 9.50$, $p = 0.003$), indicating that participants responded significantly more quickly to verbs with high construction frequency than low construction frequency, and this represented a medium effect ($r = 0.32$). The main effect of lexical frequency was significant ($F(1, 82) = 5.31$, $p = 0.024$), indicating that participants responded significantly more quickly to verbs with high lexical frequency than low lexical frequency, and this represented a small effect ($r = 0.25$). The interaction between construction frequency and lexical frequency was significant ($F(1, 82) = 7.43$, $p = 0.008$) and represented a small effect ($r = 0.29$). The interaction revealed no significant effect of construction frequency on responses to verbs with high lexical frequency ($t(84) = -0.42$, $p = 0.674$, $r = 0.05$) and a medium, significant effect of construction frequency on responses to verbs with low lexical frequency ($t(83) = -3.64$, $p < 0.001$, $r = 0.37$).

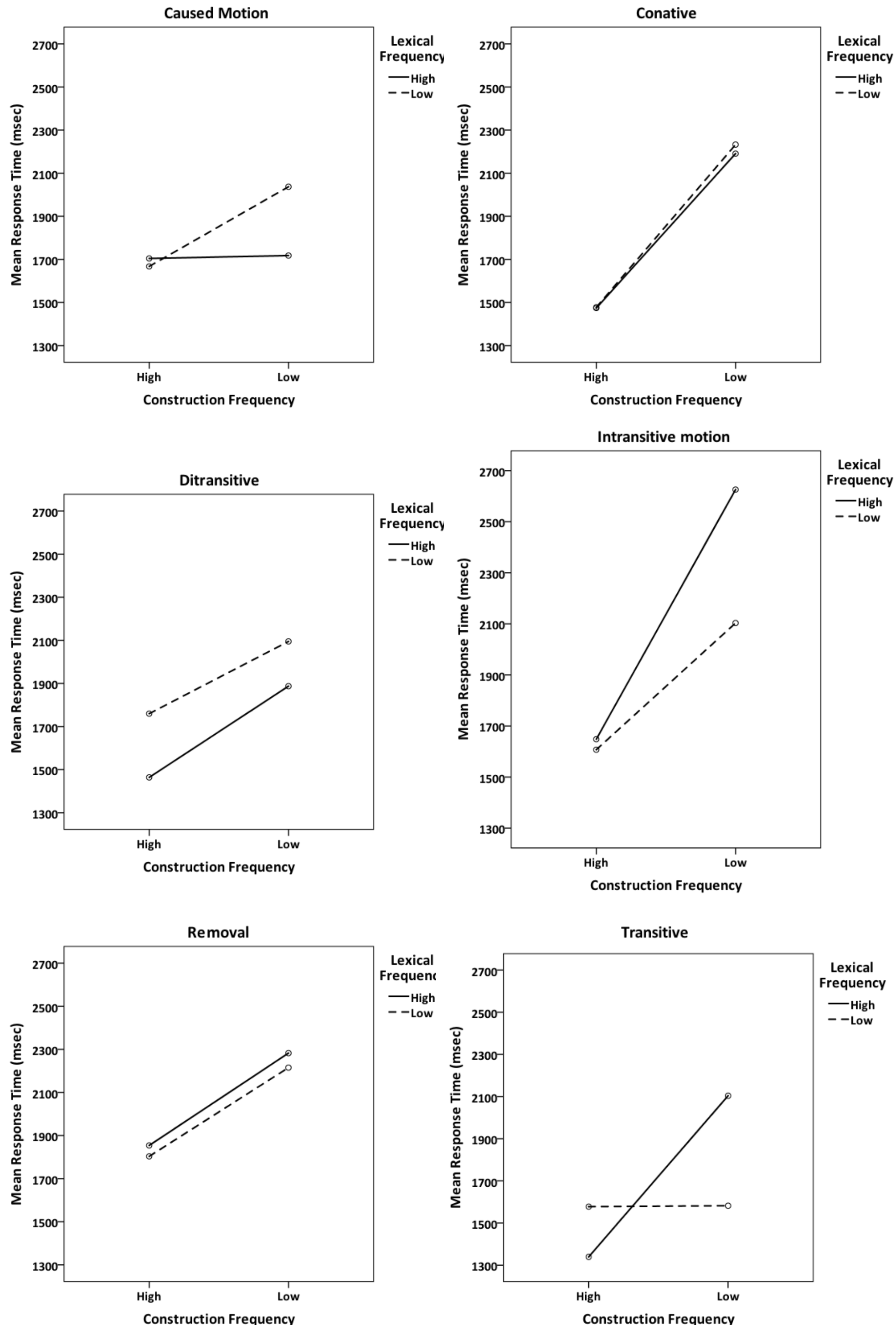


Figure 7.4 Interaction graphs for constructions in grammaticality judgement task on raw data from typical participants

This represented a medium effect (r = 0.46). The main effect of lexical frequency was significant ($F(1, 82) = 6.79$, $p = 0.011$), indicating that participants responded significantly more quickly to verbs with high lexical frequency than low lexical frequency, and this represented a small effect (r = 0.28). The interaction between construction frequency and lexical frequency was significant ($F(1, 82) = 17.82$, $p < 0.001$) and represented a medium effect (r = 0.42). The interaction revealed no significant effect of construction frequency on responses to verbs with high lexical frequency ($t(84) = 0.87$, $p = 0.389$, $r = 0.09$) and a strong, significant effect of construction frequency on responses to verbs with low lexical frequency ($t(83) = 6.39$, $p < 0.001$, $r = 0.57$).

Conative

Table 7.12 (over page) shows the summary statistics for the conative construction. A repeated-measures ANOVA on the raw data revealed a significant main effect of construction frequency ($F(1, 71) = 53.90$, $p < 0.001$), indicating that participants responded significantly more quickly to verbs with high construction frequency than low construction frequency. This represented a large effect (r = 0.66). The main effect of lexical frequency was not significant ($F(1, 71) = 0.11$, $p = 0.738$, $r = 0.04$), nor was the interaction between construction frequency and lexical frequency ($F(1, 71) = 0.10$, $p = 0.759$, $r = 0.04$).

A repeated-measures ANOVA on the transformed data revealed a significant main effect of construction frequency ($F(1, 71) = 154.71$, $p < 0.001$), indicating that participants responded significantly more quickly to verbs with high construction frequency than low construction frequency. This represented a large effect (r = 0.83). The main effect of lexical frequency was not significant ($F(1, 71) = 0.39$, $p = 0.534$, $r = 0.07$), nor was the interaction between construction frequency and lexical frequency ($F(1, 71) = 1.18$, $p = 0.281$, $r = 0.13$).

Ditransitive

Table 7.13 (over page) shows the summary statistics for the ditransitive construction. A repeated-measures ANOVA on the raw data revealed a significant main effect of construction frequency ($F(1, 76) = 24.27$, $p < 0.001$), indicating that participants responded significantly more quickly to verbs with high construction frequency than low construction frequency. This represented a medium effect (r = 0.49). The main effect of lexical frequency was significant ($F(1, 76) = 17.90$, $p < 0.001$), indicating that participants responded significantly more quickly to verbs with high lexical frequency than low lexical frequency, representing a medium effect (r = 0.44). The interaction between construction frequency and lexical frequency was not significant ($F(1, 76) = 0.27$, $p = 0.605$, $r = 0.06$).

Table 7.11 Summary statistics for analysis of caused motion construction in grammaticality judgement task

	M	95% CI		*SD*	min	max
		Lower bound	Upper bound			
Raw data						
high cx, high lex	1704	1551	1858	702	784	5269
high cx, low lex	1667	1510	1825	721	875	IC0
low cx, high lex	1718	1581	1855	627	877	4369
low cx, low lex	2037	1824	2250	975	1052	7372
Back transformations						
high cx, high lex	1536	1447	1639	-	784	5263
high cx, low lex	1511	1429	1602	-	875	RE4
low cx, high lex	1560	1471	1663	-	876	4367
low cx, low lex	1784	1673	1912	-	1052	7353

Note. Mean (*M*), lower and upper bound of 95% confidence interval (CI), standard deviation (*SD*), minimum value (min) and maximum value (max). *n* = 83.

Table 7.12 Summary statistics for analysis of conative construction in grammaticality judgement task

	M	95% CI		*SD*	min	max
		Lower bound	Upper bound			
Raw data						
high cx, high lex	1474	1371	1577	437	906	3152
high cx, low lex	1478	1375	1580	436	862	3331
low cx, high lex	2191	1956	2425	999	1056	5720
low cx, low lex	2233	1927	2538	1300	970	7754
Back transformations						
high cx, high lex	1372	1296	1458	-	907	3155
high cx, low lex	1379	1303	1463	-	862	3333
low cx, high lex	1910	1778	2063	-	1056	5714
low cx, low lex	1834	1688	2007	-	970	7752

Note. Mean (*M*), lower and upper bound of 95% confidence interval (CI), standard deviation (*SD*), minimum value (min) and maximum value (max). *n* = 72.

Table 7.13 Summary statistics for analysis of ditransitive construction in grammaticality judgement task

	M	95% CI		*SD*	min	max
		Lower bound	Upper bound			
Raw data						
high cx, high lex	1464	1381	1547	366	783	2805
high cx, low lex	1760	1540	1979	967	1050	7021
low cx, high lex	1887	1626	2149	1153	883	8005
low cx, low lex	2095	1858	2332	1046	1083	6186
Back transformations						
high cx, high lex	1385	1317	1461	-	784	2809
high cx, low lex	1540	1449	1643	-	1050	7042
low cx, high lex	1608	1504	1726	-	883	8000
low cx, low lex	1790	1663	1937	-	1083	6173

Note. Mean (*M*), lower and upper bound of 95% confidence interval (CI), standard deviation (*SD*), minimum value (min) and maximum value (max). *n* = 77.

A repeated-measures ANOVA on the transformed data revealed a significant main effect of construction frequency ($F(1, 76) = 61.28$, $p < 0.001$), indicating that participants responded significantly more quickly to verbs with high construction frequency than low construction frequency. This represented a large effect ($r = 0.67$). The main effect of lexical frequency was significant ($F(1, 76) = 22.30$, $p < 0.001$), indicating that participants responded significantly more quickly to verbs with high lexical frequency than low lexical frequency, and represented a medium effect ($r = 0.48$). The interaction between construction frequency and lexical frequency was not significant ($F(1, 76) = 0.12$, $p = 0.731$ $r = 0.04$).

Intransitive motion

Table 7.14 (over page) shows the summary statistics for the intransitive motion construction. A repeated-measures ANOVA on the raw data revealed a significant main effect of construction frequency ($F(1, 69) = 61.57$, $p < 0.001$), indicating that participants responded significantly more quickly to verbs with high construction frequency than low construction frequency. This represented a large effect ($r = 0.69$). The main effect of lexical frequency was significant ($F(1, 69) = 13.40$, $p < 0.001$), revealing a reverse frequency effect, as participants responded significantly more quickly to verbs with low lexical frequency than high lexical frequency. This represented a medium effect ($r = 0.40$). The interaction between construction frequency and lexical frequency was significant ($F(1, 69) = 7.87$, $p = 0.007$) and represented a medium effect ($r = 0.32$). The interaction revealed no significant effect of lexical frequency on responses to verbs with high construction frequency ($t(84) = = -0.01$, $p = 0.995$, $r = 0.00$) and a significant, medium effect of lexical frequency in the reverse direction on responses to verbs with low construction frequency ($t(69) = 3.40$, $p = 0.001$, $r = 0.38$).

A repeated-measures ANOVA on the transformed data revealed a significant main effect of construction frequency ($F(1, 69) = 78.43$, $p < 0.001$), indicating that participants responded significantly more quickly to verbs with high construction frequency than low construction frequency. This represented a large effect ($r = 0.73$). The main effect of lexical frequency was significant ($F(1, 69) = 7.98$, $p = 0.006$), revealing a reverse frequency effect, as participants responded significantly more quickly to verbs with low lexical frequency than high lexical frequency. This represented a medium effect ($r = 0.32$). The interaction between construction frequency and lexical frequency was not significant ($F(1, 69) = 2.47$, $p = .120$, $r = 0.19$); however, consistent with the analysis of the raw data, the effect of lexical frequency on responses to verbs with high construction frequency represented a non-significant effect ($t(84) = -0.27$, $p = 0.790$, $r = 0.03$), but the effect of lexical frequency on

Table 7.14 Summary statistics for analysis of intransitive motion construction in grammaticality judgement task

	M	95% CI		*SD*	min	max
		Lower bound	Upper bound			
Raw data						
high cx, high lex	1648	1502	1795	615	870	4472
high cx, low lex	1607	1478	1736	543	802	4426
low cx, high lex	2626	2306	2946	1343	920	6306
low cx, low lex	2103	1899	2307	855	1068	5422
Back transformations						
high cx, high lex	1502	1409	1606	-	870	4464
high cx, low lex	1471	1376	1579	-	802	4425
low cx, high lex	2106	1901	2360	-	920	6289
low cx, low lex	1862	1733	2015	-	1068	5435

Note. Mean (*M*), lower and upper bound of 95% confidence interval (CI), standard deviation (*SD*), minimum value (min) and maximum value (max). $n = 70$.

Table 7.15 Summary statistics for analysis of removal construction in grammaticality judgement task

	M	95% CI		*SD*	min	max
		Lower bound	Upper bound			
Raw data						
high cx, high lex	1854	1670	2038	810	803	5403
high cx, low lex	1803	1639	1968	723	921	4186
low cx, high lex	2283	2019	2546	1162	1018	7707
low cx, low lex	2215	2031	2400	812	931	5099
Back transformations						
high cx, high lex	1627	1517	1754	-	803	5405
high cx, low lex	1592	1484	1718	-	921	4184
low cx, high lex	1955	1810	2126	-	1017	7692
low cx, low lex	1946	1796	2123	-	931	5102

Note. Mean (*M*), lower and upper bound of 95% confidence interval (CI), standard deviation (*SD*), minimum value (min) and maximum value (max). $n = 77$.

Table 7.16 Summary statistics for analysis of transitive construction in grammaticality judgement task

	M	95% CI		*SD*	min	max
		Lower bound	Upper bound			
Raw data						
high cx, high lex	1340	1264	1415	340	782	2449
high cx, low lex	1578	1483	1672	428	951	2761
low cx, high lex	2104	1888	2320	977	1015	5628
low cx, low lex	1582	1465	1699	528	859	3428
Back transformations						
high cx, high lex	1262	1197	1334	-	782	2451
high cx, low lex	1477	1400	1563	-	951	2762
low cx, high lex	1809	1680	1958	-	1015	5618
low cx, low lex	1453	1370	1545	-	859	3425

Note. Mean (*M*), lower and upper bound of 95% confidence interval (CI), standard deviation (*SD*), minimum value (min) and maximum value (max). $n = 81$.

responses to verbs with low construction frequency represented a significant, medium effect ($t(69) = -2.70$, $p = 0.009$, $r = 0.31$).

Removal

Table 7.15 (back page) shows the summary statistics for the removal construction. A repeated-measures ANOVA on the raw data revealed a significant main effect of construction frequency ($F(1, 76) = 25.37$, $p < 0.001$), indicating that participants responded significantly more quickly to verbs with high construction frequency than low construction frequency. This represented a large effect ($r = 0.50$). The main effect of lexical frequency was not significant ($F(1, 76) = 0.68$, $p = 0.413$, $r = 0.09$), nor was the interaction between construction frequency and lexical frequency ($F(1, 76) = 0.01$, $p = 0.913$, $r = 0.01$).

A repeated-measures ANOVA on the transformed data revealed a significant main effect of construction frequency ($F(1, 76) = 35.56$, $p < 0.001$), indicating that participants responded significantly more quickly to verbs with high construction frequency than low construction frequency. This represented a large effect ($r = 0.56$). The main effect of lexical frequency was not significant ($F(1, 76) = 0.40$, $p = 0.532$, $r = 0.07$), nor was the interaction between construction frequency and lexical frequency ($F(1, 76) = 0.22$ $p = 0.641$, $r = 0.05$).

Transitive

Table 7.16 (back page) shows the summary statistics for the transitive construction. A repeated-measures ANOVA on the raw data revealed a significant main effect of construction frequency ($F(1, 80) = 37.33$, $p < 0.001$), indicating that participants responded significantly more quickly to verbs with high construction frequency than low construction frequency. This represented a large effect ($r = 0.56$). The main effect of lexical frequency was significant ($F(1, 80) = 6.41$, $p = 0.013$), indicating a reverse frequency effect, as participants responded significantly more quickly to verbs with low lexical frequency than high lexical frequency. This represented a small effect ($r = 0.27$). The interaction between construction frequency and lexical frequency was significant ($F(1, 80) = 49.18$, $p < 0.001$) and represented a large effect ($r = .62$). The interaction revealed a significant, large effect of construction frequency on responses to verbs with high lexical frequency ($t(80) = -7.20$, $p < 0.001$, $r = 0.63$) but no significant effect of construction frequency on responses to verbs with low lexical frequency ($t(85) = -0.74$, $p = 0.464$, $r = 0.08$).

A repeated-measures ANOVA on the transformed data revealed a significant main effect of construction frequency ($F(1, 80) = 48.15$, $p < 0.001$), indicating that participants responded significantly more quickly to verbs with high construction frequency than low construction frequency. This represented a large effect ($r = 0.61$). The main effect of lexical frequency was not significant ($F(1, 80) = 0.50$, $p = 0.480$, $r = 0.08$). The interaction between construction frequency and lexical frequency was significant ($F(1, 80) = 72.28$, $p < 0.001$) and represented a large effect ($r = 0.69$). The interaction revealed a significant, large effect of construction frequency on responses to verbs with high lexical frequency ($t(80) = 9.27$, $p < 0.001$, $r = 0.72$) but no significant effect of construction frequency on responses to verbs with low lexical frequency ($t(85) = -0.41$, $p = 0.681$, $r = 0.04$).

Summary of results from individual argument structure constructions

The main effect of construction frequency proved robust in all constructions, as it was identified as a medium or large effect in the analysis of each of the six constructions. A main effect of lexical frequency was identified in the caused motion and ditransitive constructions, where it was smaller than the effect of construction frequency. Analyses of the intransitive motion and transitive constructions revealed a reverse effect of lexical frequency in the context of an interaction: the reverse lexical frequency effect occurred only in response to verbs that were low in construction frequency.

7.1.5 Additional analyses

Verbs were selected for inclusion in Phase 2 based on their values of construction frequency and lexical frequency. In the selection of materials, it was not possible to control for other factors known to affect single-word processing. This section reports the results of additional analyses that investigated whether the number of target responses and response times were affected by variables relating to attributes of the participants or items included in the study.

Analyses in this section were carried out on the number of target responses and the raw and transformed values of response times from all 86 participants included in the analysis of the grammaticality judgement task. Note that because the inverse transformation has the effect of reversing the distribution of raw response time values, so the left tail of the distribution became the right tail, negative correlations with inverse values reported below represent a real positive relationship.

7.1.5.1 Participant variables

Age

Pearson's correlation coefficients were used to explore the effect of age on the number of target responses participants produced and participants' response times. There was no significant relationship between participants' age in years and their number of target responses ($r = 0.12$, $p = 0.131$). There was no significant relationship between participants' age and their response times to verbs in the raw data ($r = 0.12$, $p = 0.133$); however, this relationship reached significance in the transformed data (r = -0.20, p = 0.036), indicating that participants' response times increased with age.

In order to further explore this finding, Table 7.17 shows the correlations between age and response times in the transformed data for each condition. The effect of age was significant only for responses to verbs with high construction frequency and, across all conditions, represented a low magnitude relationship.

Table 7.17 Correlations for relationship between participants' age and response times in grammaticality judgement task by condition

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
Pearson's correlation coefficient *r*	-0.25**	-0.21*	-0.16	-0.08
p-value	0.009	0.024	0.073	0.227

Note. * p < 0.05. ** p < 0.01. Correlations on transformed data; N = 86.

Age was included as a covariate in the by-participant ANOVA on transformed data from all 86 participants, in order to explore the effects of construction frequency and lexical frequency whilst accounting for age. In this analysis, the effect of construction frequency was significant (F(1, 84) = 15.31, p < 0.001), indicating that participants responded significantly more quickly to verbs with high construction frequency than low construction frequency. This represented a medium effect (r = 0.39). The main effect of lexical frequency was not significant (F(1, 84) = 2.41, p = 0.124, r = 0.17), indicating that participants showed no significant difference between response times to verbs of high and low lexical frequency. The interaction between construction frequency and lexical frequency was not significant (F(1, 84) = 0.08, p = 0.775, r = 0.03).

The interaction between construction frequency and age was significant (F(1, 84) = 4.95, p = 0.029) and represented a small effect (r = 0.24). The interaction between lexical frequency and age was not significant (F(1, 84) = 1.76, p = 0.188, r = 0.14), nor was the two-way interaction between construction frequency, lexical frequency and age (F(1, 84) = 0.47, p = 0.497, r = -0.07). Results of this ANCOVA confirm that the effect of construction frequency was robust after correcting for age.

Education

Pearson's correlation coefficients were used to explore the effect of education on the number of target responses participants produced and their response times. There was a mild, positive correlation between the number of years participants spent in education and their number of target responses (r = 0.21, one-tailed p = 0.025). This correlation was likely influenced by the performance of the single individual who produced the lowest number of target responses in Phase 2, accepting only 58% of the verbs included to elicit a 'yes' judgement, as he also spent the least amount of time in education, ten years. However, other participants who also spent ten years in education performed at ceiling in the task. Results from the non-parametric correlation coefficient support this observation, as Spearman's rho was not significant ($r_s = 0.12$, $p = 0.142$).

There was no significant effect of the number of years participants spent in education on their mean response times to verbs in the raw data ($r = 0.11$, $p = 0.168$) or the transformed data ($r = -0.10$, $p = 0.169$).

Gender

Independent samples *t*-tests were used to examine differences in the number of target responses and response times between women ($n = 42$) and men ($n = 44$). There was no significant difference in the number of target responses produced by women and men ($t(84) = 0.45$, $p = 0.655$). There was no significant difference in response times between women and men in the raw data ($t(84) = -1.38$, $p = 0.170$) or in the transformed data ($t(84) = 1.43$, $p = 0.157$).

Handedness

Independent samples *t*-tests were used to examine differences in the number of target responses and response times between right- ($n = 76$) and left-handed ($n = 10$) participants. There was no significant difference between the number of target responses of right- and left-handed participants ($t(9.35) = 0.83$, $p = 0.429$). There was no significant difference in response times between right- and left-handed participants in the raw data ($t(84) = -0.30$, $p = 0.763$) or in the transformed data($t(84) = 0.70$, $p = 0.484$).

7.1.5.2 Item variables

Pearson's correlation coefficients were used to examine the effect of variables known to influence single-word processing on response times to verbs in the dataset, in the raw and transformed data. These variables were introduced in Section 4.4.3 and are briefly described below.

Age-of-acquisition

Age-of-acquisition refers to the age at which a word is learned, and ease of processing is associated with earlier acquired lexemes. Ratings for the age-of-acquisition of verbs in the current study were taken from data published by Kuperman et al. (2012). These researchers presented adults with uninflected word forms and asked them to rate the age at which they acquired the word. Ratings were available for 23 of the 24 verbs included in the grammaticality judgement task. There was no significant relationship between typical participants' mean response times to verbs and verbs' age-of-acquisition ratings in the raw data (r = -0.06, p = 0.400) or in the transformed data (r = 0.04, p = 0.420).

Imageability

Imageability is a semantic variable that refers to how easy or difficult it is to picture a word. More highly imageable words are associated with more efficient processing. Ratings for the imageability of verbs in the current study were taken from data published by Cortese and Fugett (2004), who asked adults to rate imageability on a scale of 1 (low imageability) to 7 (high imageability). Ratings were available for all 24 verbs included in the grammaticality judgement task. There was no significant relationship between typical participants' mean response times to verbs and verbs' imageability ratings in the raw data (r = 0.15, p = 0.249) or in the transformed data (r = -0.09, p = 0.345).

Lexical decision times

The British Lexicon Project published response times to single words from a lexical decision task (Keuleers et al., 2012). Lexical decision times can be taken as indicating the ease of processing a single word in isolation. There was a significant correlation between mean response times to verbs in the grammaticality judgement task and mean response times to the verbs in the lexical decision task, in the raw data (r = 0.36, p = 0.043) and in the transformed data (r = -0.38, p = 0.035). This correlation indicated that verbs which elicited shorter or longer response times in a lexical decision task also elicited shorter or longer response times in the grammaticality judgement task. This result can be expected in the present study, because both tasks involve the process of single-word recognition.

Section 4.4.3 described the component processes involved in the completion of the grammaticality judgement task. Participants must first read verbs as single words, and then decide whether or not those verbs can be used in the sentence stimulus. High lexical frequency was expected to facilitate single-word reading (see Section 2.2.1.1), and lexical frequency was identified as a main effect having a small but significant effect on response times in the grammaticality judgement task in the dataset containing transformed response times from 86 typical participants (see Section 7.1.3.2). In order to explore whether the effect of lexical frequency observed in the grammaticality judgement task could be attributed solely to single-word processing, lexical decision times for each verb from the British Lexicon Project were included as a covariate in by-item ANOVAs on raw and transformed data from 86 participants.

The by-item ANOVA on raw data revealed a significant main effect of construction frequency ($F(1, 19) = 19.91$, $p < 0.001$), indicating verbs with high construction frequency elicited significantly shorter response times than verbs with low construction frequency. This represented a large effect ($r = 0.72$). The main effect of lexical frequency was not significant ($F(1, 19) = 0.02$, $p = 0.896$, $r = 0.03$), nor was the interaction between construction frequency and lexical frequency ($F(1, 19) = 0.48$, $p = 0.495$, $r = 0.16$).

The by-item ANOVA on transformed data revealed a significant main effect of construction frequency ($F(1, 19) = 19.83$, $p < 0.001$), indicating verbs with high construction frequency elicited significantly shorter response times than verbs with low construction frequency. This represented a large effect ($r = 0.71$). The main effect of lexical frequency was not significant ($F(1, 19) = 0.13$, $p = 0.723$, $r = 0.08$), nor was the interaction between construction frequency and lexical frequency ($F(1, 19) = 0.63$, $p = 0.437$, $r = 0.18$).

These findings are similar to those reported in Section 7.1.3.2 for by-item ANOVAs that did not include lexical decision times as a covariate. Results from the ANCOVA demonstrate that the effect of construction frequency cannot be attributed solely to the process of single-word recognition, because the effect survives in a statistical model that contains lexical decision times as a covariate. The effect of lexical frequency, however, can be attributed to the process of single-word recognition, because it does not appear as a significant effect in the statistical model containing lexical decision times as a covariate.

Orthography

The effect of three orthographic characteristics of verbs on the mean response times they elicited were examined. Coltheart's N refers to the number of words that differ in only one letter from a target word. OLD20 refers to the Levenstein distance between a target word and the twenty words most similar to it, where Levenstein distance is the number of deletions, insertions or substitutions required to transform the target word. These measures were included in data available through the British Lexicon Project (Keuleers et al., 2012), and data were available for all 24 verbs included in the grammaticality judgement task. The number of letters in each verb was also considered. None of these orthographic characteristics proved to affect response times.

There was no significant relationship between the mean response times to verbs and verbs' Coltheart's N in the raw data ($r = 0.00$, $p = 0.497$) or in the transformed data ($r = -0.01$, $p = 0.479$). There was no significant relationship between the mean response times to verbs and verbs' OLD20 in the raw data ($r = 0.01$, $p = 0.479$) or in the transformed data ($r = -0.01$, $p = 0.483$). Finally, there was no significant relationship between mean response times to verbs and the number of letters they contained in the raw data (r = -0.26, p = 0.108) or in the transformed data (r = 0.24, p = 0.132).

7.1.6 Summary of results from typical participants in grammaticality judgement task

All analyses reported in this section, including by-participant, by-item and minF' analysis of raw and transformed data from 86 typical participants and the subset of 50 participants who produced only target responses, identified a medium or large effect of construction frequency. Participants responded more quickly to verbs with high construction frequency than low construction frequency, and this finding was consistent across the six constructions in the task and robust against the effect of age.

Lexical frequency was revealed as a small but significant main effect in the by-participant analysis of transformed data from 86 participants. However, it is likely this effect was attributable to the process of single-word recognition, as lexical frequency was not a significant main effect in an ANCOVA that included lexical decision times as a covariate. Lexical frequency was identified as a significant main effect in the analysis of the caused motion and ditransitive constructions, where participants responded more quickly to verbs with high lexical frequency than low lexical frequency. It represented a small and medium effect for these respective constructions.

By-participant analyses identified a significant interaction, revealing an effect of lexical frequency in the predicted direction on responses to verbs with high construction frequency but a reverse effect of lexical frequency on responses to verbs with low construction frequency. This reverse effect of lexical frequency on responses to verbs with low construction frequency was evident in the analysis of the intransitive motion and transitive constructions only.

7.2 Results from participants with aphasia

Results from participants with aphasia in the Phase 2 grammaticality judgement task are reported below. Throughout this section, participants are listed in order of their scores in the written sentence comprehension subtest from the *Comprehensive Aphasia Test*, beginning with participants with the highest scores. The number of target responses and response times were analysed for the group of participants with aphasia, for each individual with aphasia and by construction. Section 7.2.3 compares the performance of individual participants with aphasia to typical participants using effect scores and Bayesian Standardised Difference Tests, as described in Section 4.4.2.2.

Table 7.18 Number of target responses from participants with aphasia in grammaticality judgement task

	All items *n* = 60	No judgements *n* = 30	Yes judgements *n* = 30	high cx, high lex *n* = 6	high cx, low lex *n* = 6	low cx, high lex *n* = 6	low cx, low lex *n* = 6
BF	51	24	27	6	6	4	5
CJ	48	21	27	6	4	6	5
GW	51	22	29	6	6	5	6
HM	48	24	24	6	6	4	4
IC	42	12	30	6	6	6	6
JF	46	23	23	4	5	4	5
KT	41	20	21	5	5	5	2
MJ	40	13	27	6	6	5	4
PD	50	23	27	6	5	5	5
RE	46	20	26	6	6	4	4
SP	40	27	13	5	2	2	1
UT	33	12	21	5	5	3	3
VH	23	11	12	5	4	3	0
WD	34	14	20	6	3	5	3
Mdn	44	20.5	25	6	5	4.5	4
M	42.4	19	23.4	5.6	4.9	4.4	3.8

7.2.1 Number of target responses from participants with aphasia in grammaticality judgement task

7.2.1.1 Number of target responses from individual participants with aphasia in grammaticality judgement task

Table 7.18 shows the number of target responses that participants with aphasia produced in the grammaticality judgement task. Section 5.4.1 described how stimuli were presented for participants with aphasia. To recap, participants with aphasia experienced trials blocked by construction, and the first two trials within a block were practice items which were not included in the task for typical participants. For participants with aphasia, this resulted in the task containing a total of 60 verbs: each of the six constructions was associated with 10 verbs (4 verbs to elicit 'yes' judgements + 4 verbs to elicit 'no' judgements + 2 practice items, including one 'yes' judgement and one 'no' judgement).

The number of target responses that participants produced in response to all items in the task, as well as to 'no' and 'yes' items separately, is shown in the first three columns of Table 7.18. The final four columns show the number of participants' target responses to verbs in each of the four frequency conditions under consideration.

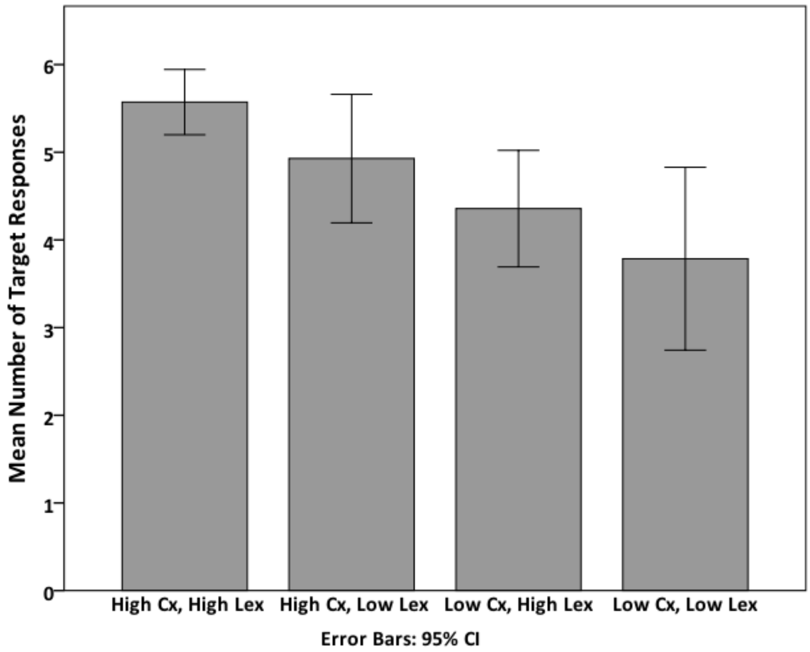


Figure 7.5 Mean number of target responses produced by participants with aphasia in Phase 2 grammaticality judgement task

7.2.1.1 Number of target responses from group of participants with aphasia in grammaticality judgement task

Figure 7.5 shows the mean number of target responses the group of participants with aphasia produced in response to verbs in each condition. The mean number of responses is reported, despite the use of non-parametric statistical tests, due to the limited range in the number of responses.

In order to investigate main effects of construction frequency and lexical frequency at the group level, participants' scores from the four frequency conditions were collapsed into two groups in order to compare the number of target responses to verbs with high and low frequency. Verbs were collapsed in two ways: once to explore a difference in construction frequency, and again to explore a difference in lexical frequency. To illustrate, the effect of construction frequency was inspected by comparing the number of target responses to verbs in two groups: high construction frequency, including verbs in the conditions [high cx, high lex] and [high cx, low lex], and low construction frequency, including verbs in the conditions [low cx, high lex] and [low cx, low lex]. The effect of lexical frequency was inspected by comparing the number of target responses in the conditions [high cx, high lex] and [low cx, high lex] to the number of target responses in the conditions [high cx, low lex] and [low cx, low lex]. Two Wilcoxon signed-rank tests were used to make the comparisons. Effect sizes are reported as *r*, following the correct formula from Field (2009).

Table 7.19 Results from Wilcoxon signed-rank tests on differences in number of target responses from participants with aphasia between conditions of grammaticality judgement task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	-			
high cx, low lex	T = 2.00 p = 0.071 r = -0.34	-		
low cx, high lex	T = 0.00 p = 0.004 r = -0.54	T = 15.00 p = 0.189 r = -0.25	-	
low cx, low lex	T = 2.50 p = 0.004 r = -0.55	T = 2.00 p = 0.014 r = -0.46	T = 10.50 p = 0.142 r = -0.28	-

Participants with aphasia produced significantly more target responses to verbs with high construction frequency than to verbs with low construction frequency (T = 2.50, two-tailed p = 0.004), and this represented a large effect (r = -0.54). There was a numerical difference in the number of target responses to verbs with different levels of lexical frequency, with a greater number of target responses to verbs with high lexical frequency than low lexical frequency, and this difference was approaching significance (T = 10.00, two-tailed p = 0.072, r = -0.34).

In order to explore interactions between conditions, Wilcoxon signed-rank tests were performed to investigate differences in the number of target responses between each condition in the task. Results indicated that the effect of lexical frequency on responses to verbs with high construction frequency was approaching significance (T = 2.00, p = 0.071), and this represented a medium effect (r = -0.34). In contrast, there was no significant effect of lexical frequency on responses to verbs with low construction frequency (T = 10.50, p = 0.142, r = -0.28). Results of the full set of comparisons are shown in Table 7.19 (over page).

To summarise, construction frequency represented a large, significant main effect on the number of target responses that participants with aphasia produced in the grammaticality judgement task. Lexical frequency represented a medium-sized main effect that was approaching significance. Lexical frequency had a significant effect on responses to verbs with high construction frequency, but not low construction frequency.

7.2.1.2 Number of target responses from group of participants with aphasia in grammaticality judgement task by construction

The analyses described in Section 7.2.1.2 were performed separately for each of the six argument structure constructions included in Phase 2. Figure 7.6 shows the mean number of target responses participants with aphasia produced in each condition of the six argument

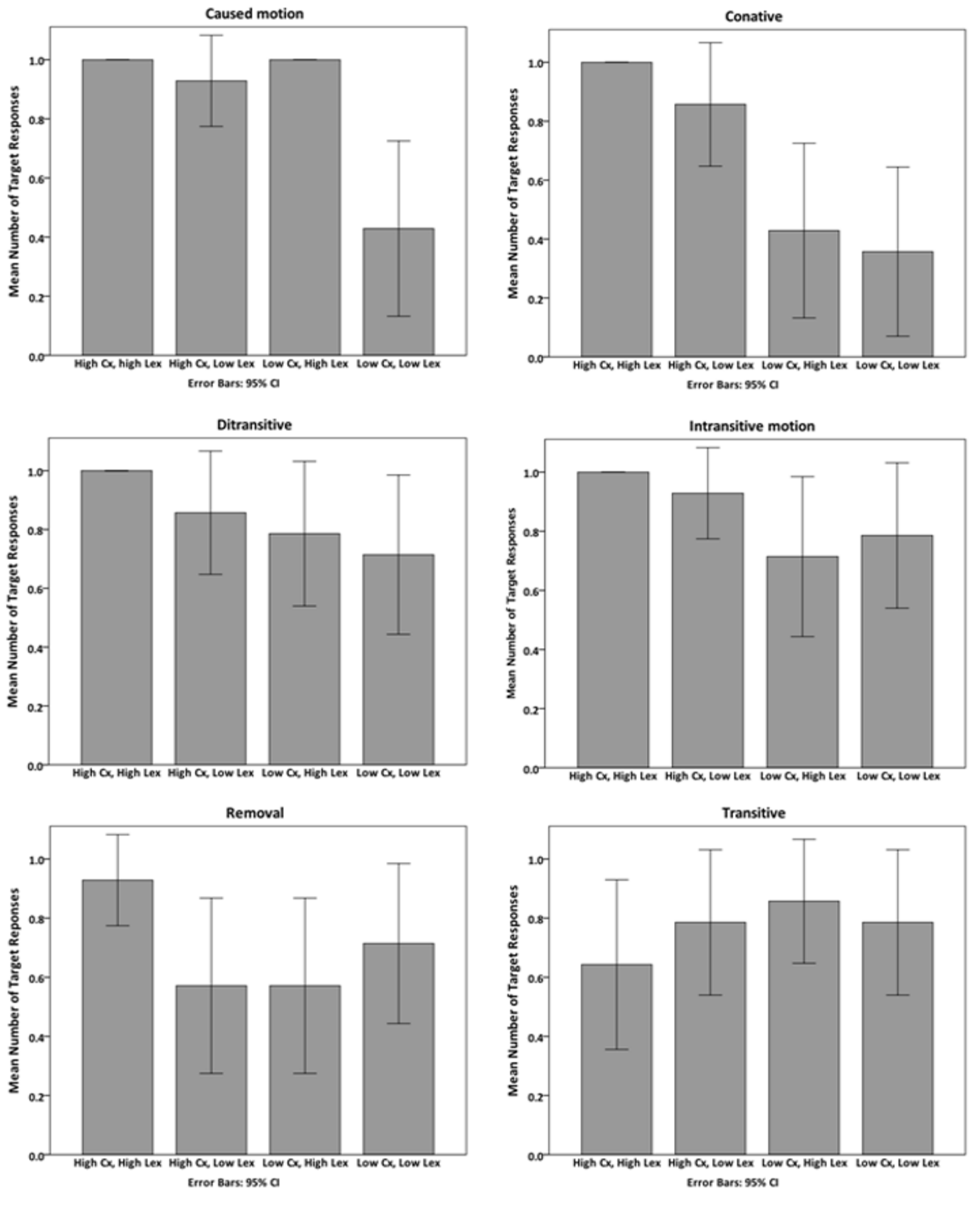


Figure 7.6 Mean number of target responses from participants with aphasia in each condition per construction in grammaticality judgement task

structure constructions in the task. Note that only one verb was included in each frequency condition per construction, so in this instance means were more informative than medians in reporting, despite the use of a non-parametric statistical test.

Table 7.20 Results from Wilcoxon signed-rank tests on number of target responses from participants with aphasia for argument structure constructions in Phase 2 grammaticality judgment task

	Construction Frequency						Lexical Frequency					
	high freq $\bar{x}$	low freq $\bar{x}$	*T*	*p*	*z*	*r*	high freq $\bar{x}$	low freq $\bar{x}$	*T*	*p*	*z*	*r*
Caused motion	1.93	1.43	0	0.008**	-2.65	-0.50	2.00	1.36	0	0.007**	-2.71	-0.51
Conative	1.86	0.79	6	0.008**	-2.67	-0.50	1.43	1.21	3	0.180	-1.34	-0.25
Ditransitive	1.86	1.50	0	0.025*	-2.24	-0.42	1.79	1.57	6	0.317	-1.00	-0.19
Intransitive motion	1.93	1.50	0	0.034*	-2.12	-0.40	1.71	1.71	5	1.000	0.00	0.00
Removal	1.50	1.29	3	0.180	-1.34	-0.25	1.50	1.29	9	0.380	-0.88	-0.17
Transitive	1.43	1.64	3	0.180	-1.34	-0.25	1.50	1.57	9	0.739	-0.33	-0.06

Note. Mean number of participants' target responses in high and low frequency conditions (maximum 2); Wilcoxon test statistic *T*; two-tailed *p*-value; *z*-score; effect size *r*. * $p < 0.05$. ** $p < 0.01$.

Results from the two Wilcoxon signed-rank tests on the main effects of construction frequency and lexical frequency on the number of target responses produced to each individual construction are shown in Table 7.20. Results from the full set of Wilcoxon signed-rank tests performed between conditions for the six constructions are available in Appendix K.

Caused motion

There were large, significant main effects of construction frequency ($T = 0.00$, $p = 0.008$, $r = -0.50$) and lexical frequency ($T = 0.00$, $p = 0.007$, $r = -0.51$) on the number of target responses produced to the caused motion construction. This was due to the fact that the number of target responses in the [low cx, low lex] condition was significantly lower than in the [high cx, high lex] condition ($T = 0.00$, $p = 0.005$, $r = -0.53$), the [high cx, low lex] condition ($T = 0.00$, $p = 0.008$, $r = -0.50$) and the [low cx, high lex] condition ($T = 0.00$, $p = 0.005$, $r = -0.53$).

Conative

There was a significant main effect of construction frequency on the number of target responses produced to the conative construction ($T = 6.00$, $p = 0.008$), and this represented a large effect ($r = -0.50$). The main effect of lexical frequency was not significant, nor was any interaction.

Ditransitive

There was a significant main effect of construction frequency on the number of target responses produced to the ditransitive construction ($T = 0.00$, $p = 0.025$), and this represented a medium effect ($r = -0.42$). This was due to a significantly greater number of target responses produced in the [high cx, high lex] condition compared to the [low cx, high lex] condition ($T = 0.00$, $p = 0.083$, $r = -0.33$) and the [low cx, low lex] condition ($T = 0.00$, $p = 0.046$, $r = -0.38$). The main effect of lexical frequency was not significant.

Intransitive motion

There was a significant main effect of construction frequency on the number of target responses produced to the intransitive motion construction ($T = 0.00$, $p = 0.034$), and this represented a medium effect ($r = -0.40$). This main effect was not evident in the difference between [high cx, low lex] and [low cx, low lex] conditions ($T = 2.50$, $p = 0.317$, $r = -0.19$). The main effect of lexical frequency was not significant.

Removal

There were no significant main effects of construction frequency or lexical frequency on the number of target responses produced to the removal construction. However, significantly more target responses were produced in the [high cx, high lex] condition compared to the [high cx, low lex] condition (T = 0.00, p = 0.025, r = -0.42) and the [low cx, high lex] condition (T = 0.00, p = 0.025, r = -0.42).

Transitive

There were no significant main effects of construction frequency or lexical frequency on the number of target responses produced to the transitive construction. There were no significant differences between any of the four conditions.

7.2.2 Response time results from participants with aphasia in grammaticality judgement task

7.2.2.1 Response time results from individual participants with aphasia in grammaticality judgement task

Table 7.21 shows median response times from participants with aphasia in the grammaticality judgement task. Only the times for participants' target responses were included in this dataset. That is, median response times reflect instances where participants responded 'yes' to an item intended to elicit a 'yes' judgement and 'no' to an item intended to elicit a 'no' judgement. Therefore, each cell in Table 7.21 reflects the median of a different number of values, depending on the number of target responses each participant produced, which is shown in Table 7.18.

7.2.2.1 Response time results from group of participants with aphasia in grammaticality judgement task

Response times from the group of participants with aphasia were analysed as described in Section 7.2.1.2 for the number of target responses. To explore main effects, participants' median response times to verbs in the four frequency conditions were collapsed into two groups in order to compare high and low frequency. Response times were collapsed in two ways in order to explore effects of construction frequency and lexical frequency. Two Wilcoxon signed-rank tests were used to compare median response times of the group of participants with aphasia to verbs with high and low construction frequency and lexical frequency, and effect sizes are reported as r. Response times from the group of participants with aphasia are shown in Figure 7.7.

Table 7.21 Median response times of target responses from participants with aphasia in grammaticality judgement task, shown in milliseconds

	All items	No judgements	Yes judgements	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
BF	3178 (11431)	3695 (11660)	2783 (11430)	1984 (927)	2079 (1236)	3616 (2081)	2538 (14173)
CJ	4985 (4341)	6697 (4722)	4218 (3577)	4258 (951)	2824 (1204)	3045 (2002)	4904 (3504)
GW	3307 (2695)	3975 (3134)	2988 (2137)	2458 (566)	2589 (450)	3182 (3896)	2829 (663)
HM	2140 (2117)	2901 (1763)	1916 (2423)	1315 (596)	1750 (535)	2106 (4385)	2495 (3837)
IC	4364 (3377)	8033 (3755)	3609 (1911)	2178 (1440)	2922 (801)	2815 (2623)	4188 (1841)
JF	3098 (4757)	4779 (5823)	2578 (1638)	1934 (668)	2171 (470)	3994 (2263)	2838 (780)
KT	3855 (3IC)	3854 (3886)	3953 (3318)	1834 (1567)	4557 (3085)	3788 (1783)	5809 (1499)
MJ	8572 (14IC)	16739 (21748)	7014 (5748)	6066 (2507)	5956 (3969)	7667 (3429)	7234 (3782)
PD	4935 (3525)	5220 (3419)	4500 (3665)	4022 (3372)	4009 (1354)	3953 (5772)	8439 (3499)
RE	5053 (4138)	5568 (3464)	4737 (4617)	3487 (1313)	4270 (2272)	4989 (1763)	5712 (1889)
SP	6387 (2938)	6595 (3094)	5939 (2590)	5939 (2036)	5666 (334)	4436 (1691)	9673 (-)
UT	4267 (2178)	4431 (2132)	4021 (2248)	4802 (1485)	3192 (3815)	3376 (354)	4437 (1325)
VH	4485 (1947)	6990 (2157)	4055 (665)	4006 (308)	4562 (936)	3821 (908)	(-) (-)
WD	8667 (7610)	10220 (9797)	8224 (4658)	5847 (5565)	12042 (6253)	10037 (3302)	4022 (2025)
Mdn	4425	5394	4038	3747	3601	3805	4437
M	4807	6407	4324	3581	4185	4345	5009

Note. Standard deviations in brackets. Alternating rows in grey for ease of reference. (-) indicates no data is available. Bottom rows show values of group median (Mdn) and mean (*M*).

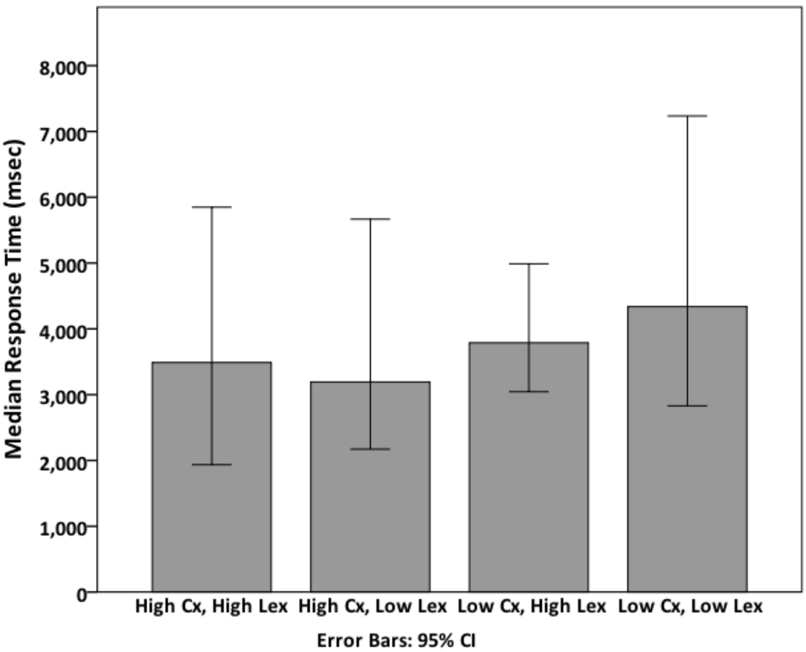


Figure 7.7 Median response times of participants with aphasia in Phase 2 grammaticality judgement task

Participants with aphasia showed significantly shorter response times to verbs with high construction frequency compared to verbs with low construction frequency (T = 13.00, two-tailed p = 0.013), and this represented a medium effect (r = -0.47). There was a numerical difference between response time to verbs with high lexical frequency and verbs with low lexical frequency, with shorter response times to verbs with high lexical frequency, but the difference did not reach significance (T = 36.00, two-tailed p = 0.300, r = -0.20).

Results from Wilcoxon signed-rank tests on participants' median response times between conditions in the task are shown in Table 7.22.

Table 7.22 Results from Wilcoxon signed-rank tests on response times from participants with aphasia between conditions in grammaticality judgement task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	-			
high cx, low lex	T = 33.00 p = 0.221 r = -0.23	-		
low cx, high lex	T = 25.00 p = 0.084 r = -0.33	T = 44.00 p = 0.594 r = -0.10	-	
low cx, low lex	T = 10.00 p = 0.013 r = -0.47	T = 13.00 p = 0.023 r = -0.43	T = 30.00 p = 0.279 r = -0.20	-

In sum, construction frequency had a significant main effect on response times from participants with aphasia in the grammaticality judgement task. There was no significant main effect of lexical frequency, and no significant interactions.

Power estimates indicated that with 14 participants and an alpha-level of 0.05, the study achieved 84% power to detect the effect of construction frequency, indicating it was fully powered for a parametric analysis. However, the study was underpowered to detect the effect of lexical frequency: to achieve 80% power at an alpha-level of 0.05, using the mean and standard deviations from current results, the required sample size was approximately 800 participants for parametric analysis. Shieh, Jan and Randles (2007) reported a sample size of approximately 10 as necessary to achieve 80% power at an alpha-level of 0.05 for the Wilcoxon-signed rank test to detect large group differences and sample sizes over 100 to detect small differences, with samples sizes rising to over 400 to detect the smallest differences between groups.

Given the results of the power analysis for a parametric approach, in conjunction with Shieh et al.'s observations, the current non-parametric analysis was likely fully powered to detect the large effect of construction frequency, but not the small effect of lexical frequency.

Notably, most researchers in behavioural science would likely consider the sample size necessary to detect an effect as small as lexical frequency was in this study to be infeasible.

7.2.2.2 Response time results from group of participants with aphasia in grammaticality judgement task by construction

The analyses described in Section 7.2.2.2 were performed separately for each of the six argument structure constructions in Phase 2. Results from the two Wilcoxon signed-rank tests on the main effects of construction frequency and lexical frequency are shown for each construction in Table 7.23 (over page). Results from the full set of Wilcoxon signed-rank tests comparing differences between each condition are provided in Appendix L. Figure 7.8 shows median response times in each condition for the six constructions in the task. Error bars were unable to be displayed for the conative construction because of the low number of responses in the [low cx, high lex] and [low cx, low lex] conditions.

Caused motion

There were no significant main effects of construction frequency or lexical frequency on response times to the caused motion construction. There were no significant differences between any conditions.

Conative

There were no significant main effects of construction frequency or lexical frequency on response times to the conative construction. There were no significant differences between any conditions.

Ditransitive

There were no significant main effects of construction frequency or lexical frequency on response times to the ditransitive construction. However, response times in the [low cx, low lex] condition were significantly longer than response times in the [low cx, high lex] condition ($T = 4.00$, $p = 0.050$, $r = -0.43$).

Intransitive motion

There was a significant main effect of construction frequency on response times to the intransitive motion construction ($T = 14.00$, $p = 0.028$), and this represented a medium effect ($r = -0.42$). This effect of construction frequency reached significance only for responses to verbs with low lexical frequency ($T = 0.00$, $p = 0.005$, $r = -0.57$).

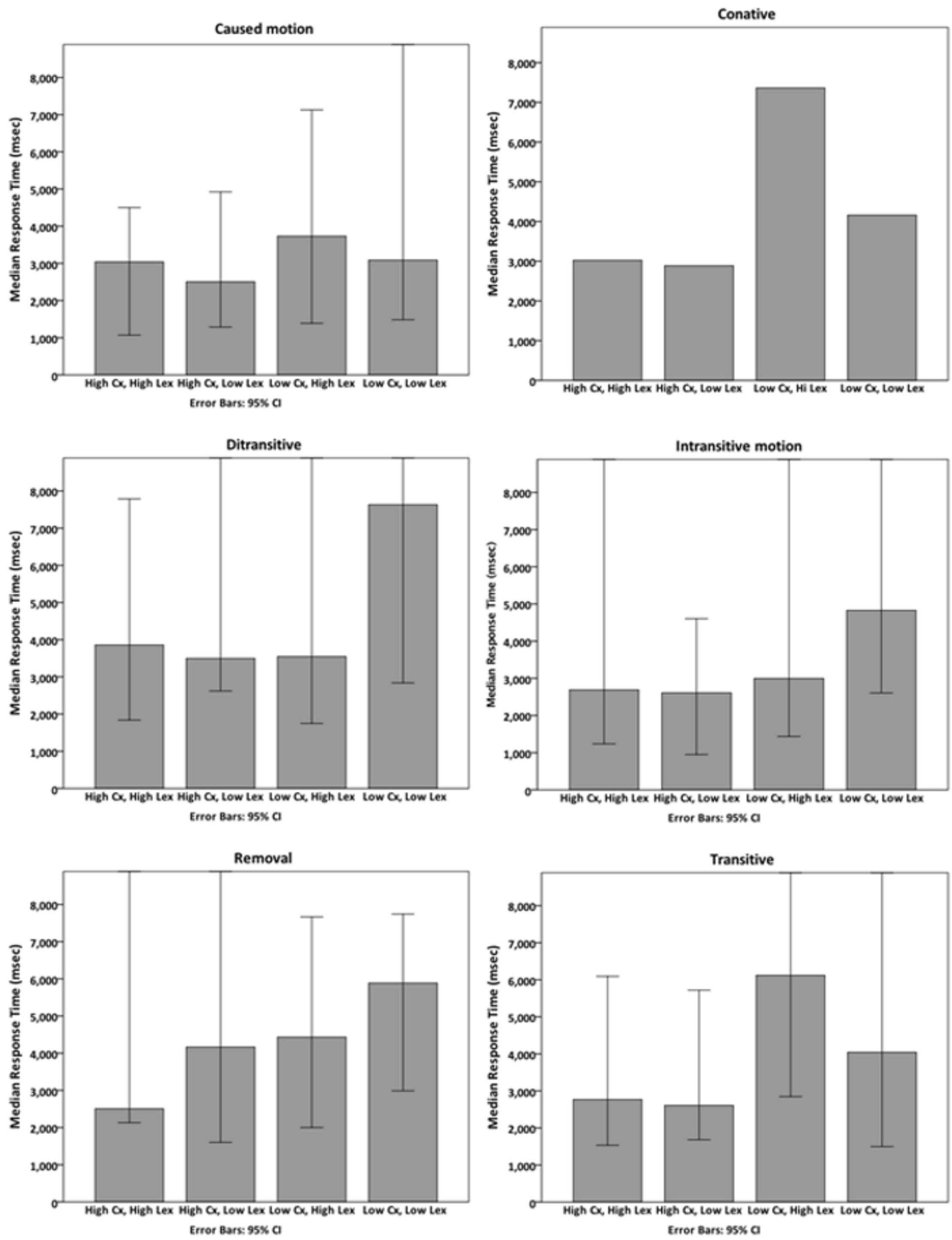


Figure 7.8 Median response times from participants with aphasia per condition for each construction in grammaticality judgement task

Table 7.23 Results from Wilcoxon signed-rank tests on median response times from participants with aphasia for argument structure constructions in Phase 2 grammaticality judgement task

	Construction Frequency						Lexical Frequency					
	high freq mdn	low freq mdn	*T*	*p*	*z*	*r*	high freq mdn	low freq mdn	*T*	*p*	*z*	*r*
Caused motion	3761	4661	43	0.551	-0.60	-0.16	4043	4738	25	0.152	-1.43	-0.28
Conative	3566	6551	6	0.176	-1.35	-0.29	4351	3747	22	0.101	-1.64	-0.31
Ditransitive	4884	5966	26	0.173	-1.36	-0.26	4839	5968	22	0.182	-1.33	-0.26
Intransitive motion	3800	RE2	14	0.028*	-2.20	-0.42	4618	5109	48	0.778	-0.28	-0.05
Removal	5164	5601	26	0.534	-0.62	-0.12	5418	5088	20	0.767	-0.30	-0.06
Transitive	3814	4909	25	0.272	-1.10	-0.22	4542	3710	29	0.722	-0.36	-0.07

Note. Mean of participants' median response times to verbs in high and low frequency conditions (maximum $n = 2$); Wilcoxon test statistic T; two-tailed p-value; z-score; effect size r. * $p < 0.05$.

Removal

There were no significant main effects of construction frequency or lexical frequency on response times to the removal construction. There were no significant differences between any conditions.

Transitive

There were no significant main effects of construction frequency or lexical frequency on response times to the transitive construction. There were no significant differences between any conditions.

7.2.2.3 Response time results from participants with aphasia in grammaticality judgement task by individual

Main effects of construction frequency and lexical frequency on response times was investigated for each participant with aphasia. Verbs from the four frequency conditions were collapsed into two groups, as described in Section 7.2.1.2, in order to analyse effects of construction frequency and lexical frequency within each participant with aphasia. Two Mann-Whitney tests were conducted for each participant, using the exact method due to the small number of data points. Results are shown in Table 7.24.

For ten of the 14 participants with aphasia, median response times to verbs with high construction frequency were shorter than to verbs with low construction frequency. This difference reached significance for HM and JF. For eight of the 14 participants with aphasia, median response times to verbs with high lexical frequency were shorter than to verbs with low lexical frequency. This difference did not reach significance for any of the eight.

Rahardja, Zhao and Qu (2009) observed that ‘there is no sample size requirement for the exact [Mann-Whitney] test to be valid’ (p. 317). However, Shieh et al. (2006) reported sample sizes between 16 and 20 as necessary to detect large differences between groups of equal sizes at 90% power using the non-exact (i.e. asymptotic) Mann-Whitney test. Sample sizes increased for groups with unequal sizes and smaller differences between groups (Shieh et al., 2006). In the grammaticality judgement task, sample sizes from participants with aphasia ranged from 12 to 24, and most participants had unequal group sizes due to the exclusion of non-target responses. This suggests that, while valid, the exact Mann-Whitney test may not have been powerful enough to detect differences in participants’ responses to high frequency and low frequency verbs in some instances. Indeed, 10 of the 14 participants with aphasia produced response times in the predicted direction for construction frequency, and this represented a medium effect size for most of these participants; however, the

Table 7.24 Results from Mann-Whitney tests on response times from individual participants with aphasia in Phase 2 grammaticality judgement task

	Construction Frequency							Lexical Frequency					
	n	high freq mdn	low freq mdn	*U*	*p*	*z*	*r*	high freq mdn	low freq mdn	*U*	*p*	*z*	*r*
BF	21	2009	3136	33	0.074	-1.49	-0.33	2910	2161	48	0.327	-0.49	-0.11
CJ	21	4129	4058	52	0.432	-0.21	-0.05	4129	4058	51	0.431	-0.21	-0.05
GW	23	2589	2988	44	0.095	-1.35	-0.28	3002	2672	57	0.304	-0.55	-0.02
HM	20	1619	2373	26	0.049*	-1.70	-0.38	1415	1916	39	0.218	-0.83	-0.19
IC	24	2500	3609	55	0.174	-0.98	-0.20	2178	3539	46	0.072	-1.50	-0.31
JF	18	2047	2838	14	0.009**	-2.34	-0.55	2512	2UT	38	0.448	-0.18	-0.04
KT	17	2120	4749	19	0.067	-1.56	-0.38	2532	4749	27	0.237	-0.78	-0.19
MJ	21	5956	7667	33	0.074	-1.49	-0.33	6547	6459	54	0.486	-0.07	-0.02
PD	21	4009	6438	33	0.066	-1.55	-0.34	3953	4522	44	0.234	-0.78	-0.17
RE	20	3841	5531	33	0.135	-1.16	-0.26	3761	4661	42	0.289	-0.61	-0.14
SP	10	5902	5632	10	0.500	-0.11	-0.04	5632	5902	8	0.333	-0.57	-0.18
UT	16	4412	3489	23	0.246	-0.76	-0.19	3812	3765	64	0.360	-0.42	-0.11
VH	12	4104	3821	13	0.500	-0.09	-0.03	3914	4562	11	0.230	-0.85	-0.25
WD	17	7958	9282	28	0.240	-0.77	-0.19	8924	5422	24	0.202	-0.91	-0.22

Note. Number of items in each analysis (*n*); median verb production latencies in high and low frequency conditions in milliseconds; Mann-Whitney test statistic *U*; one-tailed *p*-value calculated via exact method; *z*-score; effect size *r*. * $p < 0.05$. ** $p < 0.01$.

difference reached significance for only two participants (see Table 7.24). Future research could increase the number of items included in within-participant comparisons to address this issue.

7.2.3 Comparison of participants with aphasia to typical participants in grammaticality judgement task

Section 4.4.2.2 explained how comparisons of participants with aphasia to typical participants were made using Bayesian Standardised Difference Tests (BSDTs) and effect scores. To review, BSDTs compare the difference between scores in two tasks obtained by a clinical participant to the difference in scores obtained by the typical population. In the present context, 'tasks' were defined as responses to high and low frequency verbs. For each participant with aphasia, a BSDT was carried out once to investigate a difference in construction frequency, and again to investigate a difference in lexical frequency. BSDTs were conducted to examine differences in response times between participants with aphasia and typical participants.

The BSDT does assume that the control distribution is normally distributed. Crawford, Garthwaite, Azzalini, Howell and Laws (2006) recommended that transformations be used to normalised the control distribution before applying their methods, if possible. An inverse transformation was found to successfully normalise the distribution of response time data from typical participants. Following Crawford et al. (2006), this transformation was applied to response time data from both typical participants and participants with aphasia before the BSDT was computed. Results from both the raw data and transformed data are reported below.

Differences in response times to verbs with high and low frequency are represented as effect scores, which were calculated on responses to construction frequency and lexical frequency. To recap, effect scores represent differences in response times to low and high frequency items as a proportion, with positive values indicating effects in the predicted direction and negative values indicating effects in the reverse direction. For example, an effect score of 0.5 represents response times that were 50% longer to low frequency verbs than high frequency verbs.

Effect scores are shown in Figure 7.9 (over page), which shows effect scores for construction frequency in dark-coloured bars and effect scores for lexical frequency in lighter-coloured bars for each participant with aphasia, in addition to the effect score from the group of typical participants. Asterisks denote significant results from Bayesian Standardised Difference Tests. Full results from all BSDTs are available in Appendix M.

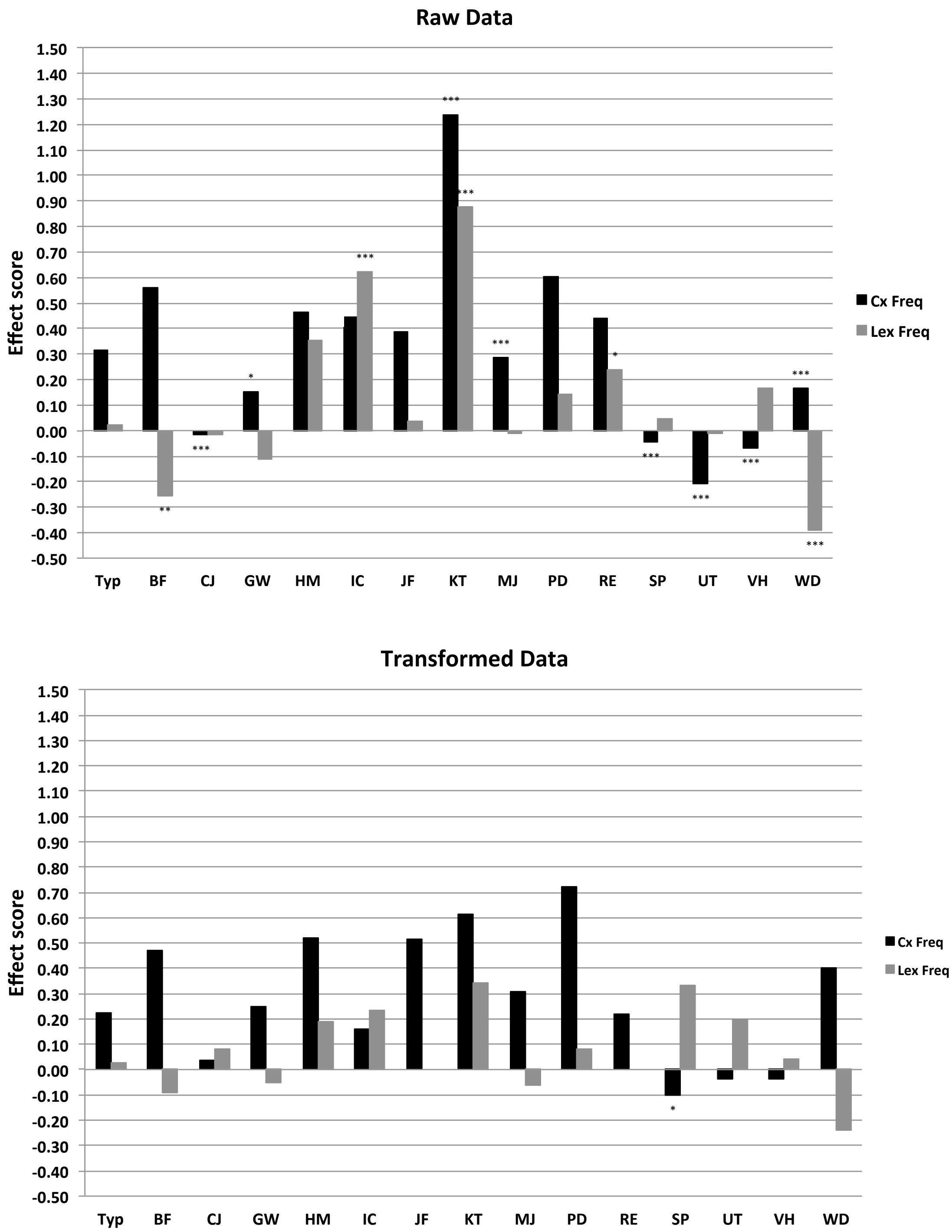


Figure 7.9 Effect scores in grammaticality judgement task

Note. Effect scores for typical participants shown on left (Typ). Asterisks denote significant results from Bayesian Standardised Difference Tests.

* $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

Results from analyses of raw data are shown in the upper panel, and results from analyses of transformed data are shown in the lower panel.

Construction frequency

Based on results from the raw data, eight of the 14 participants with aphasia performed significantly differently from typical participants. GW, KT, MJ and WD all produced longer response times to verbs with low construction frequency than high construction frequency: their effect scores were in the predicted direction. KT showed a greater effect of construction frequency than typical participants, and GW, MJ and WD showed smaller effects of construction frequency than typical participants. CJ, SP, UT and VH showed a reverse effect of construction frequency, as they produced longer response times to verbs with high construction frequency than low construction frequency.

In the transformed data, only SP's reverse effect of construction frequency remained significant. CJ did not show the reverse effect of construction frequency that he did in the raw data. Upon inspection, the raw data showed his median response times to verbs in the two conditions were very close, with a difference of only 71 milliseconds (see Table A26 in Appendix M). In this instance, the mean of transformed data likely represents a more accurate summary of his performance than medians. Additionally, GW, MJ and WD, who in the raw data demonstrated smaller effects of construction frequency than typical participants, in the transformed data showed larger effects. The size of their effects compared to typical participants was not consistent across analyses and therefore cannot be interpreted as reliable. In contrast, the effect scores of KT, UT and VH remained in the same direction compared to typical participants in the transformed data as in the raw data, but did not reach significance.

Lexical frequency

Based on results from the raw data, five participants with aphasia performed significantly differently from typical participants. IC, KT and RE showed an effect of lexical frequency in the predicted direction, and their effect scores were greater than typical participants. BF and WD showed a reverse effect of lexical frequency, producing longer response times to verbs with high lexical frequency than low lexical frequency.

None of these participants with aphasia were significantly different from typical participants in the analysis of the transformed data. However, IC and KT also showed greater effects of lexical frequency compared to typical participants, and BF and WD also showed reverse effects of lexical frequency compared to typical participants in the transformed data, as they did in the raw data.

7.2.4 Summary of results from participants with aphasia in grammaticality judgement task

This section provides a summary of findings from participants with aphasia, in addition to a brief comparison to results from typical participants. Further discussion of differences between typical participants and participants with aphasia in provided in Chapter 9.

Number of target responses

Group-level analyses of participants with aphasia revealed a large and significant main effect of construction frequency on the number of target responses they produced in the task. This main effect of construction frequency was evident in the analysis of the caused motion, conative, ditransitive and intransitive motion constructions. The main effect of lexical frequency on the number of target responses approached significance, but it was evident only in the analysis of the caused motion construction. Pairwise tests indicated lexical frequency had a significant effect on responses to verbs with high construction frequency but not low construction frequency. Only responses to the removal construction showed this interaction.

Though the number of target responses from typical participants was not subject to group-level analysis, because most participants performed at or near ceiling in the grammaticality judgement task, typical participants produced the lowest number of target responses to the [low cx, high lex] condition. In contrast, participants with aphasia produced the lowest number of target responses to the [low cx, low lex] condition.

Response times

Group-level analyses of response time data from participants with aphasia revealed a significant main effect of construction frequency. This effect of construction frequency reached significance only in the analysis of the intransitive motion construction. There was no significant effect of lexical frequency or interactions on response times from participants with aphasia.

At the group-level, participants with aphasia performed similarly to typical participants in showing the main effect of construction frequency but not lexical frequency; however, the typical interaction was not evident in the analysis of response times from participants with aphasia. Lexical frequency had an effect on response times from typical participants in the predicted direction for verbs with high construction frequency but in the reverse direction for verbs with low construction frequency. In contrast, lexical frequency had an effect on response times from participants with aphasia in the predicted direction for verbs with both high and low construction frequency.

Individuals with aphasia

Within-individual analyses demonstrated that 10 of the 14 participants produced shorter response times to verbs with high construction frequency than verbs with low construction frequency. This difference reached significance for HM and JF. A total of eight participants with aphasia produced shorter response times to verbs with high lexical frequency than verbs with low lexical frequency, but this difference did not reach significance for any individual.

Results from Bayesian Standardised Difference Tests on raw data demonstrated that KT showed a greater effect of construction frequency compared to typical participants; GW, MJ and WD showed smaller effects of construction frequency compared to typical participants; and CJ, SP, UT and VH showed reverse effects of construction frequency. IC, KT and RE showed greater effects of lexical frequency compared to typical participants, and BF and WD showed reverse effects of lexical frequency. However, only the performance of SP remained significantly different from typical participants in the analysis of the transformed data. These participants will be discussed in more detail in the General Discussion in Chapter 9.

*

Results from the Phase 2 grammaticality judgement task confirm the importance of construction frequency on participants' performance in the grammaticality judgement task, for both typical adults and adults with aphasia. Results from the sentence completion task follow in Chapter 8.

8 Results from Phase 2 sentence completion task

This chapter reports results from the experimental sentence completion task in Phase 2. The dependent variable for all participants was their verb production latency, or the time between the onset of a written verb on screen and the production of that verb in a sentence context, preceded by a subject pronoun. Analyses of the number of target responses from participants with aphasia were also performed. Results from typical participants are reported in Section 8.1, followed by results from participants with aphasia in Section 8.2. Findings from the grammaticality judgement task and sentence completion task are compared in the General Discussion of Chapter 9.

8.1 Results from typical participants

The dataset for the sentence completion task contained responses from 87 typical participants in Phase 2. The sentence completion task contained 24 verbs, as a set of four verbs was associated with each of the six argument structure constructions included in the task. A different set of verbs was used in the sentence completion task than the grammaticality judgement task. The total number of data points in the complete dataset was 2088.

The type of analyses performed on data from the grammaticality judgement task described in Chapter 7 was also performed on data from the sentence completion task, reported below.

8.1.1 Screening and cleaning procedures for data from typical participants

Section 4.4.1.2 reviewed the reasons why data was excluded from analysis of the sentence completion task. Section 8.1.1.1 reports the number of responses excluded from analysis.

8.1.1.1 Excluded and missing data

Four types of responses were excluded from analyses due to participant behaviour: (1) productions that included self-corrections or comments which affected the measurement of the verb production latency; (2) productions that contained non-target constructions; (3) productions in which the articulation of the verb was incomplete due to the termination of the audio recording; and (4) no response. A total of 19 of these response types were identified from 16 typical participants, totalling 0.01% of the 2088 data points in the dataset. Ten of these 19 responses were produced to verbs in the [low cx, high lex] frequency condition.

In addition to the responses excluded due to participant behaviour, some responses were excluded due to experimental error. Trials with start times of less than 400 milliseconds were excluded from analyses, because the Praat script used to automatise the extraction of

Table 8.1 Number of excluded responses from typical participants to verbs in each condition of sentence completion task

Frequency condition	Responses excluded due to participant behaviour	Responses excluded due to experimental error	Total number of excluded responses
high cx, high lex	1	40	41
high cx, low lex	4	8	12
low cx, high lex	10	21	31
low cx, low lex	4	15	19
TOTAL	19	84	103

Table 8.2 Total number of excluded responses from typical participants to verbs in sentence completion task as a function of construction and frequency condition

	high cx, high lex		high cx, low lex		low cx, high lex		low cx, low lex	
Caused motion	LEAVE	12	PASS	2	TELL	10	DRIVE	3
Conative	TALK	1	SHOUT	1	MOVE	8	SING	2
Ditransitive	GIVE	4	OWE	4	PAY	2	SLIP	1
Intransitive motion	GO	3	WALK	1	GET	6	CUT	2
Removal	TAKE	20	BUY	1	FEEL	3	REACH	7
Transitive	SEE	1	LOVE	3	USE	2	PICK	4

Note. *N* = 87.

verb production latencies inaccurately identified the start of participants’ speech in these instances (see Section 4.4.1.2).

Table 8.1 shows the number of responses excluded from each condition due to participant behaviour and experimental error. A total of 103 responses were excluded, totalling 4.93% of the 2088 data points in the dataset. Table 8.2 shows the total number of excluded responses to each verb in the sentence completion task.

One participant was excluded from further analysis of the sentence completion task because only three of his responses remained in the dataset after the exclusion criteria were applied. The exclusion of this participant decreased the total number of excluded responses in the dataset from 103 to 82. The total number of data points in the dataset after the removal of this participant was 1981.

In order to ensure that results were not affected by missing values due to excluded data, ANOVAs were carried out on data from the subset of participants who had no excluded data, in addition to the entire group of participants.

8.1.1.2 Transformation of the data

Verb production latencies were subject to an inverse transformation in order to correct the non-normality of the ex-Gaussian response time distribution. Values in the dataset were transformed, and then averaged over participants within conditions. ANOVAs are reported for both raw and transformed versions of the dataset.

Figure 8.1 shows the distribution of verb production latencies from the 86 typical participants in the raw data (left pane) and transformed data (right pane). The distribution of the raw data was significantly different from normal by the Kolmogorov-Smirnov test ($D(86) = 0.13$, $p = 0.001$), but the distribution of the transformed data was not ($D(86) = 0.06$, $p = 0.200$).

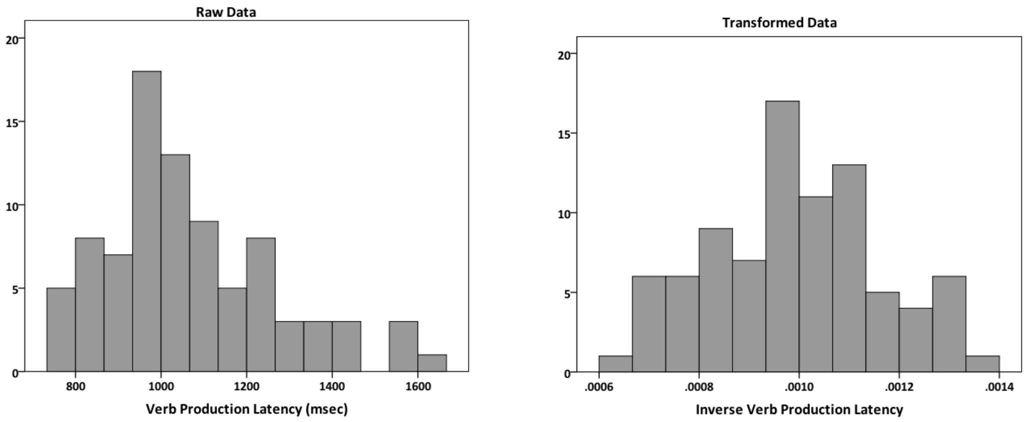


Figure 8.1 Distribution of raw and transformed verb production latencies from typical participants in sentence completion task

The following sections report analyses of verb production latencies. Section 8.1.2 reports verb production latencies to each verb in the sentence completion task. Section 8.1.3 reports the results of ANOVAs on data from the sentence completion task, including ANOVAs on data from the subset of 36 participants who had no excluded responses in the task and ANOVAs on data from all 86 typical participants included in the analysis. ANOVAs were carried out by-participant and by-item, permitting the calculation of Clark's (1973) minF'. Section 8.1.4 reports results of ANOVAs carried out on data from each of the six argument structure constructions individually. Finally, Section 8.1.5 reports results from additional analyses which investigated whether participants' responses were affected by other variables relating to the participants or items included in the task.

8.1.2 Responses to verbs in sentence completion task from typical participants

The mean verb production latency for each verb in the sentence completion task is shown in Table 8.3 (over page). Values were averaged over raw data from the 86 typical participants included in the analyses.

8.1.3 ANOVAs on data from sentence completion task

This section reports results from ANOVAs that were performed on data from the sentence completion task. A repeated-measures, by-participant ANOVA was carried out on

Table 8.3 Mean verb production latencies for verbs in sentence completion task as a function of construction and frequency condition, based on raw data in milliseconds

	high cx, high lex		high cx, low lex		low cx, high lex		low cx, low lex	
Caused motion	LEAVE	1075 (290)	PASS	1068 (273)	TELL	1089 (329)	DRIVE	1048 (259)
Conative	TALK	1067 (286)	SHOUT	1036 (282)	MOVE	1097 (348)	SING	1056 (289)
Ditransitive	GIVE	1005 (252)	OWE	1221 (304)	PAY	1073 (291)	SLIP	1066 (241)
Intransitive motion	GO	982 (193)	WALK	1086 (263)	GET	1074 (291)	CUT	1109 (262)
Removal	TAKE	1023 (246)	BUY	1061 (233)	FEEL	1092 (280)	REACH	1083 (252)
Transitive	SEE	1082 (298)	LOVE	1086 (244)	USE	1022 (266)	PICK	1053 (226)
Mean (*SD*)		1039 (41)		1093 (65)		1075 (28)		1069 (23)

Note. Standard deviation of means in brackets. Non-target responses excluded from reporting. Alternating rows in grey for ease of reference. *N* = 86.

Table 8.4 Pearson's correlation coefficients for verb production latencies from subset of typical participants in each condition of sentence completion task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
Raw values				
high cx, high lex	-			
high cx, low lex	0.87**	-		
low cx, high lex	0.91**	0.78**	-	
low cx, low lex	0.88**	0.81**	0.89**	-
Inverse values				
high cx, high lex	-			
high cx, low lex	0.88**	-		
low cx, high lex	0.90**	0.82**	-	
low cx, low lex	0.86**	0.81**	0.88**	-

Note. ** $p < 0.01$ (one-tailed). $n = 36$.

Table 8.5 Summary statistics for by-participant analyses of verb production latencies from participants with no excluded responses in sentence completion task

	M	95% CI		*SD*	min	max
		Lower bound	Upper bound			
Raw data						
high cx, high lex	1023	965	1081	171	777	1577
high cx, low lex	1108	1045	1172	187	793	1505
low cx, high lex	1054	990	1118	190	782	1617
low cx, low lex	1069	1010	1127	173	823	1618
Back transformations						
high cx, high lex	978	934	1027	-	768	1473
high cx, low lex	1057	1004	1117	-	782	1493
low cx, high lex	1001	949	1060	-	750	1502
low cx, low lex	1027	978	1081	-	802	1512

Note. Mean (*M*), lower and upper bound of 95% confidence interval (CI), standard deviation (*SD*), minimum value (min) and maximum value (max). $n = 36$.

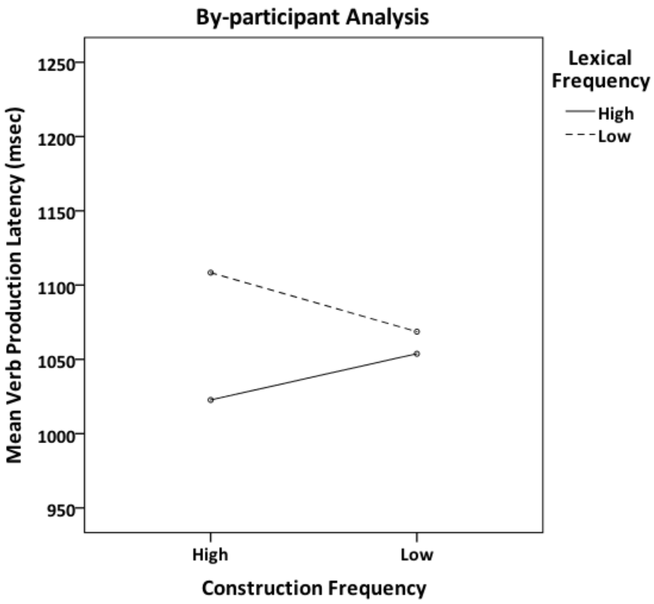


Figure 8.2 Interaction graph for by-participant analysis of raw data from subset of typical participants with no excluded responses in sentence completion task

production latencies to verbs in each of the four frequency conditions. A by-item ANOVA was also carried out on production latencies to verbs in each of the four conditions. Clark's (1973) minF' was then calculated to examine the effects of construction frequency, lexical frequency and their interaction when both participants and items were treated as random effects. Throughout this section, effect sizes are reported as Pearson's correlation coefficient *r*.

8.1.3.1 ANOVAs on data from participants with no excluded responses

36 typical participants had no data excluded from the analysis of the sentence completion task. This dataset contained a total of 864 data points. Figure 8.2 shows the interaction graph from the by-participant ANOVA on raw data.

By-participant ANOVAs

Pearson's correlation coefficients for participants' response times among conditions in the sentence completion task are shown in Table 8.4, from the 36 participants included in this analysis. Significant correlations of high magnitude were observed between participants' response times among all conditions, underpinning the repeated-measures design for both raw and transformed data.

Table 8.5 shows the summary statistics for the by-participant analyses of the dataset containing verb production latencies from participants with no excluded responses in the sentence completion task. Statistics for the raw dataset are shown in the upper panel, in milliseconds, and the values from the back transformations are shown below.

A repeated-measures ANOVA on the raw data revealed a significant main effect of lexical frequency ($F(1, 35) = 17.50$, $p < 0.001$), indicating that participants produced verbs with high lexical frequency with significantly shorter latencies than verbs with low lexical frequency. This represented a large effect ($r = 0.58$). The main effect of construction frequency was not significant ($F(1, 35) = 0.11$, $p = 0.745$, $r = 0.06$). The interaction between construction frequency and lexical frequency was significant ($F(1, 35) = 14.94$, $p < 0.001$) and represented a large effect ($r = 0.55$). The interaction reflected the fact that there was a large, significant effect of lexical frequency on the production of verbs with high construction frequency ($t(35) = -5.49$, $p < 0.001$, $r = 0.68$), but no significant effect of lexical frequency on the production of verbs with low construction frequency ($t(35) = -1.02$, $p = 0.317$, $r = 0.17$). Results of the full set of post-hoc *t*-tests are shown in Table 8.6.

A repeated-measures ANOVA on the transformed data revealed a significant main effect of lexical frequency ($F(1, 35) = 27.68$, $p < 0.001$), indicating that participants produced verbs with high lexical frequency with significantly shorter latencies than verbs with low lexical frequency. This represented a large effect ($r = 0.66$). The main effect of construction frequency was not significant ($F(1, 35) = 0.04$, $p = 0.837$, $r = 0.03$). The interaction between construction frequency and lexical frequency was significant ($F(1, 35) = 10.03$, $p = 0.003$) and represented a moderate effect size ($r = 0.47$). The interaction reflected the fact that lexical frequency had a significant effect on the production of verbs with high construction frequency ($t(35) = 6.25$, $p < 0.001$), and this represented a large effect ($r = 0.73$). In contrast, the effect of lexical frequency on the production of verbs with low construction frequency was not significant ($t(35) = 1.93$, $p = 0.062$) and represented a medium effect size ($r = 0.31$). Results of the full set of post-hoc *t*-tests are shown in Table 8.7.

By-item ANOVAs

Summary statistics for the by-item analyses of verb production latencies from typical participants with no excluded responses in the sentence completion task are available in Appendix N. A by-item ANOVA on the raw data revealed a significant main effect of lexical frequency ($F(1, 20) = 4.75$, $p = 0.041$), indicating that verbs with high lexical frequency elicited significantly shorter production latencies than verbs with low lexical frequency. This represented a moderate effect size ($r = 0.44$). The main effect of construction frequency was not significant ($F(1, 20) = 0.03$, $p = 0.864$, $r = 0.04$).

Table 8.6 Post-hoc paired-samples *t*-tests on raw data from typical participants with no excluded responses in the sentence completion task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	-			
high cx, low lex	$t(35) = -5.49$ $p < 0.001$ $r = 0.68$	-		
low cx, high lex	$t(35) = -2.38$ $p = 0.023$ $r = 0.37$	$t(35) = 2.62$ $p = 0.013$ $r = 0.40$	-	
low cx, low lex	$t(35) = -3.25$ $p = 0.003$ $r = 0.48$	$t(35) = 2.14$ $p = 0.039$ $r = 0.34$	$t(35) = -1.02$ $p = 0.317$ $r = 0.17$	-

Note. Significance at $p = 0.008$ with Bonferroni correction (0.05 ÷ 6 = 0.008).

Table 8.7 Post-hoc paired-samples *t*-tests on transformed data from typical participants with no excluded responses in the sentence completion task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	-			
high cx, low lex	$t(35) = 6.25$ $p < 0.001$ $r = 0.73$	-		
low cx, high lex	$t(35) = 2.00$ $p = 0.053$ $r = 0.32$	$t(35) = -3.33$ $p = 0.002$ $r = 0.49$	-	
low cx, low lex	$t(35) = 3.81$ $p = 0.001$ $r = 0.54$	$t(35) = -1.87$ $p = 0.070$ $r = 0.30$	$t(35) = 1.93$ $p = 0.062$ $r = 0.31$	-

Note. Significance at $p = 0.008$ with Bonferroni correction (0.05 ÷ 6 = 0.008).

The interaction between construction frequency and lexical frequency was not significant ($F(1, 20) = 2.44$, $p = 0.134$, $r = 0.33$).

A by-item ANOVA on the transformed data revealed a significant main effect of lexical frequency ($F(1, 20) = 8.25$, $p = 0.009$), indicating that verbs with high lexical frequency elicited significantly shorter production latencies than verbs with low lexical frequency. This represented a large effect ($r = 0.54$). The main effect of construction frequency was not significant ($F(1, 20) = 0.01$, $p = 0.910$, $r = 0.02$). The interaction between construction frequency and lexical frequency was not significant ($F(1, 20) = 2.17$, $p = 0.156$, $r = 0.31$).

minF'

Cohen's (1973) minF' was calculated in order to examine the effect of lexical frequency, construction frequency and their interaction when both participants and items were treated as random effects. In the raw data, the main effect of lexical frequency was approaching significance (minF'(1, 39.42) = 3.74, $p < 0.10$) and represented a small effect ($r = 0.29$). The main effect of construction frequency was not significant (minF'(1, 38.88) = 0.02, p = ns, $r = 0.02$), and the interaction between construction frequency and lexical frequency was not significant (minF'(1, 32.64) = 2.10, p = ns, $r = 0.25$). In the transformed data, lexical frequency yielded a significant main effect (minF'(1, 42.33) = 6.36, $p < 0.025$) with a moderate effect size ($r = 0.36$). The main effect of construction frequency was not significant (minF'(1, 37.50) = 0.01, p = ns, $r = 0.02$), and the interaction between construction frequency and lexical frequency was not significant (minF'(1, 35.64) = 1.78, p = ns, $r = 0.22$).

Interim summary

All analyses of data from the subset of 36 typical participants who had no excluded responses in the sentence completion task revealed a significant main effect of lexical frequency, indicating that verb production latencies were significantly shorter for verbs with high lexical frequency than for verbs with low lexical frequency. By-participant analyses identified a significant interaction between construction frequency and lexical frequency, which showed that lexical frequency had a significant effect on the production of verbs with high construction frequency, but not low construction frequency. The main effect of construction frequency was not significant in any analysis.

8.1.3.2 ANOVAs on data from all participants

Analyses in this section were performed on data from the 86 typical participants included in the sentence completion task. After accounting for the responses that were excluded from analyses, which were described in Section 8.1.1, the dataset contained a total of 1981 data points. Figure 8.3 shows the results of the by-participant analysis of raw data in the form of the interaction graph from the ANOVA.

By-participant ANOVAs

Table 8.8 shows Pearson's correlation coefficients for participants' verb production latencies among the four conditions in the sentence completion task. Significant correlations of high magnitude were observed between verb production latencies among all conditions, supporting the repeated-measures design for both raw and transformed data.

Table 8.9 (over page) shows the summary statistics for the by-participant analyses of the dataset containing responses from the 86 typical participants included in the analysis of the sentence completion task.

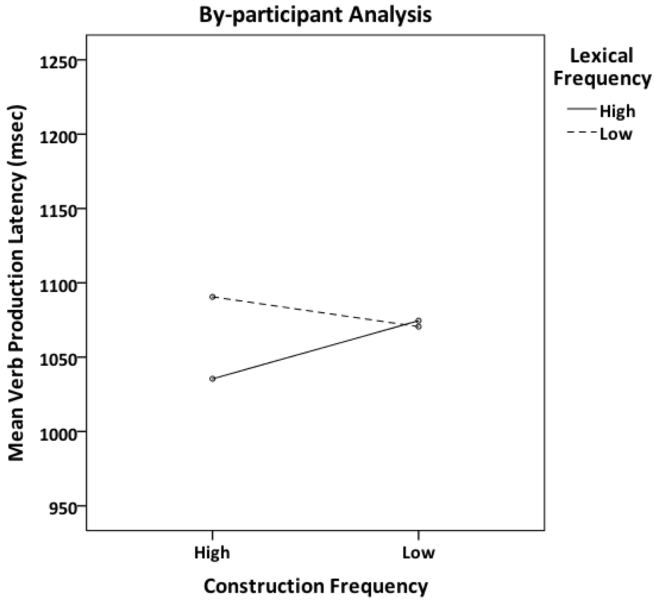


Figure 8.3 Interaction graph for by-participant analysis of raw data from all typical participants included in sentence completion task

Table 8.8 Pearson's correlation coefficients for verb production latencies from typical participants in each condition of sentence completion task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
Raw values				
high cx, high lex	-			
high cx, low lex	0.89**	-		
low cx, high lex	0.88**	0.79**	-	
low cx, low lex	0.90**	0.89**	0.88**	-
Inverse values				
high cx, high lex	-			
high cx, low lex	0.90**	-		
low cx, high lex	0.89**	0.82**	-	
low cx, low lex	0.89**	0.88**	0.88**	-

Note. ** $p < 0.01$ (one-tailed). $N = 86$.

A repeated-measures ANOVA on the raw data revealed a small but significant main effect of lexical frequency ($F(1, 85) = 8.07$, $p = 0.006$, $r = 0.29$), indicating that participants produced verbs with high lexical frequency with significantly shorter latencies than verbs with low lexical frequency. The main effect of construction frequency was not significant ($F(1, 85) = 1.12$, $p = 0.293$, $r = 0.11$). The interaction between construction frequency and lexical frequency was significant ($F(1, 85) = 22.04$, $p < 0.001$) and represented a medium effect size ($r = 0.45$). The interaction indicated that lexical frequency had a significant

Table 8.9 Summary statistics for by-participant analyses of verb production latencies from typical participants in sentence completion task

	M	95% CI		*SD*	min	max
		Lower bound	Upper bound			
Raw data						
high cx, high lex	1035	993	1078	197	700	1708
high cx, low lex	1091	1046	1135	209	723	1653
low cx, high lex	1075	1027	1122	222	694	1625
low cx, low lex	1071	1026	1115	207	742	1623
Back transformations						
high cx, high lex	981	947	1018	-	697	1610
high cx, low lex	1032	993	1074	-	721	1589
low cx, high lex	1005	965	1048	-	674	1567
low cx, low lex	1017	979	1057	-	737	1585

Note. Mean (*M*), lower and upper bound of 95% confidence interval (CI), standard deviation (*SD*), minimum value (min) and maximum value (max). $N = 86$.

Table 8.10 Post-hoc paired-samples *t*-tests on raw data from all typical participants included in the sentence completion task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	-			
high cx, low lex	$t(85) = -5.38$ $p < 0.001$ $r = 0.50$	-		
low cx, high lex	$t(85) = -3.49$ $p = 0.001$ $r = 0.35$	$t(85) = 1.06$ $p = 0.292$ $r = 0.11$	-	
low cx, low lex	$t(85) = -3.59$ $p = 0.001$ $r = 0.36$	$t(85) = 1.86$ $p = 0.066$ $r = 0.20$	$t(85) = 0.34$ $p = 0.731$ $r = 0.04$	-

Note. Significance at $p = 0.008$ with Bonferroni correction ($0.05 \div 6 = 0.008$).

Table 8.11 Post-hoc paired-samples *t*-tests on transformed data from all typical participants included in the sentence completion task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	-			
high cx, low lex	$t(85) = 5.88$ $p < 0.001$ $r = 0.54$	-		
low cx, high lex	$t(85) = 2.53$ $p = 0.013$ $r = 0.26$	$t(85) = -2.16$ $p = 0.034$ $r = 0.23$	-	
low cx, low lex	$t(85) = 4.03$ $p < 0.001$ $r = 0.40$	$t(85) = -1.56$ $p = 0.122$ $r = 0.17$	$t(85) = 1.13$ $p = 0.260$ $r = 0.12$	-

Note. Significance at $p = 0.008$ with Bonferroni correction ($0.05 \div 6 = 0.008$).

effect on the production of verbs with high construction frequency in the predicted direction ($t(85) = -5.38$, $p < 0.001$), i.e. verbs with high lexical frequency were produced with shorter latencies than verbs with low lexical frequency, and this represented a large effect ($r = 0.50$). However, lexical frequency had a non-significant effect on the production latency of verbs with low construction in the reverse direction ($t(85) = 0.34$, $p = 0.731$, $r = 0.04$), i.e. verbs with high lexical frequency were produced with longer latencies than verbs with low lexical frequency. Results of the full set of post-hoc *t*-tests are shown in Table 8.10.

A repeated-measures ANOVA on the transformed data revealed a significant main effect of lexical frequency ($F(1, 85) = 17.17$, $p < 0.001$), indicating that participants produced verbs with high lexical frequency with significantly shorter latencies than verbs with low lexical frequency. This represented a moderate effect size ($r = 0.41$). The main effect of construction frequency was not significant ($F(1, 85) = 0.39$, $p = 0.535$, $r = 0.07$). The interaction between construction frequency and lexical frequency was significant ($F(1, 85) = 11.59$, $p = 0.001$) and represented a medium effect size ($r = 0.35$). The interaction indicated that lexical frequency had a significant effect on the production of verbs with high construction frequency ($t(85) = 5.88$, $p < 0.001$), and this represented a large effect ($r = 0.54$). However, there was no significant effect of lexical frequency on the production of verbs with low construction frequency ($t(85) = 1.13$, $p = 0.260$, $r = 0.12$). Results of the full set of post-hoc *t*-tests are shown in Table 8.11.

By-item ANOVAs

Summary statistics for the by-item analyses of verb production latencies from typical participants in the sentence completion task are available in Appendix N. A by-item ANOVA on the raw data revealed no significant main effect of lexical frequency ($F(1, 20) = 1.96$, $p = 0.177$, $r = 0.30$), no significant main effect of construction frequency ($F(1, 20) = 0.11$, $p = 0.743$, $r = 0.07$) and no significant interaction between construction frequency and lexical frequency ($F(1, 20) = 2.92$, $p = 0.103$, $r = 0.36$).

A by-item ANOVA on the transformed data revealed that the main effect of lexical frequency was approaching significance ($F(1, 20) = 3.90$, $p = 0.062$) and represented a medium effect size ($r = 0.40$). The main effect of construction frequency was not significant ($F(1, 20) = 0.00$, $p = 1.00$, $r = 0.00$), and the interaction between construction frequency and lexical frequency was not significant ($F(1, 20) = 1.54$, $p = 0.229$, $r = 0.27$).

minF’

Cohen's (1973) minF’ was calculated to examine the effects of lexical frequency, construction frequency and their interaction when both participants and items were treated as random effects. In the raw data, no effects reached significance. The main effect of lexical frequency was not significant (minF'(1, 37.12) = 1.58, p = ns, r = 0.20); the main effect of construction frequency was not significant (minF'(1, 28.95) = 0.10, p = ns, r = 0.06); and the interaction between construction frequency and lexical frequency was not significant (minF'(1, 30.88) = 2.58, p = ns, r = 0.28). In the transformed data, the main effect of lexical frequency approached significance (minF'(1, 36.32) = 3.18, $p < 0.100$) and represented a small effect size ($r = 0.28$). The main effect of construction frequency was not significant (minF'(1, 24) = 0.00, p = ns, r = 0.00), and the interaction between construction frequency and lexical frequency was not significant (minF'(1, 30.83) = 1.36, p = ns, r = 0.21).

Interim summary

By-participant analyses of data from the 86 typical participants included in the sentence completion task identified a significant main effect of lexical frequency, indicating that participants produced verbs with high lexical frequency with shorter latencies than verbs with low lexical frequency. A significant interaction in the by-participant analyses indicated that lexical frequency had a significant effect on the production of verbs with high construction frequency but not low construction frequency. However, these effects did not reach significance in the by-item analysis, and therefore also not by the calculation of minF'. No analysis identified a significant main effect of construction frequency.

Power

This study was sufficiently powered to detect medium to large effects. Power estimates revealed that the by-participant analysis achieved 61% power to detect an effect of lexical frequency. The power of this analysis falls short of the recommended 80% power for a statistical test (Cohen, 1992), but lexical frequency represented a small effect size (r = 0.29). To reach 80% power at an alpha-level of 0.05, using current mean and standard deviations, a sample size of 135 participants would be required. The comparison between responses to verbs in the [high cx, high lex] and [high cx, low lex] conditions was fully powered (100%) and represented a large effect size (r = 0.50).

The study was under-powered to detect the small (r = 0.11) effect of construction frequency: to reach 80% power at an alpha-level of 0.05, a sample size of approximately 1000 participants would be necessary. To detect an effect of lexical frequency on verbs with low construction frequency, a sample size of approximately 6000 participants would be needed to reach 80% power at an alpha-level of 0.05, using current means and standard deviations. With an effect size of $r = 0.04$, this comparison did not reach Cohen's (1988) cut-off for a small effect ($r = 0.10$).

Future research could usefully pursue the effect of lexical frequency in the sentence completion task using a larger sample of older adults. However, the sample size required to investigate the effect of construction frequency is beyond the means of most behavioural research projects.

8.1.4 Construction analyses

Repeated-measures ANOVAs on participants' verb production latencies to the six argument structure constructions in the sentence completion task were performed in order to examine whether the main effect of lexical frequency was consistent across all constructions, and whether particular constructions were driving the interaction that resulted in a significant effect of lexical frequency on the production of verbs with high construction frequency but not low construction frequency. Results from each construction are shown in Figure 8.4 (over page). A 2x2x6 ANOVA was not performed at the outset because only one verb was included per condition for each construction, making the results from these analyses specific to the lexical items included for each construction.

Data from the 86 typical participants included in the analysis of the sentence completion task were included in the analyses of individual constructions. Participants with missing data were excluded from the analysis of the repeated-measures design, so a different number of participants contributed data to the analysis of each construction. ANOVAs on data from each construction are reported in the following sections, and the full set of post-hoc *t*-tests exploring interactions are available in Appendix O.

Caused motion

Table 8.12 (over page) shows the summary statistics for the caused motion construction. A repeated-measures ANOVA on the raw data revealed no significant main effect of lexical frequency ($F(1, 65) = 0.19$, $p = 0.661$, $r = 0.05$), no significant main effect of construction frequency ($F(1, 65) = 0.71$, $p = 0.404$, $r = 0.10$) and no significant interaction ($F(1, 65) = 0.94$, $p = 0.336$, $r = 0.12$).

A repeated-measures ANOVA on the transformed data revealed no significant main effect of lexical frequency ($F(1, 65) = 0.01$, $p = 0.919$, $r = 0.01$), no significant main effect of construction frequency ($F(1, 65) = 1.64$, $p = 0.205$, $r = 0.16$) and no significant interaction ($F(1, 65) = 0.82$, $p = 0.367$, $r = 0.11$).

Conative

Table 8.13 (over page) shows the summary statistics for the conative construction. A repeated-measures ANOVA on the raw data revealed no significant main effect of lexical frequency ($F(1, 77) = 0.93$, $p = 0.338$, $r = 0.11$), no significant main effect of construction frequency ($F(1, 77) = 1.89$, $p = 0.173$, $r = 0.15$) and no significant interaction ($F(1, 77) = 0.38$, $p = 0.541$, $r = 0.07$).

A repeated-measures ANOVA on the transformed data revealed no significant main effect of lexical frequency ($F(1, 77) = 0.87$, $p = 0.354$, $r = 0.11$), no significant main effect of construction frequency ($F(1, 77) = 0.45$, $p = 0.504$, $r = 0.08$) and no significant interaction ($F(1, 77) = 0.00$, $p = 0.970$, $r = 0.00$).

Ditransitive

Table 8.14 (over page) shows the summary statistics for the ditransitive construction. A repeated-measures ANOVA on the raw data revealed a significant main effect of lexical frequency ($F(1, 78) = 18.06$, $p < 0.001$, $r = 0.43$), indicating that participants produced verbs with high lexical frequency with significantly shorter latencies than verbs with low lexical frequency. The main effect of construction frequency approached significance ($F(1, 78) = 2.99$, $p = 0.088$, $r = 0.19$), indicating that participants produced verbs with high construction frequency with longer latencies than verbs with low construction frequency. The interaction between construction frequency and lexical frequency was significant ($F(1, 78) = 23.78$, $p < 0.001$) and represented a medium effect size ($r = 0.48$). The interaction revealed that lexical frequency had a significant effect on the production of verbs with high construction frequency in the predicted direction ($t(79) = -6.60$, $p < 0.001$), i.e. the production latency of verbs with high lexical frequency was shorter than the production latency of verbs with low lexical frequency, and this represented a large effect ($r = 0.60$). In contrast, lexical frequency had a non-significant effect in the reverse direction on the production of verbs with low construction frequency ($t(84) = 0.15$, $p = 0.885$, $r = 0.02$), i.e. the production latency of verbs with high lexical frequency was longer than the production latency of verbs with low lexical frequency.

A repeated-measures ANOVA on the transformed data revealed a significant main effect of lexical frequency ($F(1, 78) = 29.43$, $p < 0.001$, $r = 0.52$), indicating that participants produced verbs with high lexical frequency with significantly shorter latencies than verbs with low lexical frequency. The main effect of construction frequency was not significant ($F(1, 78) = 2.69$, $p = 0.105$, $r = 0.18$). The interaction indicated that lexical frequency had a

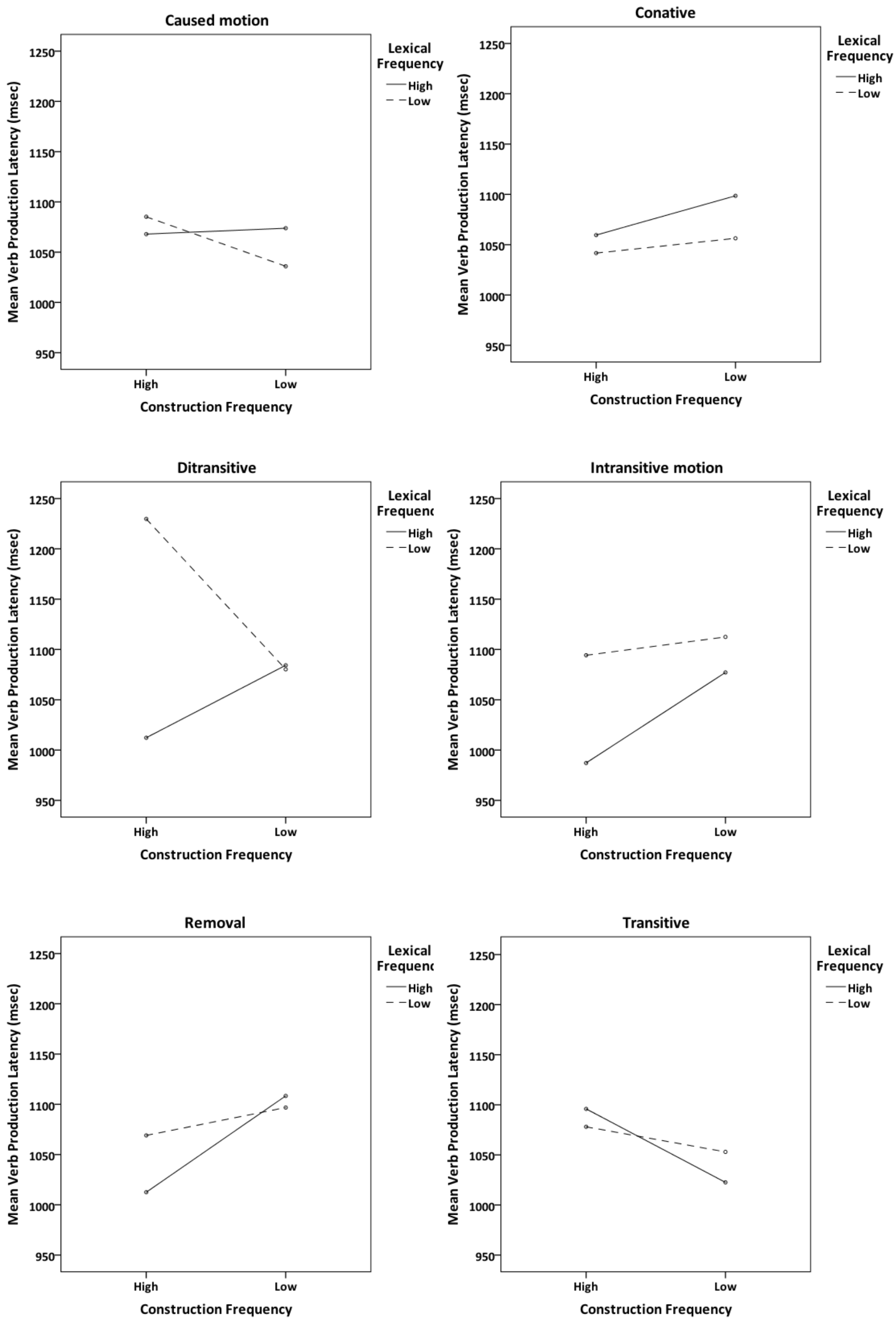


Figure 8.4 Interaction graphs of raw data from typical participants for argument structure constructions in sentence completion task

Table 8.12 Summary statistics for analysis of caused motion construction in sentence completion task

	M	95% CI		*SD*	min	max
		Lower bound	Upper bound			
Raw data						
high cx, high lex	1068	998	1138	286	540	2550
high cx, low lex	1085	1016	1154	280	700	2280
low cx, high lex	1074	998	1150	308	550	2700
low cx, low lex	1036	975	1097	247	650	1980
Back transformations						
high cx, high lex	1009	955	1071	-	540	2550
high cx, low lex	1028	974	1088	-	700	2280
low cx, high lex	1007	948	1073	-	550	2700
low cx, low lex	985	935	1041	-	650	1980

Note. Mean (*M*), lower and upper bound of 95% confidence interval (CI), standard deviation (*SD*), minimum value (min) and maximum value (max). *n* = 66.

Table 8.13 Summary statistics for analysis of conative construction in sentence completion task

	M	95% CI		*SD*	min	max
		Lower bound	Upper bound			
Raw data						
high cx, high lex	1060	996	1124	284	640	2506
high cx, low lex	1042	978	1105	282	590	2230
low cx, high lex	1099	1020	1177	350	550	2540
low cx, low lex	1056	990	1123	295	600	2490
Back transformations						
high cx, high lex	1002	954	1055	-	640	2506
high cx, low lex	979	930	1035	-	590	2230
low cx, high lex	1013	956	1078	-	550	2540
low cx, low lex	989	937	1048	-	600	2490

Note. Mean (*M*), lower and upper bound of 95% confidence interval (CI), standard deviation (*SD*), minimum value (min) and maximum value (max). *n* = 78.

Table 8.14 Summary statistics for analysis of ditransitive construction in sentence completion task

	M	95% CI		*SD*	min	max
		Lower bound	Upper bound			
Raw data						
high cx, high lex	1012	955	1069	255	640	2190
high cx, low lex	1230	1161	1299	309	730	2500
low cx, high lex	1084	1019	1150	293	560	2460
low cx, low lex	1080	1026	1135	244	620	1910
Back transformations						
high cx, high lex	959	913	1010	-	640	2190
high cx, low lex	1169	1114	1228	-	730	2500
low cx, high lex	1020	968	1079	-	560	2460
low cx, low lex	1030	983	1083	-	620	1910

Note. Mean (*M*), lower and upper bound of 95% confidence interval (CI), standard deviation (*SD*), minimum value (min) and maximum value (max). *n* = 79.

significant effect on the production of verbs with high construction frequency ($t(79) = 8.33$, $p < 0.001$), and this represented a large effect ($r = 0.68$). In contrast, lexical frequency did not have a significant effect on the production of verbs with low construction frequency ($t(84) = 0.40$, $p = 0.691$, $r = 0.04$).

Intransitive motion

Table 8.15 (over page) shows the summary statistics for the intransitive motion construction. A repeated-measures ANOVA on the raw data revealed a significant main effect of lexical frequency ($F(1, 77) = 14.14$, $p < 0.001$, $r = 0.39$), indicating that participants produced verbs with high lexical frequency with significantly shorter latencies than verbs with low lexical frequency. The main effect of construction frequency was significant ($F(1, 77) = 7.07$, $p = 0.010$, $r = 0.29$), indicating that participants produced verbs with high construction frequency with significantly shorter latencies than verbs with low construction frequency. The interaction between construction frequency and lexical frequency approached significance ($F(1, 77) = 3.43$, $p = 0.068$) and represented a small effect ($r = 0.21$). The interaction indicated that lexical frequency had a significant effect on the production of verbs with high construction frequency ($t(83) = -4.12$, $p < 0.001$), and this represented a medium effect size ($r = 0.41$); however, lexical frequency did not have a significant effect on the production of verbs with low construction frequency ($t(79) = -1.32$, $p = 0.192$, $r = 0.15$).

A repeated-measures ANOVA on the transformed data revealed a significant main effect of lexical frequency ($F(1, 77) = 19.43$, $p < 0.001$, $r = 0.45$), indicating that participants produced verbs with high lexical frequency with significantly shorter latencies than verbs with low lexical frequency. The main effect of construction frequency was significant ($F(1, 77) = 5.66$, $p = 0.020$, $r = 0.26$), indicating that participants produced verbs with high

construction frequency with significantly shorter latencies than verbs with low construction frequency. The interaction between construction frequency and lexical frequency was not significant ($F(1, 77) = 0.96$, $p = 0.331$, $r = 0.11$).

Removal

Table 8.16 (over page) shows the summary statistics for the removal construction. A repeated-measures ANOVA on the raw data revealed no significant main effect of lexical frequency ($F(1, 59) = 1.18$, $p = 0.281$, $r = 0.14$). The main effect of construction frequency was significant ($F(1, 59) = 8.33$, $p = 0.005$, $r = 0.35$), indicating that participants produced verbs with high construction frequency with significantly shorter latencies than verbs with

Table 8.15 Summary statistics for analysis of intransitive motion construction in sentence completion task

	M	95% CI		*SD*	min	max
		Lower bound	Upper bound			
Raw data						
high cx, high lex	987	943	1032	197	670	1640
high cx, low lex	1094	1033	1155	271	640	2090
low cx, high lex	1077	1010	1144	296	650	1960
low cx, low lex	1112	1053	1172	265	740	2030
Back transformations						
high cx, high lex	953	916	993	-	670	1640
high cx, low lex	1037	986	1092	-	640	2090
low cx, high lex	1006	952	1067	-	650	1960
low cx, low lex	1061	1015	1113	-	740	2030

Note. Mean (*M*), lower and upper bound of 95% confidence interval (CI), standard deviation (*SD*), minimum value (min) and maximum value (max). *n* = 78.

Table 8.16 Summary statistics for analysis of removal construction in sentence completion task

	M	95% CI		*SD*	min	max
		Lower bound	Upper bound			
Raw data						
high cx, high lex	1013	952	1073	233	680	2010
high cx, low lex	1069	1013	1125	216	713	1610
low cx, high lex	1108	1041	1176	263	730	1900
low cx, low lex	1097	1031	1163	256	740	2110
Back transformations						
high cx, high lex	974	932	1021	-	680	2010
high cx, low lex	1028	977	1084	-	713	1610
low cx, high lex	1053	996	1117	-	730	1900
low cx, low lex	1049	997	1106	-	740	2110

Note. Mean (*M*), lower and upper bound of 95% confidence interval (CI), standard deviation (*SD*), minimum value (min) and maximum value (max). *n* = 60.

Table 8.17 Summary statistics for analysis of transitive construction in sentence completion task

	M	95% CI		*SD*	min	max
		Lower bound	Upper bound			
Raw data						
high cx, high lex	1096	1029	1163	298	650	2000
high cx, low lex	1078	1025	1132	239	690	1710
low cx, high lex	1023	965	1080	258	540	1790
low cx, low lex	1053	1003	1103	222	650	1800
Back transformations						
high cx, high lex	1028	975	1087	-	650	2000
high cx, low lex	1030	984	1081	-	690	1710
low cx, high lex	959	906	1019	-	540	1790
low cx, low lex	1011	967	1058	-	650	1800

Note. Mean (*M*), lower and upper bound of 95% confidence interval (CI), standard deviation (*SD*), minimum value (min) and maximum value (max). *n* = 79.

low construction frequency. The interaction between construction frequency and lexical frequency was not significant ($F(1, 59) = 2.57$, $p = 0.114$, $r = 0.20$).

A repeated-measures ANOVA on the transformed data revealed no significant main effect of lexical frequency ($F(1, 59) = 2.02$, $p = 0.161$, $r = 0.18$). The main effect of construction frequency was significant ($F(1, 59) = 8.14$, $p = 0.006$, $r = 0.35$), indicating that participants produced verbs with high construction frequency with significantly shorter latencies than verbs with low construction frequency. The interaction between construction frequency and lexical frequency was not significant ($F(1, 59) = 2.75$, $p = 0.103$, $r = 0.21$).

Transitive

Table 8.17 (back page) shows the summary statistics for the transitive construction. A repeated-measures ANOVA on the raw data revealed no significant main effect of lexical frequency ($F(1, 78) = 0.10$, $p = 0.754$, $r = 0.04$). The main effect of construction frequency was significant ($F(1, 78) = 6.34$, $p = 0.014$, $r = 0.27$), indicating a reverse frequency effect, as participants produced verbs with high construction frequency with significantly longer latencies than verbs with low construction frequency. The interaction between construction frequency and lexical frequency was not significant ($F(1, 78) = 1.68$, $p = 0.198$, $r = 0.15$).

A repeated-measures ANOVA on the transformed data revealed no significant main effect of lexical frequency ($F(1, 78) = 2.31$, $p = 0.132$, $r = 0.17$). The main effect of construction frequency was significant ($F(1, 78) = 5.87$, $p = 0.018$, $r = 0.26$), indicating a reverse frequency effect, as participants produced verbs with high construction frequency with significantly longer latencies than verbs with low construction frequency. The interaction between construction frequency and lexical frequency was approaching significance ($F(1, 78) = 2.79$, $p = 0.099$) and represented a small effect ($r = 0.19$), indicating that lexical frequency had a small, significant effect on the production of verbs with low construction frequency ($t(81) = 2.23$, $p = 0.028$, $r = 0.24$) but a non-significant effect on verbs with high construction frequency ($t(82) = 0.39$, $p = 0.696$, $r = 0.04$).

Summary of results from individual argument structure constructions

Lexical frequency was revealed as a significant main effect on verb production latencies for the ditransitive and intransitive motion constructions. A significant main effect of construction frequency was identified for the intransitive motion and removal constructions, and a significant reverse effect of construction frequency was identified for the transitive construction. The interaction between lexical frequency and construction frequency resulting from a significant effect of lexical frequency on the production of verbs with high construction frequency but not low construction frequency was observed for the ditransitive construction and approached significance only in the analysis of the raw data for the intransitive motion construction.

8.1.5 Additional analyses

Additional analyses were carried out in order to examine whether variables associated with the participants or items in Phase 2 affected verb production latencies in the sentence completion task. Analyses were carried out on the raw and transformed values from the 86 typical participants included in the analysis of the sentence completion task. Note that the inverse transformation reverses the direction of the dataset, so negative correlations in the transformed data represent real positive relationships.

8.1.5.1 Participant variables

Age

Pearson's correlation coefficients were used to investigate the effect of age on verb production latencies. There was a significant relationship between age and verb production latencies in the raw data (r = 0.32, one-tailed p = 0.001) and in the transformed data (r = -0.32, one-tailed p = 0.001), indicating that verb production latencies increased with age; that is, older participants took longer to produce verbs than younger participants.

Education

Pearson's correlation coefficients were used to investigate the effect of education on verb production latencies. There was no significant correlation between the number of years participants spent in education and verb production latencies in the raw data (r = 0.01, one-tailed p = 0.471) or in the transformed data (r = 0.03, one-tailed p = 0.404).

Gender

An independent samples t-test was used to examine differences in verb production latencies between women and men. In the raw data, women showed significantly shorter verb production latencies than men ($t(84)$ = -2.06, two-tailed p = 0.042, r = 0.22; women $\bar{x}$ = 1025 msec; men $\bar{x}$ = 1112 msec). This difference was approaching significance in the transformed data ($t(84)$ = 1.94, two-tailed p = 0.056, r = 0.21; women $\bar{x}$ = 975 msec; men $\bar{x}$ = 1046 msec).

Age and gender interactions

Results from the above analyses indicated that age and gender affected verb production latencies. Table 8.18 shows Pearson's correlation coefficients for the relationship between age and verb production latencies in each condition in the sentence completion task for men and women separately, in order to examine the effects of participants' characteristics on

Table 8.18 Pearson's correlation coefficients for relationship between age and verb production latencies for each condition in sentence completion task, by gender

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
Women (*n* = 43)				
raw values	0.44**	0.44**	0.45**	0.43**
transformed values	-0.42**	-0.42**	-0.41**	-0.40**
Men (*n* = 43)				
raw values	0.18	0.30*	0.13	0.15
transformed values	-0.18	-0.29*	-0.16	-0.12

Note. * *p* < 0.05 (one-tailed). ** *p* < 0.01 (one-tailed).

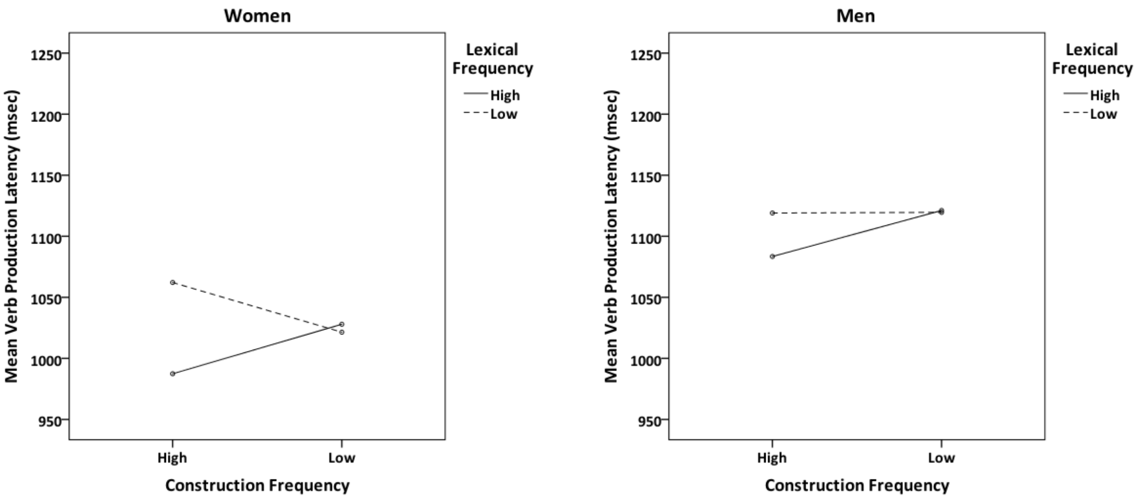


Figure 8.5 Interaction graphs of verb production latencies in raw data of sentence completion task for women (left pane) and men (right pane)

verb production latencies in more detail. Note that the negative sign of correlations for transformed values is due to the inverse transformation and in fact represents a real positive relationship.

Results of the correlational analyses indicated that the relationship between age and verb production latencies was significant in all four conditions in the sentence completion task for women only. These correlations represented a medium effect. In contrast, the correlation between age and verb production latencies in men represented a smaller effect, which was significant only in the [high cx, low lex] condition.

Age and gender were included as covariates in by-participant ANOVAs on verb production latencies from the 86 participants included in the sentence completion task in order to examine whether the effect of lexical frequency and the interaction resulting in a significant effect of lexical frequency in response to verbs with high construction frequency but not low construction frequency were robust against the effects of age and gender.

In the raw data, the main effect of lexical frequency was not significant when age and gender were included as covariates in an ANCOVA ($F(1, 83) = 0.14$, $p = 0.707$, $r = 0.04$). The interaction between lexical frequency and construction frequency did not reach significance when age and gender were included as covariates ($F(1, 83) = 2.13$, $p = 0.148$, $r = 0.16$). There were only two significant effects in the model containing the covariates: (1) the two-way interaction between lexical frequency, construction frequency and age ($F(1, 83) = 4.63$, $p = 0.034$, $r = 0.23$), and (2) the two-way interaction between lexical frequency, construction frequency and gender ($F(1, 83) = 4.12$, $p = 0.046$, $r = 0.22$). Figure 8.5 shows

verb production latencies in each condition of the sentence completion task for women and men separately, based on raw data. Women showed a significant effect of lexical frequency on the production of verbs with high construction frequency ($t(42) = -5.55$, $p < 0.001$). When a Bonferroni-corrected p-value of 0.008 was implemented to explore the interaction in men, the effect of lexical frequency on the production of verbs with high construction frequency did not reach significance ($t(42) = -2.37$, $p = 0.022$). Further, the effect of lexical frequency on the production of verbs with high construction frequency represented a large effect in women ($r = 0.65$) but a small effect in men ($r = 0.34$).

In the transformed data, the main effect of lexical frequency was not significant when age and gender were included as covariates in an ANCOVA ($F(1, 83) = 0.10$, $p = 0.756$, $r = 0.03$). The interaction between construction frequency and lexical frequency was approaching significance ($F(1, 83) = 3.01$, $p = 0.086$, $r = 0.19$). Only two other effects in the model were significant or approaching significance: (1) the two-way interaction between lexical frequency, construction frequency and age was significant ($F(1, 83) = 5.17$, $p = 0.026$, $r = 0.24$), and (2) the two-way interaction between lexical frequency, construction frequency and gender was approaching significance ($F(1, 83) = 3.37$, $p = 0.070$, $r = 0.20$). As in the raw data, the effect of lexical frequency on the production of verbs with high construction frequency represented a large effect in women ($r = 0.67$) but a small effect in men ($r = 0.39$).

8.1.5.2 Item variables

Pearson's correlation coefficients were used to examine the effect of variables known to influence single-word processing on verb production latencies in the sentence completion task. These variables were described in section 7.1.5.2.

Age-of-acquisition

Ratings from Kuperman et al.'s (2012) dataset were available for all 24 verbs included in the sentence completion task. There was a significant relationship between age-of-acquisition and verb production latencies in the raw data ($r = 0.58$, one-tailed $p = 0.002$) and in the transformed data ($r = -0.61$, one-tailed $p = 0.001$), indicating that verb production latencies were longer for later acquired verbs. However, this relationship was influenced by the outlier verb OWE, which was the latest acquired verb in the dataset with an age-of-acquisition rating of 8.61 years. Excluding the verb OWE, correlations between age-of-acquisition and verb production latencies were not significant in the raw data ($r = 0.15$, one-tailed $p = 0.249$) or in the transformed data ($r = -0.23$, one-tailed $p = 0.144$).

Imageability

Ratings from Cortese and Fugett's (2004) dataset were available for 23 of the 24 verbs included in the sentence completion task. There was no significant relationship between imageability and verb production latencies in the raw data ($r = -0.05$, $p = 0.416$) or in the transformed data ($r = -0.02$, $p = 0.471$).

Lexical decision times

Lexical decision times from Keuleers et al.'s (2012) British Lexicon Project were available for all 24 verbs included in the sentence completion task. There was a significant relationship between lexical decision times and verb production latencies in the raw data ($r = 0.41$, $p = 0.024$) and in the transformed data ($r = -0.43$, $p = 0.018$), indicating that verbs which attracted longer response times in a lexical decision task were produced with longer latencies in the sentence completion task. As above, the verb OWE was an outlier, with the longest lexical decision time of 608 milliseconds. Excluding the verb OWE, correlations between lexical decision times and verb production latencies were not significant in the raw data ($r = 0.13$, $p = 0.284$) or in the transformed data ($r = -0.17$, $p = 0.217$).

Orthography

The effect of three orthographic characteristics of verbs was investigated, including Coltheart's N, OLD20 and the number of letters in each verb. These three orthographic measures were available for all 24 verbs in the sentence completion task. None of the orthographic characteristics proved to affect verb production latencies.

There was no significant correlation between Coltheart's N and verb production latencies in the raw dataset ($r = 0.11$, one-tailed $p = 0.307$) or in the transformed dataset ($r = -0.14$, one-tailed $p = 0.265$). There was no significant correlation between OLD20 and verb production latencies in the raw dataset ($r = 0.00$, one-tailed $p = 0.497$) or in the transformed dataset ($r = 0.03$, one-tailed $p = 0.440$). There was no significant correlation between the number of letters in each verb and verb production latencies in the raw dataset ($r = -0.02$, one-tailed $p = 0.457$) or in the transformed dataset ($r = 0.03$, one-tailed $p = 0.443$).

8.1.6 Summary of results from typical participants in sentence completion task

By-participant analyses revealed a significant main effect of lexical frequency on verb production latencies, indicating that typical participants produced verbs with high lexical frequency with shorter latencies than verbs with low lexical frequency. An interaction between lexical frequency and construction frequency indicated that lexical frequency had a significant effect on responses to verbs with high construction frequency, but not verbs with low construction frequency. Responses to the six argument structure constructions in the sentence completion task were more variable. This variability likely reflects item-specific responses, because only one verb was included per condition for each construction in the task. The effects uncovered in the main analyses were driven by responses to the ditransitive and intransitive motion constructions.

Further analyses revealed effects of age and gender on verb production latencies. Production latencies were significantly positively correlated with age in each condition for women, but not men. When age and gender were included as covariates in an ANCOVA, the main effect of lexical frequency and the interaction between lexical frequency and construction frequency did not reach significance. However, the ANCOVA revealed significant two-way interactions between lexical frequency, construction frequency and age and between lexical frequency, construction frequency and gender, indicating that the effect of lexical frequency on the production of verbs with high construction frequency represented a large effect for women, but a small effect for men.

8.2 Results from participants with aphasia

Results from participants with aphasia in the Phase 2 sentence completion task are reported in this section. Responses from 11 participants with aphasia were included in the analysis of the sentence completion task. Upon assessment, UT was unable to read aloud the sentence stimuli. VH and WD did not complete the sentence completion task because their spoken output was severely affected by apraxia. As in Chapter 7, participants are listed throughout this section in order of their scores in the written sentence comprehension subtest of the *Comprehensive Aphasia Test*, beginning with participants with the highest scores.

For participants with aphasia, each construction was associated with a set of six verbs in the sentence completion task. Two verbs per construction were included as practice items and did not form part of the analysis. Only responses to the 24 verbs in the frequency conditions in the design of Phase 2 were included in analyses reported below.

In the following, Section 8.2.1 presents analyses of the number of target responses participants with aphasia produced in the task. Section 8.2.2 presents analyses of verb

Table 8.19 Number of responses from participants with aphasia included in sentence completion task

	All items *N* = 24	high cx, high lex *n* = 6	high cx, low lex *n* = 6	low cx, high lex *n* = 6	low cx, low lex *n* = 6
BF	24	6	6	6	6
CJ	15	5	5	2	3
GW	22	5	5	6	6
HM	21	5	5	6	5
IC	20	6	5	5	4
JF	23	6	5	6	6
KT	24	6	6	6	6
MJ	19	5	6	6	2
PD	19	5	4	5	5
RE	24	6	6	6	6
SP	15	5	4	4	2
Mdn	21	5	5	6	5
M	20.5	5.5	5.2	5.3	4.6

production latencies, including analyses at the group level, by construction and within individual participants with aphasia. Finally, Section 8.2.3 compares the performance of participants with aphasia to typical participants. Section 8.2.4 concludes with a summary of the results from participants with aphasia in the sentence completion task.

8.2.1 Number of target responses from participants with aphasia in sentence completion task

8.2.1.1 Number of responses from individual participants with aphasia in sentence completion task

As described in Section 4.4.1.2, responses from participants with aphasia were excluded from analysis for several reasons. As for typical participants, the following types of responses were excluded due to participant behaviour: (1) productions that included self-corrections or comments which affected the accurate measurement of the verb production latency; (2) productions that contained non-target constructions; (3) productions that were incomplete due to the termination of the audio recording; and (4) no response. In addition, responses that contained a non-target verb, or speech errors on the production of the verb, were also excluded. The number of responses from each participant with aphasia included in the analysis of the sentence completion task is shown in Table 8.19.

No responses from participants with aphasia were excluded due to experimental error.

8.2.1.2 Number of responses from group of participants with aphasia in sentence completion task

Figure 8.6 shows the mean number of target responses from the group of participants with aphasia in the sentence completion task per condition. Mean number of responses are

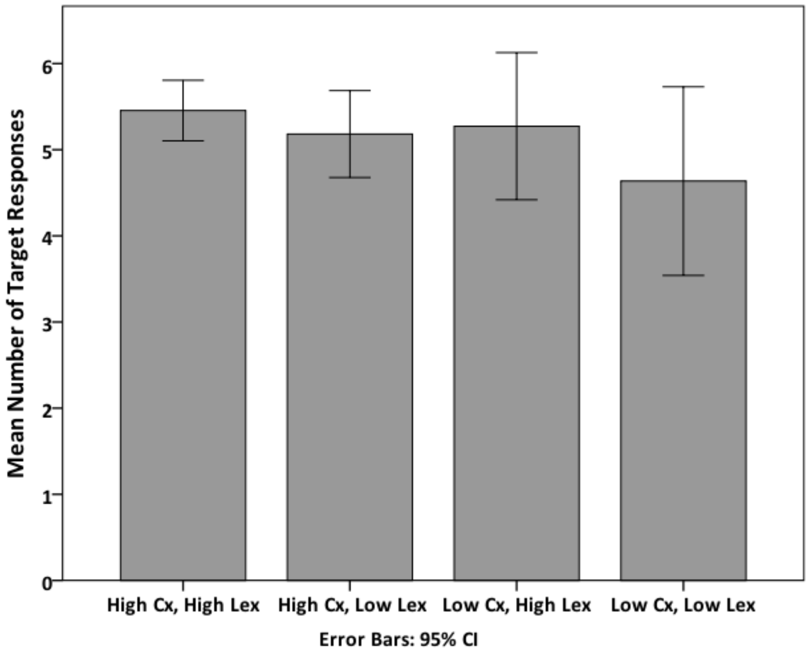


Figure 8.6 Mean number of responses from participants with aphasia to conditions in sentence completion task

shown, despite the use of non-parametric statistical tests, because of the limited range of the number of responses.

As described in Chapter 7, the four frequency conditions were collapsed in two ways in order to investigate the main effects of construction frequency and lexical frequency on the number of responses from the group of participants with aphasia. Comparisons were made using the Wilcoxon signed-rank test. The main effect of construction frequency was examined by comparing the number of participants' responses in the conditions [high cx, high lex] and [high cx, low lex] to the number of participants' responses in the conditions [low cx, high lex] and [low cx, low lex]. The main effect of lexical frequency was examined by comparing the number of participants' responses in the conditions [high cx, high lex] and [low cx, high lex] to the number of participants' responses in the conditions [high cx, low lex] and [low cx, low lex].

Participants with aphasia produced significantly more target responses containing verbs with high lexical frequency than low lexical frequency ($T = 2.50$, two-tailed $p = 0.047$, $r = -0.42$). There was no significant difference between the number of target responses from participants with aphasia containing verbs with high and low construction frequency ($T = 10.50$, two-tailed $p = 0.290$, $r = -0.23$).

In order to investigate interactions between the main effects of lexical frequency and construction frequency, Wilcoxon signed-rank tests were performed to compare the number of responses participants produced in each condition. Participants with aphasia produced a

Table 8.20 Results from Wilcoxon signed-rank tests on differences in number of responses between conditions in sentence completion task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	-			
high cx, low lex	*T* = 3.00 *p* = 0.180 *r* = -0.29	-		
low cx, high lex	*T* = 9.00 *p* = 0.739 *r* = -0.07	*T* = 5.00 *p* = 0.480 *r* = -0.15	-	
low cx, low lex	*T* = 1.00 *p* = 0.077 *r* = -0.38	*T* = 7.50 *p* = 0.262 *r* = -0.24	*T* = 2.00 *p* = 0.131 *r* = -0.32	-

greater number of target responses in the [high cx, high lex] condition than the [low cx, low lex] condition, and this difference was approaching significance ($T = 1.00$, $p = 0.077$, $r = -0.38$). No other difference between conditions reached significance. Results of the full set of comparisons are shown in Table 8.20.

To summarise, lexical frequency had a significant main effect on the number of responses participants with aphasia produced in the sentence completion task, and this represented a moderate effect size. Participants with aphasia produced more target responses containing [high cx, high lex] verbs than [low cx, low lex] verbs, and this difference was approaching significance. There were no other significant interactions.

8.2.1.3 Number of responses from group of participants with aphasia in sentence completion task by construction

The analyses described in Section 8.2.1.2 were performed separately for each of the six argument structure constructions in Phase 2. Figure 8.7 (over page) shows the mean number of responses in each condition to the six constructions in the sentence completion task. Note only one verb was included per condition for each construction, so the mean represents a more sensitive method of reporting than the median in this instance, despite the use of non-parametric statistical tests.

Results from the two Wilcoxon signed-rank tests on the number of responses containing verbs with high and low construction frequency and lexical frequency are shown for each construction individually in Table 8.21 (over pages). Results of the full set of Wilcoxon signed-rank tests for each construction are provided in Appendix P.

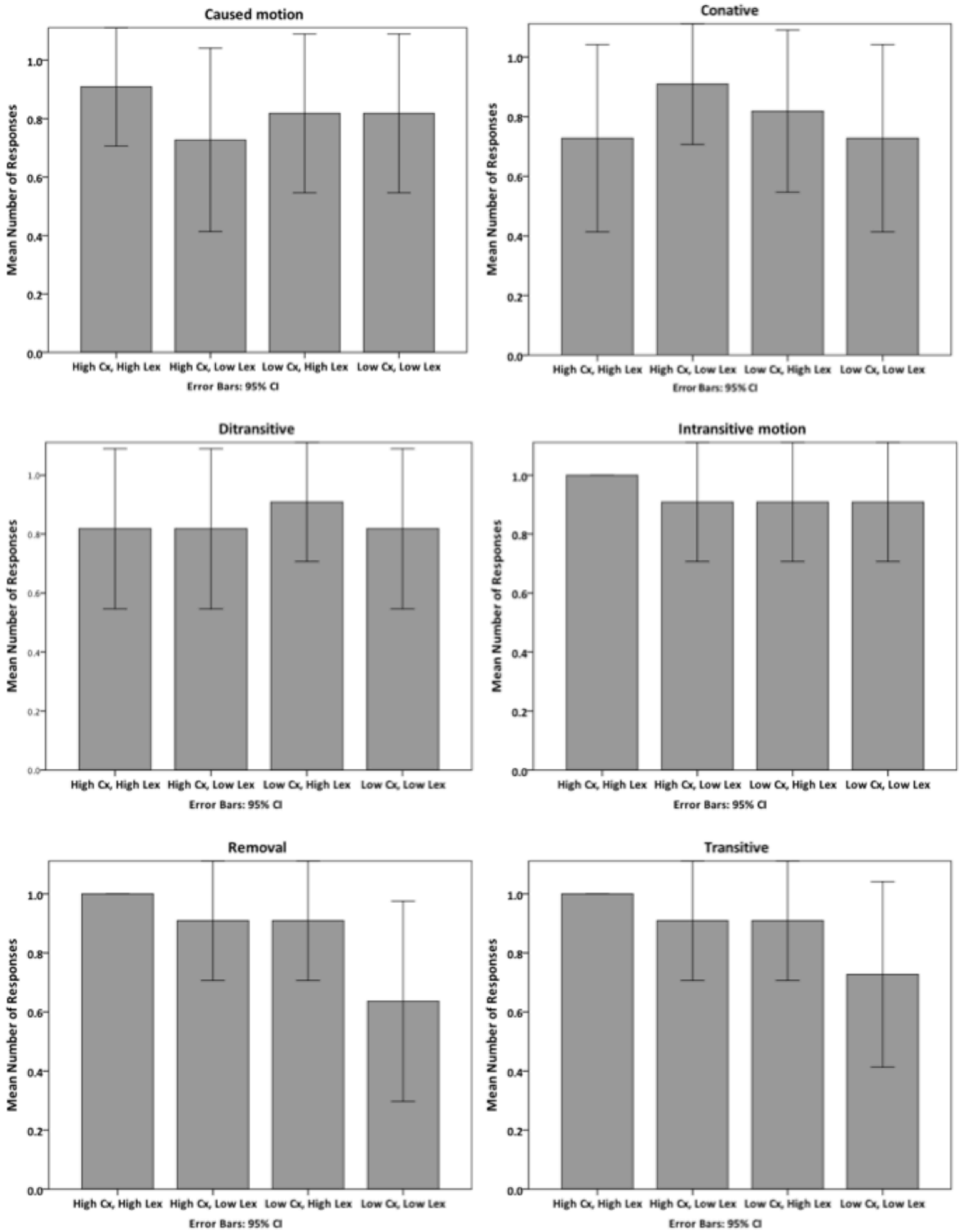


Figure 8.7 Mean number of responses from participants with aphasia per condition for each construction in the sentence completion task

Caused motion

There were no significant main effects of lexical frequency or construction frequency on the number of target responses participants with aphasia produced to the caused motion construction. None of the differences between conditions was significant.

Conative

There were no significant main effects of lexical frequency or construction frequency on the number of target responses participants with aphasia produced to the conative construction. None of the differences between conditions was significant.

Ditransitive

There were no significant main effects of lexical frequency or construction frequency on the number of target responses participants with aphasia produced to the ditransitive construction. None of the differences between conditions was significant.

Intransitive motion

There were no significant main effects of lexical frequency or construction frequency on the number of target responses participants with aphasia produced to the intransitive motion construction. None of the differences between conditions was significant.

Removal

There was a significant main effect of lexical frequency on the number of responses participants with aphasia produced to the removal construction ($T = 0$, $p = 0.046$), indicating that participants produced significantly more target responses containing verbs with high lexical frequency than low lexical frequency. The main effect of lexical frequency represented a moderate effect size ($r = -0.43$). Participants with aphasia produced significantly more target responses containing verbs in the [high cx, high lex] condition than the [low cx, low lex] condition ($T = 0.00$, $p = 0.046$, $r = -0.43$). The effect of lexical frequency on responses containing verbs with low construction frequency was approaching significance ($T = 1.00$, $p = 0.083$, $r = -0.37$).

Transitive

There were no significant main effects of lexical frequency or construction frequency on the number of target responses participants with aphasia produced to the transitive construction. However, participants produced more target responses containing verbs in the [high cx, high lex] condition than verbs in the [low cx, low lex] condition, and this difference was approaching significance ($T = 0.00$, $p = 0.083$, $r = -0.37$).

Table 8.21 Results from Wilcoxon signed-rank tests on number of target responses from participants with aphasia for argument structure constructions in Phase 2 sentence completion task

	Construction Frequency						Lexical Frequency					
	high freq $\bar{x}$	low freq $\bar{x}$	*T*	*p*	*z*	*r*	high freq $\bar{x}$	low freq $\bar{x}$	*T*	*p*	*z*	*r*
Caused motion	1.64	1.64	10.5	1.000	0.00	0.00	1.73	1.55	7	0.414	-0.82	-0.17
Conative	1.64	1.55	0	0.317	-1.00	-0.21	1.55	1.64	0	0.317	-1.00	-0.21
Ditransitive	1.64	1.73	2	0.564	-0.58	-0.12	1.73	1.64	2	0.564	-0.58	-0.12
Intransitive motion	1.91	1.82	1	0.655	-0.45	-0.10	1.91	1.82	0	0.317	-1.00	-0.21
Removal	1.91	1.55	2.5	0.157	-1.41	-0.30	1.91	1.55	0	0.046*	-2.00	-0.43
Transitive	1.91	1.64	0	0.180	-1.34	-0.29	1.91	1.64	0	0.180	-1.34	-0.29

Note. Mean number of participants' target responses in high and low frequency conditions (maximum 2); Wilcoxon test statistic *T*; two-tailed *p*-value; *z*-score; effect size *r*. * $p < 0.05$.

8.2.2 Verb production latencies from participants with aphasia in sentence completion task

8.2.2.1 Verb production latencies from individual participants with aphasia in sentence completion task

Table 8.22 shows the median verb production latencies from each participant with aphasia in the four conditions in the sentence completion task. The number of values contributing to each median varies according to the number of responses included in the analysis of the sentence completion task, as shown in Table 8.19.

Table 8.22 Median verb production latencies from participants with aphasia in sentence completion task, in milliseconds

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
BF	1027 (236)	1005 (175)	1005 (220)	995 (115)
CJ	1410 (1935)	3860 (1783)	1760 (141)	2470 (2907)
GW	1250 (446)	1340 (212)	1235 (129)	1365 (209)
HM	1110 (1103)	1190 (464)	1205 (340)	1270 (705)
IC	2135 (698)	1680 (590)	1290 (467)	2180 (572)
JF	1350 (446)	1140 (377)	1455 (630)	1710 (215)
KT	1285 (142)	1295 (143)	1399 (225)	1365 (274)
MJ	2090 (626)	2740 (710)	2740 (528)	2215 (163)
PD	3660 (1439)	3655 (1659)	4950 (1913)	4620 (2258)
RE	1465 (224)	1375 (234)	1195 (315)	1621 (352)
SP	1180 (389)	1040 (152)	1070 (180)	880 (113)
Mdn	1350	1340	1290	1621
M	1633	1847	1755	1881

Note. Standard deviation in brackets. Alternating rows in grey for ease of reference. Bottom rows show values of group median (Mdn) and mean (*M*).

8.2.2.2 Verb production latencies from group of participants with aphasia in sentence completion task

Median verb production latencies from participants with aphasia are shown in Figure 8.8 (over page) for each condition in the task. Two Wilcoxon signed-rank tests were used to examine the main effects of construction frequency and lexical frequency on verb production latencies from the group of participants with aphasia. There was no significant difference in production latencies between verbs with high and low lexical frequency (T = 27.00, two-tailed p = 0.594, r = -0.11), or between verbs with high and low construction frequency (T = 29.00, two-tailed p = 0.722, r = -0.08).

Wilcoxon signed-rank tests were used to investigate participants' median verb production latencies between each condition in the task. Participants with aphasia produced verbs in the [high cx, high lex] condition with significantly shorter latencies than verbs in the [low cx,

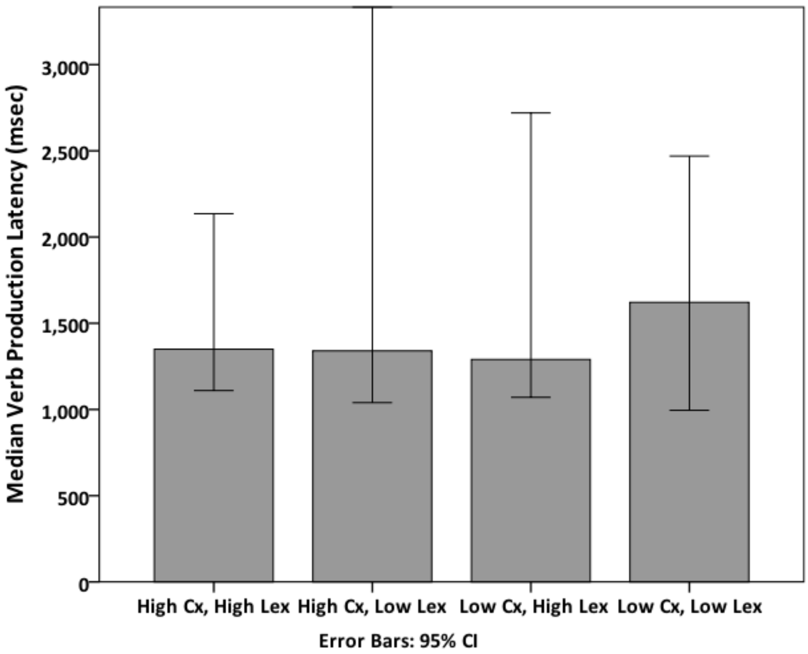


Figure 8.8 Median verb production latencies from participants with aphasia in Phase 2 sentence completion task

Table 8.23 Results from Wilcoxon signed-rank tests on verb production latencies from participants with aphasia between conditions in sentence completion task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	-			
high cx, low lex	*T* = 32.50 *p* = 0.965 *r* = -0.01	-		
low cx, high lex	*T* = 25.00 *p* = 0.477 *r* = -0.15	*T* = 24.00 *p* = 0.721 *r* = -0.08	-	
low cx, low lex	*T* = 9.00 *p* = 0.033 *r* = -0.45	*T* = 25.00 *p* = 0.477 *r* = -0.15	*T* = 24.00 *p* = 0.424 *r* = -0.17	-

low lex] condition (T = 9.00, p = 0.033), and this represented a moderate effect size (r = -0.45). Results from the full set of analyses are shown in Table 8.23.

In sum, there were no significant main effects of lexical frequency or construction frequency on verb production latencies from participants with aphasia in the sentence completion task. However, participants with aphasia produced verbs in the [high cx, high lex] condition with significantly shorter latencies than verbs in the [low cx, low lex condition]. No other interactions reached significance.

Power estimates indicated that with 11 participants and an alpha-level of 0.05, the study was under-powered to detect an effect of construction frequency or lexical frequency for parametric analysis. To reach 80% power at an alpha-level of 0.05, using the mean and standard deviations from the current results, a sample size of approximately 200 participants would be necessary to detect an effect of construction frequency and a sample size of 825 to detect an effect of lexical frequency. The necessary sample size for detecting the interaction between responses to verbs in the [high cx, high lex] and [low cx, low lex] conditions would be 25 for parametric analysis. As mentioned in Section 7.2.2.2, sample size estimates for the Wilcoxon signed-rank test range from approximately 10 for large differences to well over 100 for small differences (Shieh et al., 2007). The parametric power estimates suggest that future research might profitably investigate the observed interaction in a larger sample of adults with aphasia; however, the sample sizes required to detect main effects of construction frequency or lexical frequency in the sentence completion task suggest that these main effects are not likely to be of practical significance.

8.2.2.3 Verb production latencies from group of participants with aphasia in sentence completion task by construction

The analyses reported in Section 8.2.2.2 were performed separately for each of the six argument structure constructions included in Phase 2. Figure 8.9 (over page) shows median verb production latencies per condition for each construction in the task. Results from the two Wilcoxon signed-rank tests investigating main effects of construction frequency and lexical frequency are shown for each construction in Table 8.24 (over page). Results of the full set of Wilcoxon signed-rank tests for each construction are available in Appendix Q.

Caused motion

There were no significant main effects of lexical frequency or construction frequency on verb production latencies to the caused motion construction from participants with aphasia. None of the differences between conditions was significant.

Conative

There were no significant main effects of lexical frequency or construction frequency on verb production latencies to the conative construction from participants with aphasia. However, production latencies of verbs in the [high cx, high lex] condition were longer than verbs in the [low cx, high lex] condition, and this difference was approaching significance ($T = 3.00$, $p = 0.063$, $r = -0.45$). Production latencies of verbs in the [low cx, low lex] condition were longer than to verbs in the [high cx, low lex] condition, and this difference was also approaching significance ($T = 3.00$, $p = 0.063$, $r = -0.45$).

.

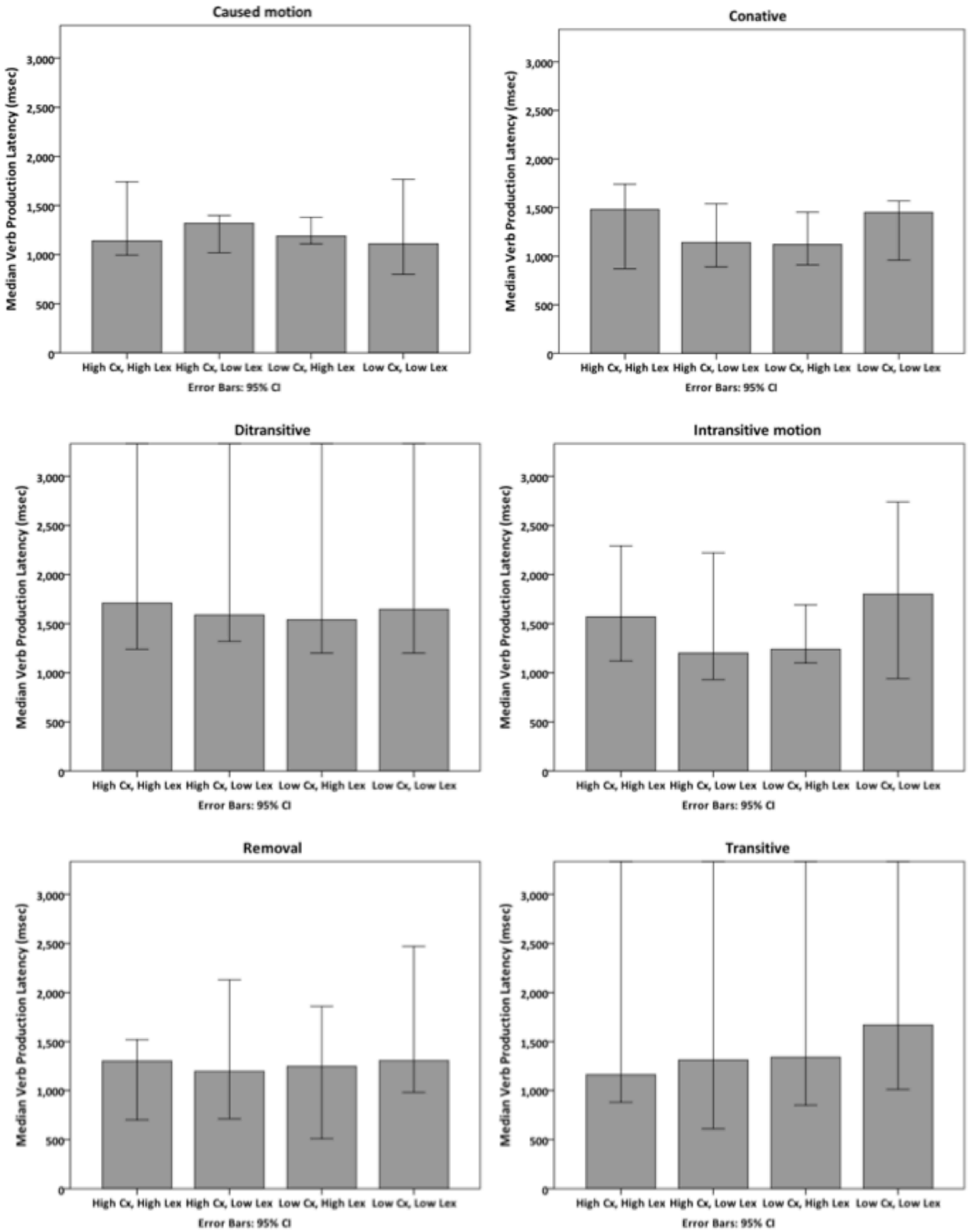


Figure 8.9 Median verb production latencies from participants with aphasia per condition for each construction in sentence completion task

Ditransitive

There were no significant main effects of lexical frequency or construction frequency on verb production latencies to the ditransitive construction from participants with aphasia. None of the differences between conditions was significant.

Intransitive motion

There were no significant main effects of lexical frequency or construction frequency on verb production latencies to the intransitive motion construction from participants with aphasia. Production latencies of verbs in the [high cx, high lex] condition were significantly longer than verbs in the [high cx, low lex] condition (T = 8.00, p = 0.047, r = -0.43). Production latencies of verbs in the [low cx, low lex] were longer than verbs in the [high cx, low lex] condition, and this difference was approaching significance (T = 8.00, p = 0.086, r = -0.38).

Removal

There were no significant main effects of lexical frequency or construction frequency on verb production latencies to the removal construction from participants with aphasia. The effect of lexical frequency on responses containing verbs with high construction frequency was approaching significance (T = 9.00, p = 0.059, r = -0.41). Additionally, production latencies of verbs in the [high cx, high lex] condition were shorter than of verbs in the [low cx, low lex] condition, and this difference was approaching significance (T = 3.50, p = 0.075, r = -0.42).

Transitive

Construction frequency had a significant main effect on verb production latencies from participants with aphasia in the transitive construction (T = 4.00, p = 0.017), indicating that participants with aphasia produced verbs with high construction frequency with shorter latencies than verbs with low construction frequency, and this represented a large effect (r = -0.52). This effect of construction frequency was approaching significance for verbs with high lexical frequency (T = 6.00, p = 0.069, r = -0.40) but was not significant for verbs with low lexical frequency (T = 3.00, p = 0.594, r = -0.11). Additionally, the effect of lexical frequency on responses to verbs with low construction frequency was approaching significance (T = 16.00, p = 0.074, r = -0.37).

Table 8.24 Results from Wilcoxon signed-rank tests on median verb production latencies from participants with aphasia for argument structure constructions in Phase 2 sentence completion task

	Construction Frequency						Lexical Frequency					
	high freq mdn	low freq mdn	*T*	*p*	*z*	*r*	high freq mdn	low freq mdn	*T*	*p*	*z*	*r*
Caused motion	2017	1809	16.5	0.262	-1.12	-0.24	1891	1859	21	0.508	-0.66	-0.14
Conative	2323	1RE	23	0.646	-0.46	-0.10	1965	1993	18	0.333	-0.97	-0.21
Ditransitive	2195	2035	22	0.953	-0.06	-0.01	1995	2254	19	0.678	-0.42	-0.09
Intransitive motion	1900	2049	21	0.508	-0.66	-0.14	1931	1950	28	0.657	-0.45	-0.09
Removal	1616	1891	13	0.139	-1.48	-0.32	1606	1813	17	0.155	-1.42	-0.30
Transitive	1656	2126	4	0.017*	-2.40	-0.52	1720	2206	16	0.241	-1.17	-0.26

Note. Mean of participants' median response times to verbs in high and low frequency conditions (maximum $n = 2$); Wilcoxon test statistic T; two-tailed p-value; z-score; effect size r. * $p < 0.05$.

8.2.2.4 Verb production latencies from participants with aphasia in sentence completion task by individual

The main effects of construction frequency and lexical frequency on verb production latencies were investigated for each participant with aphasia individually. Verbs from the four frequency conditions were collapsed in two ways, as described in Section 8.2.1.2, in order to analyse the main effect of construction frequency and lexical frequency within each participant with aphasia. Two Mann-Whitney tests were used to inspect data from each participant, using the exact method due to the small number of data points. Results are shown in Table 8.25 (over page).

Eight of the 11 participants with aphasia showed shorter production latencies of verbs with high lexical frequency than verbs with low lexical frequency; however, this difference was approaching significance only for CJ, where it represented a medium effect (r = -0.36). Six of the 11 participants with aphasia showed shorter production latencies of verbs with high construction frequency than verbs with low construction frequency; however, this difference reached significance only for JF, where it represented a medium effect (r = -0.40).

As discussed in Section 7.2.2.4, the exact Mann-Whitney test may have lacked the power to detect real differences between response times to verbs with high and low frequency from participants with aphasia, given the number of items included in the sentence completion task. Future research could include a greater number of items in a similar task to increase the power of within-participant comparisons. Note, however, that results showed low effect sizes for most participants in this task (see Table 8.25).

8.2.3 Comparison of participants with aphasia to typical participants in sentence completion task

This section reports results from Bayesian Standardised Difference Tests (BSDTs) and effects scores investigating the difference in production latencies of verbs with high and low frequency between participants with aphasia and typical participants. The 86 typical participants included in the analysis of the sentence completion task served as the comparison to participants with aphasia. As in Chapter 7, two sets of BSDTs were performed: one to investigate differences in lexical frequency, and another to investigate differences in construction frequency. Analyses were carried out on both raw and transformed data. The performance of typical participants was defined as their mean verb

Table 8.25 Results of Mann-Whitney tests on verb production latencies from individual participants with aphasia in Phase 2 sentence completion task

		Construction Frequency						Lexical Frequency					
	n	high freq mdn	low freq mdn	*U*	*p*	*z*	*r*	high freq mdn	low freq mdn	*U*	*p*	*z*	*r*
BF	24	1017	995	71.5	0.495	-0.03	-0.01	1027	995	65.5	0.362	-0.38	-0.08
CJ	15	2480	1860	24	0.477	-0.12	-0.03	1660	3165	16	0.095	-1.39	-0.36
GW	22	1295	1290	58	0.455	-0.13	-0.03	1250	1340	57.7	0.430	-0.20	-0.04
HM	21	1180	1210	54	0.479	-0.07	-0.02	1200	1230	51	0.398	-0.28	-0.06
IC	20	1980	1820	41	0.276	-0.65	-0.14	1920	1820	40	0.251	-0.72	-0.16
JF	23	1300	1585	35	0.029*	-1.91	-0.40	1370	1460	64.5	0.470	-0.09	-0.02
KT	24	1285	1399	65	0.351	-0.40	-0.08	1000	1020	71.5	0.494	-0.03	-0.01
MJ	22	2220	2525	52	0.303	-0.56	-0.12	2360	2485	48	0.219	-0.82	-0.04
PD	19	3660	4875	33	0.178	-0.98	-0.22	4875	3740	41	0.390	-0.33	-0.08
RE	24	1405	1311	64.5	0.341	-0.43	-0.09	1255	1437	51.5	0.124	-1.18	-0.24
SP	15	1060	1000	20.5	0.238	-0.77	-0.20	1120	990	16.5	0.117	-1.24	-0.32

Note. Number of items in each analysis (*n*); median verb production latencies in high and low frequency conditions in milliseconds; Mann-Whitney test statistic *U*; one-tailed *p*-value calculated via exact method; *z*-score; effect size *r*. * $p < 0.05$.

production latency to high frequency verbs and low frequency verbs. The performance of participants with aphasia was defined as their median verb production latency to high frequency verbs and low frequency verbs in the raw data and their mean verb production latency in the transformed data. For both participant groups, values were transformed before being averaged. Results of the full set of BSDTs are available in Appendix R. Effect scores based on the raw and transformed data are shown in Figure 8.10.

Lexical frequency

Based on results from the raw data, three participants with aphasia performed significantly differently from typical participants. CJ and RE showed a significantly greater effect of lexical frequency than typical participants. PD's performance was significantly different from typical participants, because she showed a reverse effect of lexical frequency.

In the transformed data, CJ and RE also showed greater effects of lexical frequency than typical participants, though the difference reached significance only for CJ. PD showed a reverse effect of lexical frequency, though it was not significant in the analysis of the transformed data. In addition, SP showed a significant reverse effect of lexical frequency, which was the same pattern he demonstrated in the raw data,

Construction frequency

Based on results from the raw data, five participants with aphasia performed significantly differently from typical participants. JF, MJ and PD showed a significantly greater effect of construction frequency than typical participants. CJ and IC performed significantly differently from typical participants, because they showed a reverse effect of construction frequency.

In the transformed data, JF, MJ and PD also showed greater effects of construction frequency than typical participants, though only JF's performance remained significantly different. As in the raw data, IC showed a reverse effect of construction frequency, but this did not reach significance in the transformed data. However, CJ's effect score, which indicated a significant reverse effect of construction frequency in the raw data, showed a non-significant effect of construction frequency in the predicted direction in the transformed data. This switch may have been influenced by the fact that CJ produced only five target responses including verbs with low construction frequency in the task. (See Table 8.19.)

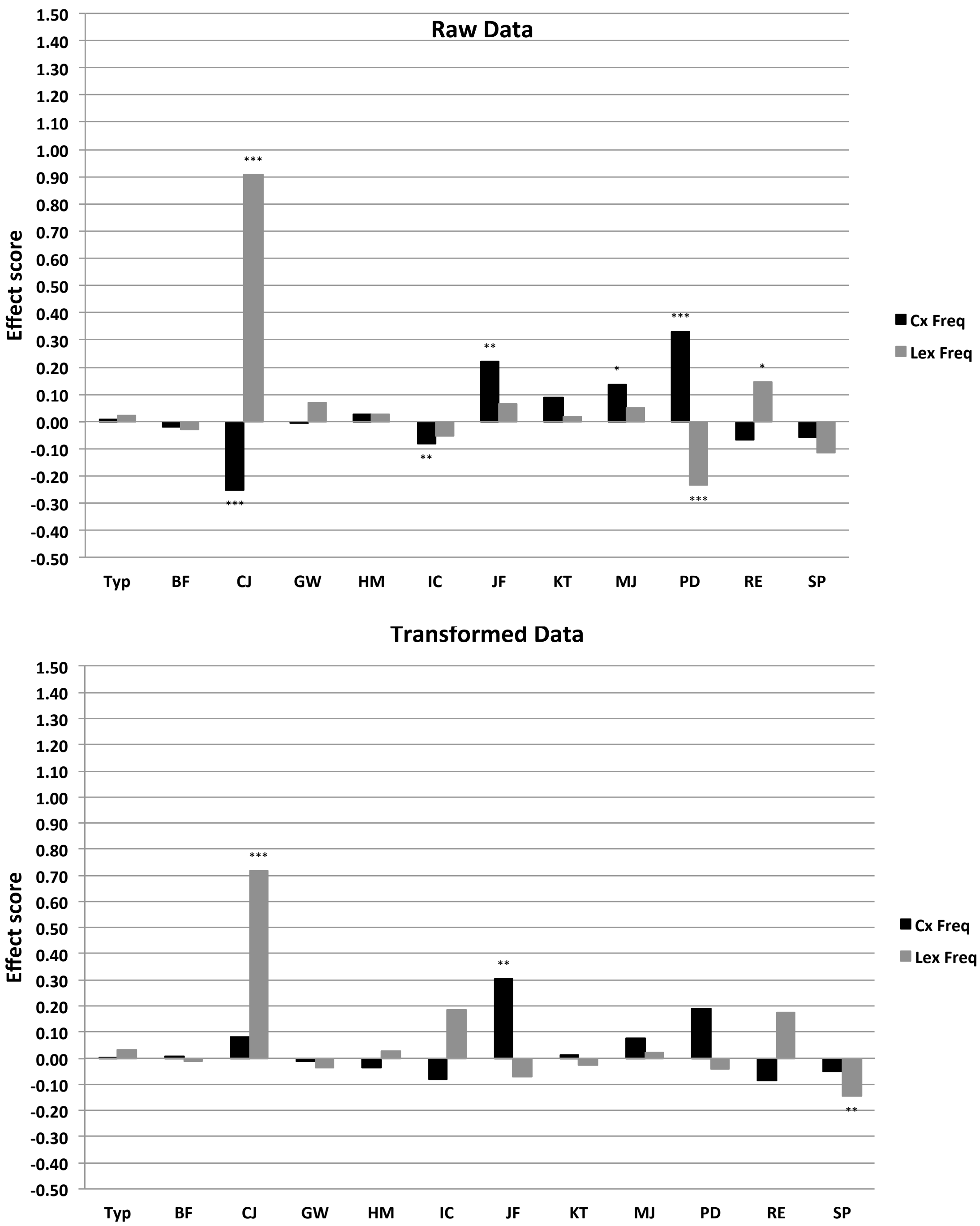


Figure 8.10 Effect scores in sentence completion task

Note. Effect scores for typical participants shown on left (Typ). Asterisks denote significant results from Bayesian Standardised Difference Tests. * $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

8.2.4 Summary of results from participants with aphasia in sentence completion task

This section summarises findings from participants with aphasia in the sentence completion task and compares them to results from typical participants.

Number of target responses

Group-level analyses revealed a significant main effect of lexical frequency on the number of target responses participants with aphasia produced in the sentence completion task, but no main effect of construction frequency. The main effect of lexical frequency reached significance only in the analysis of the removal construction. An interaction revealed a significantly greater number of target responses contained verbs in the [high cx, high lex] condition than the [low cx, low lex] condition. This interaction reached significance only in the analyses of the removal and transitive constructions.

Analyses of the effects of construction frequency and lexical frequency on the number of responses from typical participants were not performed, because only 19 responses from typical participants were excluded due to participant behaviour. The greatest number of these responses (n = 10) were excluded from the [high lex, low cx] condition, however, in contrast to participants with aphasia, who produced the fewest target responses to verbs in the [low cx, low lex] condition.

Verb production latencies

Group-level analyses showed no significant main effects of lexical frequency or construction frequency on verb production latencies from participants with aphasia in the sentence completion task. A significant interaction revealed that production latencies of verbs in the [high cx, high lex] condition were significantly shorter than verbs in the [low cx, low lex] condition. This interaction reached significance only in the analysis of the removal construction.

These findings differ to those from typical participants, whose data yielded a significant main effect of lexical frequency and a significant interaction, where lexical frequency had a significant effect on responses to verbs with high construction frequency but not low construction frequency. Explanatory factors for these results are discussed further in the General Discussion in Chapter 9.

Individuals with aphasia

Four participants with aphasia performed differently from typical participants with respect to the effect of lexical frequency. CJ and RE showed a greater effect of lexical frequency than typical participants; only analyses of the raw data identified RE as significantly different from typical participants, whereas analyses of both the raw and transformed data identified CJ as such. In addition, within-participant analysis revealed the effect of lexical frequency on CJ's verb production latencies was approaching significance. PD and SP showed a reverse effect of lexical frequency; PD's performance was significantly different from typical participants in the analysis of the raw data only and SP in the analysis of the transformed data only.

Five participants with aphasia performed differently from typical participants with respect to the effect of construction frequency. Analysis of the raw data revealed JF, MJ and PD showed a significantly greater effect of construction frequency than typical participants, but only JF's performance remained significantly different from typical participants in the analysis of the transformed data. Furthermore, within-participant analysis identified a significant effect of construction frequency on JF's verb production latencies. CJ and IC demonstrated a reverse effect of construction frequency in the analysis of the raw data; however, CJ did not demonstrate a reverse effect of construction frequency in the analysis of the transformed data.

*

This chapter presented analyses of data from the sentence completion task in Phase 2. The General Discussion follows in Chapter 9.

9 General Discussion

This work was grounded in the usage-based approach to language and investigated the effect of construction frequency and lexical frequency on the processing of verbs and argument structure constructions in typical adults and adults with acquired aphasia. Results confirm that prior linguistic experience of verbs in particular syntactic constructions, as revealed by frequency of occurrence, influences language processing in both typical adults and adults with aphasia.

Phase 1 aimed to examine the relationship between the number of times typical participants and participants with aphasia named verbs that could occur in eight unique argument structure constructions and verbs' lexical frequency and construction frequency. Each frequency measure was positively related to the number of times participants generated verbs in the task, and this finding was discussed in Section 3.4. Phase 1 also established that argument structure constructions, represented as sentences devoid of lexical semantic content, successfully elicited verb responses from adults with aphasia.

Phase 2 examined the effect of lexical frequency and construction frequency on responses from typical participants and participants with aphasia to verbs, given prior exposure to an argument structure construction. Participants took part in a grammaticality judgement task and sentence completion task. Phase 2 aimed to investigate differences and similarities in performance between typical participants and participants with aphasia, as well as to consider how different patterns of language ability in aphasia related to performance in the two tasks. Section 9.2 of this chapter discusses results in light of these aims.

In this chapter, Section 9.1 provides a discussion of results from typical participants in the two tasks of Phase 2, including mechanisms that may give rise to the observed effects, as well as how these may differ between the two tasks. Section 9.2 provides a discussion of results from participants with aphasia in Phase 2, supporting the aims summarised above. Section 9.3 considers the implications for clinical approaches to language in aphasia.

9.1 Discussion of results from typical participants in Phase 2

Phase 2 generated new data that demonstrated the predicted effect of linguistic frequency on sentence processing in typical adults. The current work represents an important departure from past research in this area, which has investigated frequency effects at the sentence level mainly by focussing on multi-word phrases (see Section 2.2.2) or verb bias (see Section 2.2.3). The present research employed a novel methodology which established that the likelihood of a lexical verb, given an argument structure construction, affects language processing. Crucially, sentence materials were abstract and devoid of lexical semantic information, permitting the conclusion that probabilistic relationships between verbs and syntactic constructions were responsible for the observed effects, rather than semantic relationships between lexemes in sentence materials. Only one known previous study, Johnson and Goldberg (2013), designed materials in this way, though they aimed to examine the semantics of argument structure constructions, rather than the relationship between syntactic constructions and lexical verbs.

Previous research on frequency effects in syntax primarily investigated verb bias. Most of this work employed materials containing structurally complex sentences, such as embedded sentences (Ferreira & Schotter, 2013; Garnsey et al., 1997; Jaeger, 2010; Trueswell & Kim, 1998; Trueswell et al., 1993), passive sentences (Street & Dąbrowska, 2014) or relative clauses (Reali & Christiansen, 2007). Fewer studies have considered probabilistic relationships in structurally simpler, single-clause sentences (Gahl & Garnsey, 2004; Wilson & Garnsey, 2009), like those employed in the present studies. The current work adds to this body of evidence that suggests the effect of contextual frequency cannot be attributed to strategic processing of long or complex material. Rather, probabilistic measures appear to affect the language processing system generally.

The following sections examine response mechanisms that may underlie the performance of typical participants and give rise to the frequency effects observed in Phase 2.

9.1.1 Discussion of results from the grammaticality judgement task

9.1.1.1 Main findings

In the Phase 2 grammaticality judgement task, typical participants judged verbs with high construction frequency to be grammatical in a sentence context more quickly than verbs with low construction frequency, and this was identified as a medium or large effect in all analyses. Construction frequency was revealed as a significant main effect in the analysis of all six argument structure constructions in the task and was robust against the effect of age. These findings extend those from Phase 1, where construction frequency had a greater effect than lexical frequency on the number of times typical participants produced verbs in response to most constructions. Taken together, these results confirm the importance of contextual frequency on typical participants' processing of verbs and argument structure constructions.

An interaction between construction frequency and lexical frequency was revealed in by-participant analyses, and it indicated that lexical frequency had an effect in the predicted direction on responses to verbs with high construction frequency, but not low construction frequency. This interaction was evident in responses to the intransitive motion and transitive constructions only.

9.1.1.2 Mechanisms for responding in grammaticality judgement task

This section outlines mechanisms that may account for the observed effects on participants' responses in the grammaticality judgement task.

Source of construction frequency effect in the grammaticality judgement task

Recently, researchers have emphasised the important role of prediction in language processing. In an fMRI study, Johnson et al. (2016) taught adult participants novel argument structure constructions and observed decreased activation in the ventral striatum when construction forms allowed for the prediction of upcoming actions in visual scenes. Based on results from an eye-tracking study, Kamide et al. (2003) discussed the role of prediction in human sentence comprehension as one in which various attributes of sentence elements combine to restrict possible upcoming input. Pickering and Garrod (2013) explained language comprehension as a process in which a listener uses information about the context, their interlocutor and what their own production would be in that situation to predict upcoming language input. Pickering and Garrod (2007), and Dell and Chang (2014), attributed these predictive processes to the production system. The important role of prediction in language processing is consistent with the function of prediction in models of cognitive processing more generally, such as Clark (2016).

Pertinent to the present context is the finding of McRae, Hare, Elman and Ferretti (2005) that typical adults use nouns to generate expectancies about upcoming verbs. In a single-word priming study, they found that nouns which typically occurred as the agent, patient, instrument or location of a verb primed the naming of those verbs. For example, the noun *teeth* primed the verb *brushed*, because it is a typical patient of the verb *brush*, and the noun *towel* primed the verb *drying*, because it is a typical instrument of the verb *dry*. The authors explained their results arose because the nouns and the verbs they primed shared an event schema. The noun activated a set of concepts related to the event encoded by the verb, leading participants to generate a verb in that particular semantic space, resulting in the observed priming effect.

A similar relationship holds between argument structure construction stimuli and verb targets in the present study: both encode the same type of event. Argument structure constructions referred to an abstract event, and the verb targets denoted the event's specifics. Following this logic, the construction activated an event representation shared by the target verb. Phase 1 showed that the association between constructions and verbs varies in strength, because participants were more likely to name some verbs in response to constructions than others. In the Phase 2 grammaticality judgement task, participants responded most quickly to the verbs that were most strongly associated with argument structure constructions. This main effect of construction frequency can be understood as participants' generation of verbs associated with the semantics of an argument structure construction. The reasoning is as follows.

Upon exposure to an argument structure construction, typical participants began to generate acceptable responses in terms of verbs that could occur in each sentence stimulus. Participants were most likely to generate verbs that were most strongly associated with each construction, or high in construction frequency. Participants' shorter response times to verbs with high construction frequency compared to low construction frequency resulted from a 'match' between a verb that participants were expecting, given the construction, and the verb they experienced in the task. Verbs with low construction frequency were only weakly associated with an argument structure construction, so participants took longer to respond to these verbs as a result of a 'mismatch' between a verb expected in the context of a construction and the one experienced in the task. This effect is consistent with the predictions for the study.

Lexical frequency in the grammaticality judgement task

Lexical frequency was revealed as a significant main effect in the by-participant analysis of transformed data from the 86 participants in Phase 2. However, the effect did not survive in an ANCOVA with lexical decision times as a covariate, indicating that the small effect of lexical frequency observed in this instance may have been a product of single-word recognition of the written verbs in the task, consistent with research reporting the effect of lexical frequency on this process (Brysbaert et al., 2000; Grainger, 1990; Morrison & Ellis, 1995).

Lexical frequency may not have influenced participants' responses in the grammaticality judgment task, because the relationship between verbs and argument structure constructions is primarily semantic. In the construction grammar framework outlined in Section 1.2.3, Goldberg (1995) explained how semantic principles constrain the integration of the participant roles of a verb and the argument roles of a construction. Therefore, lexical frequency may not have influenced the hypothesised process of participants' generation of verbs in response to constructions, explained above as the source of the construction frequency effect in the task.

Source of interaction in the grammaticality judgement task

An interaction in the grammaticality judgement task resulted in responses showing the effect of lexical frequency in the predicted direction for verbs with high construction frequency but not low construction frequency, i.e. response times were shorter for verbs with high compared to low lexical frequency when verbs were high in construction frequency, but not when verbs were low in construction frequency. The interaction reached significance in the analysis of the intransitive motion and transitive constructions only. In these two cases, response times were longest to verbs in the [low cx, high lex] condition. Notably, this condition elicited the greatest number of non-target responses for four of the six constructions in the task.

The interaction may have arisen because materials in Phase 2 were not controlled for verb bias, explained in Section 2.2.3 as the likelihood of a verb appearing in a particular syntactic context over others. The verbs in the [low cx, high lex] condition for the intransitive motion and transitive constructions, *lead* and *try*, respectively, may have a bias that conflicted with the argument structure construction with which they were associated in the task. In the task, the verbs appeared in the intransitive sentence, *we lead through there*, and the transitive sentence, *you try them*. Priming literature indicates that verb bias can be derived within 39 milliseconds of encountering a written verb form (Trueswell & Kim, 1998). The exceptionally long response times to [low cx, high lex] verbs in the grammaticality judgement task may therefore result from two conflicts:

(1) a mismatch between the verbs expected to occur, given the argument structure construction, and the target verb with low construction frequency; and

(2) a mismatch between the syntactic context corresponding to the verb bias of the target verb and the syntactic context of the argument structure construction provided in the task.

The effect of verb bias may also explain the interaction in the caused motion construction, where no effect of construction frequency was observed for the pair of verbs with high lexical frequency. Participants responded to the [low cx, high lex] verb in the sentence *they chuck it to me* with similar response times as for verbs with high construction frequency. In this instance, the verb *chuck* may have a strong bias to occur in the caused motion construction; indeed, there may be few alternative syntactic constructions in which it can occur. This strong bias may have caused participants to respond to it in a similar way as verbs with high construction frequency.

9.1.2 Discussion of findings from the sentence completion task

9.1.2.1 Main findings

Contrary to predictions, findings from the sentence completion task differed from findings from the grammaticality judgement task. In the sentence completion task, a main effect of lexical frequency arose in by-participant analyses only. An interaction showed that this effect of lexical frequency was significant only on responses to verbs with high construction frequency, but not low construction frequency. The interaction reached significance in both the raw and transformed data only in the analysis of the ditransitive construction. Additionally, the pattern of the interaction was evident in responses to the caused motion, intransitive motion and removal constructions, but did not reach significance in these instances. Effects in the sentence completion task were variable for each construction and were not robust against effects of age and gender.

9.1.2.2 Mechanisms for responding in sentence completion task

This section identifies aspects of the sentence completion task that influenced participants' responses, particularly accounting for the lack of a main effect of construction frequency. Because responses in this task were more varied for each construction than responses in the grammaticality judgement task, results are discussed with reference to the six individual argument structure constructions in the task.

Source of lexical frequency effect in the sentence completion task

The main effect of lexical frequency in the sentence completion task was significant in by-participant analyses collapsed across constructions. It represented a medium or large effect for the ditransitive and intransitive motion constructions. Unlike in the grammaticality judgement task, analysis showed that verb production latencies were unrelated to lexical decision times of verbs, indicating that the effect of lexical frequency could not be attributed solely to the process of single-word recognition.

The effect of lexical frequency in the sentence completion task may have arisen due to the requirement to produce phonological word forms. Jescheniak and Levelt (1994) concluded that lexical frequency is a property of the retrieval of a phonological word form, rather than the retrieval of a word's semantic or syntactic properties. In Phase 2, no effect of lexical frequency was observed in the grammaticality judgement task, arguably because that task required no phonological output.

This conclusion raises the question of why lexical frequency was not a significant effect in the analysis of all six argument structure constructions in the sentence completion task. Jescheniak and Levelt's investigation involved only single-word processing tasks. In contrast, participants in the current study may have begun to generate some expectations of upcoming verb targets upon exposure to argument structure construction stimuli (see Section 9.1.1.2), so their response times to verbs did not represent the sole process of single-word production.

Source of interaction in the sentence completion task

An interaction in the sentence completion task revealed a significant effect of lexical frequency on the production of verbs with high construction frequency but not low construction frequency. The interaction represented the largest effect on responses to the ditransitive construction and was demonstrated to a lesser extent on responses to the caused motion, intransitive motion and removal constructions.

This interaction could be argued to be similar to the interaction observed in the analysis of the grammaticality judgement task, where a typical effect of lexical frequency was produced in response to verbs with high construction frequency but not low construction frequency. Because of the consistency of this finding across tasks, one possible interpretation of the interaction relates to the degree of 'surprisal' verbs introduced to the language processing system, where surprisal is a function of a lexical item's probability in a given context (Hale, 2001). In the present work, verbs with high construction frequency have low values of surprisal, while verbs with low construction frequency have high values of surprisal. Low surprisal contexts can facilitate lexical processing (Levy, 2008), which may explain why a typical effect of lexical frequency was observed in the present work only in response to verbs with high construction frequency.

However, Section 9.1.1.2 explained the interaction in the grammaticality judgement task as arising due to responses to specific verbs, because the interaction was not evident in responses to each construction. A similar explanation applies to the interaction in the sentence completion task (see paragraph below). More data, in terms of the number of verbs associated with each construction in the research design, may be necessary in order to make reliable claims about whether the interaction observed in the present work reflects a real interaction between contextual frequency and single-word frequency measures in language processing. The present results do not exclude an interpretation relating to item-specific responses.

The interaction reached significance in the evaluation of the sentence completion task only for the ditransitive construction, where participants produced the [high cx, low lex] verb OWE with the longest latencies. The long processing times associated with the verb OWE may reflect the fact that OWE does not literally encode the event referred to by the argument structure construction. Section 1.2.3.3 explained how argument structure constructions can be polysemous, and the use of OWE in the ditransitive construction results in a sentence that refers to potential or expected transfer, rather than the construction's central sense of actual transfer. Additionally, the lexeme OWE proved to be an outlier in other, additional analyses, as it attracted the longest response times in Keuleers et al.'s (2012) lexical decision task and was the latest acquired verb in the task. These findings converge to suggest that OWE may have required longer processing for semantic, orthographic and developmental reasons.

Construction frequency in the sentence completion task

Construction frequency was not identified as a significant main effect in the analyses collapsed across argument structure constructions; however, construction frequency proved to be a significant main effect in by-participant analysis of the intransitive motion construction, where it represented a small effect, and the removal construction, where it represented a medium effect. For verbs with high lexical frequency in the ditransitive construction, and for both pairs of verbs in the conative construction, mean verb production latencies were in the predicted direction, but did not reach statistical significance. Therefore, participants' responses to most verbs in the sentence completion task were consistent with the predicted effect of construction frequency, but the effect did not reach significance like it did in the grammaticality judgement task. Section 9.1.3 considers why effects on participants' responses differed between the two tasks in Phase 2.

Participants' responses to two constructions in the sentence completion task did not show the predicted effect of construction frequency. Responses to the caused motion construction showed no significant effect of either type of frequency. In this instance, the low construction frequency verbs *tell* and *drive* may have had a strong bias that matched the construction, leading participants to respond to all four verbs in the caused motion construction similarly.

Figure 9.1 An instance of the verb *drip* in the transitive construction

Responses to the transitive construction revealed a significant reverse effect of construction frequency. This response pattern echoed responses to the transitive construction in the grammaticality judgement task, where participants showed no effect of construction frequency in response to verbs with low lexical frequency. In Phase 1, the transitive construction elicited one of the greatest numbers of verb types from typical participants (n = 164). Corresponding to the broad range of verb types that can meaningfully occur in this construction, the syntactic form of the transitive construction can encode two distinct semantic events. As described in Section 2.1.1.3 and Section 3.2.4.2, the subject of a transitive sentence containing an unergative verb encodes an agent, such as *the puppy licked the child*, whereas the subject of a transitive sentence containing an unaccusative verb encodes a theme, such as *the grocer sneezed*. Therefore, the form of the transitive construction does not refer to a single semantic event, which may be necessary to drive participants' expectations of the verbs that may appear in the task, described in Section 9.1.1.2 as the mechanism underlying the effect of construction frequency. Furthermore, findings from Kako (2006), discussed in Section 1.3.1, suggest that even transitive sentences containing intransitive or nonsense verbs are interpretable. These attributes of the transitive construction converge to suggest that the construction may not be strongly associated with any particular verb types, because it represents the basic English SVO sentence structure, and the same form can refer to distinct semantic events. Figure 9.1 demonstrates how even a prototypically intransitive verb such as *drip* can appear in the transitive construction, given a particular context.

9.1.3 Differences in Phase 2 task demands affect responses to construction frequency

Construction frequency yielded a significant main effect on participants' responses in the grammaticality judgement task, but not the sentence completion task. This may have resulted from differences between the demands imposed by the two tasks, including the required degree of semantic processing of the sentence stimuli and the required response, as well as individual differences among participants.

Sentence semantics may have influenced participants' responses to a greater extent in the grammaticality judgement task than in the sentence completion task. In responding to whether verbs could be used in argument structure constructions in the grammaticality judgement task, participants may have been influenced by whether the resulting sentence was meaningful or not, because verbs that were included in the task to elicit 'no' responses did not result in meaningful sentences. In contrast, all verbs in the sentence completion task resulted in meaningful sentences. The grammaticality judgement task therefore demanded greater engagement with the semantics of the sentence stimuli than the sentence completion task.

In response to this difference between the tasks, participants may not have derived event-level meaning from the sentence stimuli prior to revealing verb targets in the sentence completion task, affecting the response generation mechanism that may have driven the effect of construction frequency (see Section 9.1.1.2). The sentence completion task could in fact be accomplished without processing the semantics of the sentence stimulus prior to revealing the target verb; participants may have displayed both components before reading aloud the entire sentence. Such a response strategy is compatible with research showing that reading - rather than an inevitably automatic process - is a behaviour subject to task demands. Reading times differ between tasks that emphasise speed or accuracy (Stine-Morrow, Shake, Miles & Noh, 2006), and whether single-word reading occurs depends on participants' intention, given task demands, in the case of visual word recognition (Risko, Stolz & Besner, 2005) and reading aloud (O'Malley & Besner, 2012; Reynolds & Besner, 2006). This evidence suggests that simple exposure to a written sentence stimulus in the sentence completion task may have been insufficient for all participants to trigger the response generation mechanism that could have driven the effect of construction frequency.

Furthermore, the task of reading aloud does not necessarily require access to semantics. One model of reading aloud purports that the task can be accomplished via a direct relationship between spelling and sound, without access to semantics (Seidenberg & McClellan, 1989). All models of reading agree that linguistic form is assessed prior to semantics in reading (Coltheart, 2005; Just & Carpenter, 1980). The delayed processing of semantic information compared to form-based information suggests that verb production latencies may not reflect an effect of construction frequency as sensitively as might a measure of later stages of sentence processing, for example, when verbs and constructions have been meaningfully integrated. Research using eye-tracking has shown that properties of single words can affect the processing of constituents later in the sentence (Stites & Federmeier, 2015), and this effect may be exacerbated in older compared to younger adults (Lee & Federmeier, 2012). The effect of construction frequency may have been greater in the sentence completion task if assessed by a secondary task delivered during or immediately after sentence production, or using methods more sensitive to online processing, such as eye-tracking. Future research could investigate such a possibility.

Finally, individual differences may have influenced how participants processed materials in the sentence completion task. Individual variation in reading comprehension has been attributed to problem-solving abilities and language experience, including print exposure and vocabulary size (Freed, Hamilton & Long, 2017). Individuals with greater levels of print exposure show better single-word recognition and lexical access during reading (Lowder & Gordon, 2017). Individuals rely to differing extents on accessing semantics when reading aloud words with exceptional spelling-to-sound correspondences, such as *pint* (Woolams, Lambon Ralph, Madrid & Patterson, 2016). Such variation in typical participants may have contributed to different strategies of responding in the sentence completion task.

In sum, because the sentence completion task did not require a response related to the semantics of the sentence stimulus, participants may have adopted a task-dependent strategy in which they did not extract meaning from sentence stimuli in advance of processing verb targets. This strategy may have been adopted to a greater or lesser extent by individuals in the sample, leading to an overall smaller and more variable effect of construction frequency in the sentence completion task than the grammaticality judgement task.

9.1.4 Effects of age and gender

In the grammaticality judgement task, age was weakly correlated to participants' response times, and the correlation was stronger for high construction frequency conditions, which elicited the shortest response times. Both women and men evinced this response pattern. Age-related increases in response times in forced-choice tasks have been observed in adults up to the age of 89 (Bugg, Zook, DeLosh, Davalos & Davis, 2006; Deary, Liewald & Nissan, 2011). Woods, Wyma, Yund, Herron and Reed (2015) revealed that 80% of age-related increases in response times is due to slowing in the time required to make a selection, and the remaining increase is due to slowing in the time taken to initiate movement. Because the effect of age was greatest when motor responses were quickest, the effect of age in the grammaticality judgement task may have arisen from age-related slowing of motor responses.

In the sentence completion task, two-way interactions between construction frequency, lexical frequency and age and gender arose. One interaction showed that women's verb production latencies increased with age in all four conditions in the task, but men's did not. A second interaction indicated that the effect of lexical frequency on the production of verbs with high construction frequency represented a large effect for women, but a small effect for men. The remainder of this section considers reasons for these findings in more detail.

With regard to the first interaction, the extraction of meaning from sentences must be recognised as a cognitively demanding process (Stine-Morrow, Miller & Hertzog, 2006).

Older readers allocate more time in sentence reading to conceptual integration, the process by which event-level meaning is derived from sentences, in the context of a secondary task (Smiler, Gagne & Stine-Morrow, 2003). Compared to younger adults, older adults process sentences less effectively, as shown by smaller, delayed ERP responses to highly predictable words in a sentence context (Federmeier & Kutas, 2005) and longer reading times (Stine & Hindman, 1994). As Section 9.1.3 explained, this demanding process is not necessarily required in the sentence completion task. Women are also more resilient against age-related cognitive decline than men (McCarrey, An, Kitner-Triolo, Ferrucci & Resnick, 2016), suggesting that they may have more cognitive resources available generally than men. Additionally, women may have more familiarity with written text and thus more regularly extract semantics from written language due to higher rates of reading among older women as compared to older men (Jacobs, Hammerman-Rozenberg, Cohen & Stessman, 2008). Therefore, women may adopt a strategy in the sentence completion task that involves the semantic processing of written materials to a greater extent than men.

This argument reduces to the identification of sex differences in cognitive self-regulation, where cognitive self-regulation refers to unconscious 'decisions about allocation of effort, selection of processing strategies…and the speed at which [a] task should be completed' (Stine-Morrow et al., 2006, p. 585). In the sentence completion task, self-regulation strategies may have differed between men and women, as men allocated less effort to the more demanding task of extracting semantics from written stimuli, but women accomplished this process due to more robust cognitive function and greater familiarity with text than men. Little research has been devoted to examining sex differences in this area, but cognitive self-regulation improves with age (Hennecke & Freund, 2010), which may explain why such a difference arose in the population of older adults included in the study.

Women's production latencies increased with age in the sentence completion task, because the semantic system expands with age. Older adults have larger vocabularies than younger adults (Verhaeghen, 2003). Consequently, longer response times in older women reflected the memory search of a semantic system of increased size compared to younger women, according to the view of ageing advanced by Ramscar and colleagues (Ramscar, Hendrix, Love & Baayen, 2013; Ramscar, Hendrix, Shaoul, Milin & Baayen, 2014). These researchers interpret the positive relationship between age and response times in information processing tasks as the cost of learning in an experience-dependent cognitive system. In the context of verb and argument structure processing, older adults continue to encounter new words used as verbs and experience verbs in different argument structure constructions over the course of the 30 years that participants' ages spanned in Phase 2. As described above, it is possible that age did not affect men's production latencies in the sentence completion task because men may not have engaged with the meaning of the task materials to the same extent as women. A similar interaction between age and gender may not have arisen in the grammaticality judgement task because there could have been less variability in response strategies in the task, given that responses demanded more recourse to semantics than in the sentence completion task.

A second interaction in the sentence completion task revealed that the effect of lexical frequency on the production of verbs with high construction frequency represented a large effect for women, but a small effect for men. This finding may result from higher reading rates in older women than men, as mentioned above (Jacobs et al., 2008). Increased exposure to words with high lexical frequency could cause women's large effect of lexical frequency in the task, whereas men's comparatively decreased exposure would lead to a smaller effect of lexical frequency. As discussed in Section 9.1.2.2, this interaction may not have been observed in the grammaticality judgement task because no phonological output was required.

9.2 Discussion of results from participants with aphasia in Phase 2

This section discusses findings from participants with aphasia in Phase 2. Group-level results from participants with aphasia are addressed in Section 9.2.1, where they are compared to findings from typical participants, in line with the second aim of Phase 2. In support of the third aim of Phase 2, individual participants with aphasia are discussed in Section 9.2.2 with reference to their profiles of language abilities. The clinical implications of findings from Phase 1 and Phase 2 follow in Section 9.3.

9.2.1 Discussion of group-level findings from participants with aphasia

9.2.1.1 Discussion of findings from the grammaticality judgement task

Construction frequency had a significant main effect on the number of target responses and response times from participants with aphasia in the Phase 2 grammaticality judgement task. Participants with aphasia produced significantly more target responses to verbs with high construction frequency than verbs with low construction frequency for four of the six constructions, and their response times were shorter to verbs with high construction frequency than verbs with low construction frequency for all six constructions, where the difference reached significance for the intransitive motion construction.

The effect of lexical frequency on responses in the grammaticality judgement task was less robust. Participants with aphasia produced significantly more target responses to verbs with high lexical frequency than low lexical frequency for only one of the six constructions. Participants produced shorter response latencies to verbs with high compared to low lexical frequency for three of the six constructions, but none of these differences reached significance.

Analysis of the number of target responses revealed an interaction that indicated a significant effect of lexical frequency on responses to verbs with high construction frequency, but not low construction frequency. This interaction was revealed in the analysis of typical participants' response times.

Results from participants with aphasia were consistent with findings from typical participants, with reference to the main effect of construction frequency and its interaction with lexical frequency. Section 1.2.3 described how the present work interrogated the links between verbs as single words and syntactic constructions in the network architecture of grammar. The significant effect of construction frequency on responses from participants with aphasia suggests that these 'lexical links' can remain intact in the language system of adults with aphasia, and the strength of the association between verbs and syntactic constructions is moderated by frequency. Findings support Gahl and Menn's (2016) claim that 'the effects of past linguistic experience are evident in the language of people with aphasia, just as they are in the language of all other speakers' (p. 1373).

Results from the Phase 2 grammaticality judgement task extend findings from Phase 1, which identified some evidence for the effect construction frequency in aphasia. Section 4.1.1 summarised some ways in which the tasks in Phase 2 may have been more accessible to participants with aphasia than the verbal fluency task in Phase 1. Given the reduced demands of the grammaticality judgement task in Phase 2, the effect of the relationship between verbs and syntactic constructions is clearly apparent: together with results from Phase 1, results from the grammaticality judgement task reveal that frequency effects at grain sizes larger than the single word can affect language processing in aphasia. This conclusion addresses the gap in the literature on multi-word frequency effects in aphasia identified in Section 2.2.2 and complements previous findings on the effect of verb bias on reading in aphasia (DeDe 2013a, 2013b; Gahl, 2002; Gahl et al., 2003). Linguistic experience, in terms of how frequently adults have experienced verbs in syntactic contexts, affects language processing in aphasia.

9.2.1.2 Discussion of findings from the sentence completion task

At the group level, lexical frequency had a significant main effect on the number of target responses participants with aphasia produced in the sentence completion task. Greater numbers of target responses to verbs with high compared to low lexical frequency were observed in response to five of the six constructions, where the difference reached significance for the removal construction. Participants demonstrated shorter verb production latencies of high lexical frequency verbs in response to five of the six constructions, but these differences did not reach significance. These findings are consistent with those from typical participants, who demonstrated a significant effect of lexical frequency in the sentence completion task, as measured by verb production latencies.

In contrast, construction frequency did not have an effect on the number of target responses or verb production latencies at the group level for participants with aphasia. Participants produced more target responses to verbs with high construction frequency than low construction frequency in response to four of the six constructions, but none of the differences reached significance. Participants produced high construction frequency verbs with shorter latencies than low construction frequency verbs in response to three of the six constructions, and the difference reached significance for only the transitive construction. The lack of a main effect of construction frequency in the sentence completion task agreed with findings from typical participants.

A significant interaction revealed that participants with aphasia produced more target responses and shorter verb production latencies to verbs in the [high cx, high lex] condition than the [low cx, low lex] condition. This contrasts with the significant interaction from typical participants, which Section 9.1.2.2 attributed in part to long production latencies of the verb OWE in the ditransitive construction. In contrast, participants with aphasia showed no significant differences in target responses or verb production latencies in response to the ditransitive construction. Rather, the interaction between construction frequency and lexical frequency may signify that the sentence context provided in the task prior to the presentation of the verb supported the processing of verbs with high lexical frequency, whose associations between semantics and phonology may be well-preserved (DeDe; 2012; Yap, Tse & Balota, 2009). The sentence context did not support the processing of verbs with low construction frequency, and responses in the [low cx, low lex] condition were further hindered by verbs with low lexical frequency, which may have weaker connections between semantics and phonology. (See Section 9.2.2.3 for further discussion of this interaction.)

9.2.2 Discussion of individual participants with aphasia in Phase 2

This section discusses individual participants with aphasia in more detail, examining their profile of language abilities in relation to the within-participant analyses reported in Section 7.2.2.4 and Section 8.2.2.4, and results from Bayesian Standardised Difference Tests (BSDTs) that compared response times of individual participants with aphasia to typical participants, reported in Section 7.2.3 and Section 8.2.3. Participants with aphasia who produced reliable effects in the analyses of both raw and transformed data, in that the effects were in the same direction in both analyses and effect scores were consistently greater or less than those of typical participants, were identified in Section 7.2.3 and Section 8.2.3.

Only GW and HM are not included in this discussion. GW showed a significantly smaller effect of construction frequency than typical participants in the grammaticality judgement task based on the analysis of raw data, but this effect was non-significant and of greater magnitude than typical participants in the analysis of the transformed data. No other effect score for lexical frequency or construction frequency from GW or HM was significantly different from typical participants in the Phase 2 tasks.

9.2.2.1 Construction frequency in individuals with aphasia

JF - Effect of construction frequency in both tasks

JF showed significant effects of construction frequency in both tasks. Within-participant analyses revealed a large effect of construction frequency on response times in the grammaticality judgement task and a medium effect of construction frequency on verb production latencies in the sentence completion task. Insofar as JF showed an effect of construction frequency in the sentence completion task, he was atypical, as he produced a significantly greater effect of construction frequency than typical participants, as revealed by BSDTs in the analysis of both the raw and transformed data. JF was diagnosed with anomic aphasia. His mild language impairment did not affect verb or sentence comprehension, naming or sentence processing.

JF's pattern of performance was similar to findings from research on the effect of verb bias on reading in aphasia. In self-paced reading studies, DeDe (2013a) reported that four participants with aphasia showed a larger effect of verb bias than typical participants when reading sentences containing verbs with transitive or intransitive biases. DeDe (2013b) found that four participants with aphasia showed a larger effect of verb bias than typical participants when reading sentences containing verbs with direct object or sentence complement biases. Results from these studies suggest that the strength of the relationship between verbs and the syntactic contexts in which they appear can affect some individuals with aphasia to a greater degree than typical participants, in both silent reading (DeDe 2013a, 2013b) and reading aloud (the present study). Future research could explore why some individuals with aphasia appear more sensitive to the relationship between verbs and syntactic contexts than others.

KT - Effect of construction frequency in grammaticality judgement task

KT showed a greater effect of construction frequency than typical participants in the grammaticality judgement task, and within-participant analysis indicated this represented a moderate effect size that was approaching significance. KT's verb production latencies in the sentence production task were in the direction of this effect but did not reach significance. KT was diagnosed with fluent transcortical sensory aphasia, and she showed no evidence of verb or sentence processing impairments.

MJ and PD - Effect of construction frequency in sentence completion task

Like JF, MJ and PD showed a greater effect of construction frequency in the sentence completion task than typical participants. Unlike JF, differences in verb production latencies to verbs with high and low construction frequency did not reach significance in within-participant analysis. However, the effect of construction frequency on the response times of MJ and PD in the grammaticality judgement task was approaching significance and represented a moderate effect size for both MJ and PD.

Along with JF and KT, MJ and PD represent the diversity of individuals with aphasia who show effects of construction frequency to a greater extent than typical adults. MJ was diagnosed with non-fluent Broca's aphasia, and PD was diagnosed with fluent conduction aphasia. Unlike JF and KT, MJ and PD both showed impaired naming of nouns and verbs, in addition to performing outside the normal range in the anagram task.

SP, UT and VH - Reverse effect of construction frequency in grammaticality judgement task

Three participants with aphasia performed significantly differently than typical participants in the grammaticality judgement task, because their response times to verbs with high construction frequency were longer than their response times to verbs with low construction frequency. These differences did not reach significance in within-participant analyses. This section will consider possible causes of their performance with reference to the mechanisms of responding that were outlined for typical participants in Section 9.1.1.2.

Within the group of participants with aphasia, SP, UT and VH produced the lowest number of target responses in the grammaticality judgement task. SP was also one of two participants who produced the lowest number of target responses in the sentence completion task. In both tasks, SP produced verbs with high construction frequency with longer latencies than verbs with low construction frequency. However, this difference was identified as significantly different than typical participants only in the grammaticality judgement task, and it did not reach significance in within-participant analysis. UT and VH did not complete the sentence completion task.

SP, UT and VH were unique in the group of participants with aphasia in Phase 2 to have impaired naming and word fluency scores in the context of verb and sentence comprehension deficits. All three participants showed impaired comprehension of written sentences and verbs. These problems indicate difficulty in deriving meaning from task stimuli, including the extraction of event-level meaning from sentence stimuli and lexical meaning from verb targets. Additionally, these three participants produced the lowest scores in the word fluency subtest of the *Comprehensive Aphasia Test* (CAT), and SP and UT showed impaired naming of both actions and objects. The combination of impairments in word fluency and naming suggests a deficit in the ability to access word forms from the semantic system. Section 9.1.1.2 argued that the generation of verbs in response to a meaningful argument structure construction drove the effect of construction frequency in the grammaticality judgement task, due to a match between participants' expectations and their experience in the task. SP, UT and VH may have been unable to generate verb forms in response to a construction because of problems extracting meaning from sentence stimuli, as well as having reduced resources for that meaning to activate related verb forms.

Results from SP, UT and VH are consistent with findings from TP in Phase 1. TP produced the lowest scores in the Phase 1 language assessments, and most of his responses in the spontaneous verbal fluency task were judged to be ungrammatical. Section 3.4.2.1 interpreted this as evidence for weakened lexical links in the network architecture of grammar. A similar conclusion applies to the performance of SP, UT and VH in Phase 2. For these more severely impaired participants, the frequency with which verbs appeared in argument structure constructions had no effect on language processing.

IC - Reverse effect of construction frequency in sentence completion task

IC produced longer response times to verbs with high construction frequency than low construction frequency in the sentence completion task, but the difference did not reach significance in within-participant analysis or by a BSDT on transformed data. IC did not show this pattern of performance in the grammaticality judgement task.

IC presented as a typical case of conduction aphasia. The discussion of LM in Phase 1 in Section 3.4.2.1 concluded that grammatical processing can be intact in individuals with phonological coding deficits, like those in conduction aphasia. IC's response times in the sentence completion task, a more sensitive measure than that used in Phase 1, suggest that the effect of construction frequency on language processing in conduction aphasia may be modality-specific, because he showed a typical effect of construction frequency in the grammatically judgement task, but not the sentence completion task. However, because IC's reverse effect of construction frequency in the sentence completion task was not robust in all analyses, further research is needed.

9.2.2.2 Lexical frequency in individuals with aphasia

RE - Effect of lexical frequency in both tasks

Compared to typical participants, RE showed a greater effect of lexical frequency than typical participants in both Phase 2 tasks. This result is consistent with DeDe's (2012) finding that lexical frequency can affect sentence reading in aphasia to a greater extent than in typical participants. Four participants in DeDe's study showed greater effects of lexical frequency in a self-paced reading task than typical participants, and they produced shorter response times to high frequency words. DeDe (2012; Yap et al., 2009) interpreted large effects of lexical frequency as indicating weakened lexical representations, in terms of associations between phonological and semantic representations. This conclusion could explain RE's errors in single-word comprehension and naming tasks, described in Section 6.4.3. Problems in accessing the phonological word form from a semantic representation could lead to the semantic errors RE produced in single-word processing tasks.

CJ - Effect of lexical frequency in sentence completion task

In the sentence completion task, CJ showed a significantly greater effect of lexical frequency than typical participants, in the analysis of both raw and transformed data. Within-participant analysis identified lexical frequency as a medium effect on verb production latencies, which was approaching significance. CJ showed no effect of lexical frequency on response times in the grammaticality judgement task.

Section 6.5.5 identified CJ as having a reading difficulty that affected the production of single words. All his errors in the single-word reading subtest of the CAT involved the production of a word orthographically related to the target. As noted in Section 9.1.2.2, lexical frequency may be interpreted as a property of the phonological word form (Jescheniak and Levelt, 1994). CJ could have shown an effect of lexical frequency in the sentence completion task due to the demands of reading aloud. However, the other two participants observed in Section 6.5.5 as having reading difficulties, MJ and PD, did not show an effect of lexical frequency. The current findings therefore suggest that lexical frequency does not affect the language production of all individuals with aphasia who have single-word reading difficulties.

BF and WD - Reverse effect of lexical frequency in grammaticality judgement task

BF and WD produced longer response times to verbs with high lexical frequency than low lexical frequency in the grammaticality judgement task. Their performance was significantly different from typical participants in the raw data, but the reverse effect of lexical frequency was not significant in within-participant analyses. BF was diagnosed with fluent anomic aphasia and achieved the highest score of all participants with aphasia in Phase 2 in the written sentence comprehension subtest of the CAT. In contrast, WD was diagnosed with Wernicke's aphasia and produced the lowest score of all participants in this assessment.

A lack of typical lexical frequency effects in language comprehension in aphasia has been analysed as impairment to semantic control mechanisms that ensure the correct meaning of a word is accessed when needed (Hoffman et al., 2001; see Section 2.2.1.1). The current study did not include a range of semantic assessments, making it difficult to evaluate this claim in the present context. However, BF performed within the normal range on all input tasks in the CAT, whereas WD often selected the semantic distractor (see Section 6.4.3).

PD and SP - Reverse effect of lexical frequency in sentence completion task

PD and SP produced longer verb production latencies in response to verbs with high lexical frequency than low lexical frequency in the sentence completion task. This reverse effect of lexical frequency was not significant in within-participant analysis, which indicated an insubstantial effect for PD but a medium effect for SP.

PD was diagnosed with conduction aphasia but, as described in Section 6.4.3, produced responses in reading aloud tasks that were indicative of surface dyslexia. Section 9.1.2.2 attributed a lexical frequency effect to the production of phonological output, but it does not appear that PD's reading strategy was sensitive to whole-word frequency. Anecdotally, she depended on letter-to-sound correspondences in reading aloud, which may account for the practically negligible effect of lexical frequency on her verb production latencies.

Reverse effects of lexical frequency in language production in aphasia have been described with relation to the complexity of semantic content encoded by verbs (Breedin et al., 1998) or the distinctiveness of the referent encoded by nouns (Marshall et al., 2001) (see Section 2.2.1.2). Given that targets in the Phase 2 sentence completion task were all lexical verbs, it is unclear how this semantic explanation might apply to the performance of SP, who was diagnosed with Wernicke's aphasia. As mentioned above, further assessments of semantic processing would be needed to interrogate this possibility.

9.2.2.3 Interface between lexical frequency and construction frequency in aphasic sentence processing

Across the two tasks in Phase 2, results from participants with aphasia indicated better performance to verbs in the [high cx, high lex] condition than verbs in the [low cx, low lex] condition. Group-level analyses revealed a difference between these two conditions that was significant or approaching significance for the number of target responses and response times in both the grammaticality judgement task and sentence completion task. The number of target responses was greatest to [high cx, high lex] verbs and lowest to [low cx, low lex] verbs in both tasks, and median response times in both tasks were shortest to [high cx, high lex] verbs and longest to [low cx, low lex] verbs. These differences represented a medium or large effect in all instances.

Several individuals with aphasia also showed this pattern of responses. In terms of the number of target responses, HM, KT, MJ, PD, RE, UT VH and WD produced this response pattern in the grammaticality judgment task; IC in the sentence completion task; and SP in both tasks. In terms of response times, IC, KT and RE produced this response pattern in the grammaticality judgement task; GW in the sentence completion task; and HM in both tasks. These findings suggest that, in combination, high levels of both construction frequency and lexical frequency functioned to ease sentence processing in aphasia, while low levels of frequency functioned to tax sentence processing. At the sentence level, experience of both verbs as single words, as well as verbs in particular syntactic contexts, affected sentence processing in aphasia.

These findings illuminate research that reported inconsistent effects of frequency on verb processing, introduced in Section 2.2.1.2. Bastiaanse et al. (2016) concluded that lexical frequency does not affect verb retrieval. The authors attributed this finding to the complexity of a verb's lemma, to which they ascribed all information regarding argument structure and thematic roles. Gahl and Menn (2016, p. 1368) noted that 'verb-specific frequency counts' hold more promise than lexical frequency in encoding the effect of linguistic experience with respect to verbs. The current work supports this hypothesis: verb-specific frequency counts in terms of how strongly verbs are associated with an argument structure construction, in conjunction with lexical frequency, proved to affect sentence processing in aphasia in Phase 2. Future research could expand the number of verbs included in sentence processing tasks to further investigate the effects of construction frequency and lexical frequency, using a regression-based design to complement the factorial design employed in this project.

Clinically, this finding indicates that language interventions for aphasia which target the sentence level may benefit from accounting for both the lexical frequency of verbs as single words, in addition to the contextual frequency of verbs in specific syntactic contexts. These measures could be manipulated to ease processing or increase the challenge of therapeutic materials for clients. Further clinical implications arising from the present work are discussed in the following section.

9.3 Clinical implications

9.3.1 Construction-based approaches to the treatment of sentence processing deficits in aphasia

Materials in the current study were based on eight syntactic constructions identified in Goldberg's (1995; Johnson & Goldberg, 2013) construction grammar framework. In Phase 1, these constructions successfully elicited verbs from participants with aphasia. In the Phase 2 grammaticality judgement task, participants with aphasia were able to distinguish between verbs that could and could not occur in these constructions. These results demonstrate that this set of argument structure constructions could usefully expand the number of constructions that currently form the basis of treatment for sentence processing deficits in aphasia.

Current interventions targeting sentence processing deficits in aphasia focus on the level of syntactic constructions, but the selection of these syntactic constructions has been informed by generative linguistic theory. Treatment of Underlying Forms (TUF) is a programme that capitalises on similarities among abstract, non-canonical syntactic constructions, including the passive (e.g *Alex was kissed by Sam*), object cleft sentences (e.g. *it was Alex who Sam kissed*) and wh-questions (e.g. *who did Sam kiss*). TUF has been successful in improving sentence comprehension and production in individuals with agrammatic Broca's aphasia (Thompson & Shapiro, 2005). However, it is unclear how closely the constructions contained in the TUF intervention programme align with the clinical needs and goals of clients with aphasia.

Clients, family members and therapists identify communication as an important area for aphasia rehabilitation (Pettit, Tönsing & Dada, 2016), and most clients and their families articulate clinical goals as functional outcomes (Wallace et al., 2016). Sherratt et al. (2011) noted that communication goals often targeted spoken communication, rather than written language. These findings raise the question of whether the non-canonical sentence structures included in TUF are the most appropriate linguistic targets for remediation, because they are rarely attested in language use. Roland, Dick and Elman (2007) examined the frequency of syntactic structures in several corpora. They reported that cleft sentences accounted for less than 0.1% of all sentences, and object cleft sentences are 13 times less likely than subject cleft sentences. They also found that passive sentences primarily occurred in written rather than spoken language.

Roland et al. (2007) provided data corresponding to the frequency of some of the constructions included in the current study: 30% of verbs in the British National Corpus were used in the transitive construction; 11% in the intransitive; 17% before a prepositional phrase, such as the conative construction; 7% before an object and preposition, such as the caused motion and removal constructions; and 1% in the ditransitive. The set of constructions included in the current work could complement the set of constructions identified by generative linguistic theory as targets for intervention in aphasia therapy, providing fuller treatment coverage of the types of sentences clients are likely to use in everyday language.

9.3.2 Item selection for treatment generalisation

Another way in which results from the current study bear on clinical approaches to aphasia rehabilitation pertains to the items that are selected for treatment. Plaut (1996) explored whether the treatment of items that are more or less typical category members promotes generalisation to untreated items. He studied the behaviour of a connectionist model designed to represent single-word reading in dyslexia as the learned, arbitrary relationship between orthography and semantics. Typicality was defined as the closeness in semantic space between an item and a prototype, in terms of the number of semantic features they shared in the model. Results showed that generalisation to untreated items was greater when the treated set contained atypical items. Treatment of atypical items resulted in generalisation to typical items, but treatment of typical items did not result in generalisation to atypical items.

Plaut (1996) reasoned that this finding arose because atypical category members indicate both the central tendency of and variability allowed within a category. Similarities among atypical exemplars encode central dimensions of the category, whilst differences among atypical exemplars demonstrate the degree of acceptable variability within the category. A set of atypical items thus contains information that permits generalisation of recovery to both typical and atypical items.

Anderson, Herbert and Cowell (in press) argued that argument structure constructions are semantic categories, because they elicit frequency distributions that have been interpreted as denoting prototypical members of a category. Anderson et al. observed that the heavily skewed distributions of verbs generated in response to argument structure constructions in Phase 1 can be described as Zipfian (Zipf, 1935), and this pattern is a nontrivial aspect of language (Piantadosi, 2014). Researchers have interpreted exemplars that are most frequently generated in verbal fluency tasks to represent prototypical category members (Battig & Mottague, 1969; Chang, 1986). Thus Plaut's (1996) hypothesis about semantic categories can potentially apply to the treatment of verbs in argument structure constructions.

To illustrate, the three verbs typical adults most often generated in response to the conative construction in Phase 1 were *look*, *laugh* and *run*. Verbs that were generated only once in response to the conative construction were semantically related to these high-frequency verbs and included types of looking, like *frown*, *glance*, *glower*, *smirk* and *sneer*; types of laughing, like *jeer* and *titter*; and other motion verbs, like *come*, *fly*, *dance*, *skip* and *swim*. These low-frequency, or atypical, verbs are semantically related to the high-frequency, or prototypical, verbs. Furthermore, the semantic similarity among the range of verbs associated with the argument structure construction – in this case, all these verbs can occur with a goal – can be attributed to the syntactic form of the argument structure construction, in this example, the preposition phrase headed by *at*. In this way, Plaut's (1996) method for selecting items for treatment in the rehabilitation of sentence processing deficits in aphasia could capitalise on clients' residual semantic processing abilities.

Plaut's (1996) reasoning on item selection has been implemented in a treatment for lexical retrieval deficits in aphasia. Kiran (2007) described the success of a naming therapy based on treating typical or atypical exemplars of semantic categories, such as vegetables, birds and clothing. Six of the nine individuals with aphasia who completed the therapy showed generalisation of treatment gains from trained, atypical items to untrained, typical items. Where this idea has been applied to sentence-level interventions in aphasia, typicality has been defined with reference to the complexity of syntactic structures, like those included in the Treatment of Underlying Forms, including cleft sentences, the passive and wh-questions (Thompson, Shapiro, Kiran and Sobeck's (2003) Complexity Account of Treatment Efficacy). To date, no research has considered typicality within a set of verbs associated with a single argument structure as a basis for treatment in aphasia.

9.3.3 Theories of agrammatism

One participant in Phase 2, MJ, was diagnosed with non-fluent Broca's aphasia, and he showed signs of agrammatism in his expressive and receptive language. He also demonstrated sensitivity to the effect of construction frequency. He produced a greater number of target responses to verbs in the [high cx, high lex] condition than to verbs in the [low cx, low lex] condition in both tasks. In the grammaticality judgement task, within-participant analysis identified a moderate effect of construction frequency on response times, which was approaching significance. In the sentence completion task, his construction frequency effect score was significantly greater than typical participants in the analysis of the raw data.

Currently, no explanations of agrammatism can account for MJ's pattern of performance. Theories of agrammatism have been articulated with the aim of explaining why certain syntactic constructions prove more difficult than others for individuals with agrammatism, e.g. the comprehension of reversible passives (Grodzinsky's (1986) trace-deletion hypothesis) and the production of morphemes related to tense (Friedmann and Grodzinsky's (1997) tree pruning hypothesis). No available theory can account for differential performance within a single syntactic construction, as MJ demonstrated. Results from the current study suggest that agrammatic language processing may be sensitive not only to the complexity of structural differences between word- and sentence-level constructions, but also properties derived from language use, specifically, the closeness of the association between verbs and argument structures. Future research could pursue this hypothesis in a larger sample of agrammatic participants.

9.4 Conclusion

This research investigated the effect of past linguistic experience, as indexed by frequency of occurrence, on the processing of verbs and argument structures in adults with and without acquired aphasia. Results from typical participants confirm that the frequency of verbs in particular syntactic constructions affects language processing, and the expression of this effect is subject to task demands.

Results from adults with aphasia were generally consistent with findings from typical participants. Response patterns from individuals with aphasia varied, but results from most participants with aphasia were consistent with the response pattern of typical participants. Results from the current investigation may inform treatment approaches to sentence processing deficits in aphasia, item selection for those treatments and theories of agrammatism. This project demonstrated how investigations grounded in usage-based linguistic theory have the potential to broaden the scope of knowledge about language in aphasia, especially with reference to language processing at levels greater than the single word.

References

Abrams, D. A., Ryali, S., Chen, T., Balaban, E., Levitin, D. J., & Menon, V. (2013). Multivariate activation and connectivity patterns discriminate speech intelligibility in Wernicke's, Broca's, and Geschwind's areas. *Cerebral Cortex, 23*(7), 1703-1714.

Allen, K., Pereira, F., Botvinick, M., & Goldberg, A. E. (2012). Distinguishing grammatical constructions with fMRI pattern analysis. *Brain and Language, 123*, 174-182.

Anderson, E., Herbert, R., & Cowell, P. (accepted). The role of frequency in the association between verbs and argument structure constructions. Online proceedings of the UK-CLA Meetings, available from http://uk-cla.org/proceedings

Arnon, I., & Cohen Priva, U. (2013). More than words: The effect of multi-word frequency and constituency on phonetic duration. *Language and Speech, 56*(3), 349-371.

Arnon, I., & Cohen Priva, U. (2014). Time and again: The changing effect of word and multiword frequency on phonetic duration for highly frequent sequences. *The Mental Lexicon, 9*(3), 377-400.

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language, 62*(1), 67-82.

Arroyo-Anlló, E., Lorber, M., Rigaleau, F., & Gil, R. (2011). Verbal fluency in Alzheimer's disease and aphasia. *Dementia, 11*(1), 5-18.

Aston, G., & Burnard, L. (1998). *The BNC handbook: exploring the British National Corpus with SARA*. Edinburgh: Edinburgh University Press.

Baguley, T. (2012). Comparing correlations: independent or dependent (overlapping or non-overlapping). Retrieved from https://seriousstats.wordpress.com/2012/02/05/comparing-correlations/

Baldo, J., Schwartz, S., Wilkins, D., & Dronkers, N. (2010). Double dissociation of letter and category fluency following left frontal and temporal lobe lesions. *Aphasiology, 24*(12), 1593-1604.

Balota, D., Cortese, M., Sergent-Marshall, S., Spieler, D., & Yap, M. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133*(2), 283-316.

Barbieri, E., Basso, A., Frustaci, M., & Luzzatti, C. (2010). Arugment structure deficits in aphasia: new perspective on models of lexical production. *Aphasiology, 24*(11), 1400-1423.

Barnhart, H., Haber, M., & Lin, L. (2007). An overview on assessing agreement with continuous measures. *Journal of Biopharmaceutical Statistics, 17*(4), 529-569.

Bastiaanse, R., & Jonkers, R. (1998). Verb retrieval in action naming and spontaneous speech in agrammatic and anomic aphasia. *Aphasiology, 12*(11), 951-969.

Bastiaanse, R., & van Zonneveld, R. (2005). Sentence production with verbs of alternating transitivity in agrammatic Broca's aphasia. *Journal of Neurolinguistics, 18*, 57-66.

Bastiaanse, R., Edwards, S., & Rispens, J. (2002). *VAST: The verb and sentence test*. London: Pearson.

Bastiaanse, R., Wieling, M., & Wolthuis, N. (2016). The role of frequency in the retrieval of nouns and verbs in aphasia. *Aphasiology, 30*(11), 1221-1239.

Battig, W., & Montague, W. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut Category Norms. *Journal of Experimental Psychology, 80*(3), 1-46.

Baxter, D., & Warrington, E. (1985). Category specific phonolocail dysgraphia. *Neuropsychologia, 23*(5), 653-666.

Beattie, G., & Butterworth, B. (1979). Contextual probability and word frequency as determinants of pauses and errors in spontaneous speech. *Language and Speech, 22*(3).

Beckner, C., Blythe, R., Bybee, J., Christiansen, M., Croft, W., Ellis, N., . . . Schoenemann, T. (2009). Language is a complex adaptive system: position paper. In N. C. Ellis & D. Larsen-Freeman (Eds.), *Language as a complex adaptive system* (pp. 1-26). Chichester: Wiley-Blackwell.

Bencini, G., & Goldberg, A. (2000). The contribution of argument structure constructions to sentence meaning. *Journal of Memory and Language, 43*, 640-651.
Bergen, B., & Chang, N. (2005). Embodied Construction Grammar in a simulation-based language understanding. In J. Ostman & M. Fried (Eds.), *Construction Grammars: Cognitive grounding and theoretical extensions* (pp. 147-190). Amsterdam/Philadelphia: John Benjamins Publishing Company.
Berndt, R. (1998). Sentence processing in aphasia. In M. Sarno (Ed.), *Acquired Aphasia* (pp. 229-267). San Diego, California: Academic Press.
Berndt, R. S., Haendiges, A. N., Mitchum, C. C., & Sandson, J. (1997b). Verb retrieval in aphasia 2. Relationship to sentence processing. *Brain and Language, 56*, 107-137.
Berndt, R. S., Mitchum, C. C., Haendiges, A. N., & Sandson, J. (1997a). Verb retrieval in aphasia 1. Characterizing single word impairments. *Brain and Language, 56*, 68-106.
Berndt, R., Mitchum, C., & Haendiges, A. (1996). Comprehension of reversible sentences in "agrammatism": a meta-analysis. *Cognition, 58*, 289-308.
Bird, H., & Franklin, S. (1996). Cinderella revisited: A comparison of fluent and non-fluent aphasic speech. *Journal of Neurolinguistics, 9*(3), 187-206.
Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior and Research Methods, Instruments, & Computers, 33*(1), 73-79.
Black, M., & Chiat, S. (2003). Noun-verb dissociations: a multi-faceted phenomenon. *Journal of Neurolinguistics, 16*, 231-250.
Bland, J., & Altman, D. (1996). Transformations, means, and confidence intervals. *BMJ, 312*, 1079.
Bock, K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18*, 355-387.
Bock, K., & Loebell, H. (1990). Framing sentences. *Cognition, 35*(1), 1-39.
Bock, K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review, 99*(1), 150-171.
Boersma, P., & Weenink, D. (2017). Praat: doing phonetics by computer (Version 6.0.30). Retrieved from http://www.praat.org/
Booth, J., Burman, D., Meyer, J., Gitelman, D., Parrish, T., & Mesulam, M. (2002). Modality independence of word comprehension. *Human Brain Mapping, 16*, 251-261.
Bose, A., van Lieshout, P., & Square, P. A. (2007). Word frequency and bigram frequency effects on linguistic processing and speech motor performance in individuals with aphasia and normal speakers. *Journal of Neurolinguistics, 20*(1), 65-88.
Boyle, M., & Coelho, C. (1995). Application of semantic feature analysis as a treatment for aphasic dysnomia. *American Journal of Speech-Language Pathology, 4*, 135-138.
Brady, M., Kelly, H., Godwin, J., & Enderby, P. (2012). Speech and language therapy for aphasia following stroke *Cochrane Database of Systematic Reviews 5*(CD000425).
Breedin, S., Saffran, E., & Schwartz, M. (1998). Semantic factors in verb retrieval: An effect of complexity. *Brain and Language, 63*, 1-31.
Bresnan, J., Cueni, A., Nikitina, T., & Baayen, H. (2007). Predicting the dative alternation. In G. Bouma, I. Kramer & J. Zwarts (Eds.), *Cognitive Foundations of Interpretation* (pp. 69-94): Royal Netherlands Academy of Ats and Sciences.
Broca, P. (1861). Perte de la parole, ramollissement chronique de destruction partielle du lobe anterieur gauche du cerveau. *Bulletins de la Societe d'Anthropologie de Paris* (seance du 18 avril), 235-238.
Bronner, S., & Shippen, J. (2015). Biomechanical metrics of aesthetic perception in dance. *Experimental Brain Research, 233*, 3565-3581.
Brysbaert, M. (1996). Word frequency affects naming latency in Dutch when age of acquisition is controlled. *European Journal of Cognitive Psychology, 8*(2), 185-193.
Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: A review of recent developments and implications for the

choice of frequency estimates in German. *Experimental Psychology, 58*(5), 412-424.

Brysbaert, M., Lange, M., & Van Wijnendaele, I. (2000). The effects of age-of-acquisition and frequency-of-occurence in visual word recognition: Further evidence from the Dutch language. *European Journal of Cognitive Psychology, 12*(1), 65-85.

Bugg, J., Zook, N., DeLosh, E., Davalos, D., & Davis, H. (2006). Age differences in fluid intelligence: Contributions of general slowing and frontal decline. *Brain and Cognition, 62*, 9-16.

Bybee, J. (2006). From Usage to Grammar: The Mind's Response to Repetition. *Language, 82*(4), 711-733.

Bybee, J. (2007). *Frequency of use and the organization of language*. Oxford: Oxford University Press.

Bybee, J. (2010). *Language, usage and cognition*. Cambridge: Cambridge University Press.

Bybee, J., & Hopper, P. (2001). *Frequency and the emergence of linguistic structure*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Bybee, J., & Scheibman, J. (1999). The effect of usage on degrees of constituency: the reducation of *don't* in English. *Linguistics, 37*(4), 575-596.

Bybee, J., & Thompson, S. (1997). Three frequency effects in syntax. *Proceedings of the 23rd Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Pragmatics and Grammatical Structure* (pp. 378-388).

Byng, S., Kay, J., Edmundson, A., & Scott, C. (1990). Aphasia tests reconsidered. *Aphasiology, 4*(1), 67-91.

Caplan, D., & Futter, C. (1986). Assignment of thematic roles to nouns in sentence comprehension by an agrammatic patient. *Brain and Language, 27*, 117-134.

Caramazza, A., & Hillis, A. (1991). Lexical organization of nouns and verbs in the brain. *Nature, 349*(6312), 788-790.

Caramazza, A., Capitani, E., Rey, A., & Berndt, R. (2001). Agrammatic Broca's aphasia is not associated with a single pattern of comprehension performance. *Brain and Language, 76*, 158-184.

Carragher, M., Conroy, P., Sage, K., & Wilkinson, R. (2012). Can impairment-focused therapy change the everyday conversations of people with aphasia? A review of the literaure and future directions. *Aphasiology, 26*(7), 895-916.

Casenhiser, D., & Bencini, G. (2015). Argument structure constructions. In E. Dabrowska & D. Divjak (Eds.), *Handbook of Cognitive Linguistics* (pp. 579-593). Berlin/Boston: De Gruyter Mouton.

Chang, F., Bock, K., & Goldberg, A. (2003). Can thematic roles leave traces of their places? *Cognition, 90*, 29-49.

Chang, T. M. (1986). Semantic memory: facts and models. *Psychological Bulletin, 99*(2), 199-220.

Chen, Z., Mo, L., & Honomichl, R. (2004). Having the memory of an elephant: Long-term retrieval and the use of analogues in problem solving. *Journal of Experimental Psychology: General, 133*(3), 415-433.

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton

Chomsky, N. (1980). Rules and representations. *The Behavioral and Brain Sciences, 3*, 1-61.

Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht, Holland: Foris.

Chomsky, N. (1992). A minimalist program for linguistic theory. *MIT Occasional Papers in Linguistics* (Vol. 1). Cambridge, MA: MIT.

Chou, T., Booth, J., Burman, D., Bitan, T., Bigio, J., Lu, D., & Cone, N. (2006). Developmental changes in the neural correlates of semantic processing. *NeuroImage, 29*, 1141-1149.

Clark, A. (2016). *Surfing uncertainty: Prediction, action, and the embodied mind*. Oxford: Oxford University Press.

Clark, H. (1973). The language-as-fixed-effects fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12*, 335-359.

Code, C. (2013). Significant landmarks in the history of aphasia and its therapy. In I. Papathanasiou, P. Coppens & C. Potagas (Eds.), *Aphasia and Related Neurogenic Communication Disorders* (pp. 3-22). Burlington, MA: Jones & Bartlett Learning.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37-46.

Cohen, J. (1988). *Statistical power analyses for the social sciences*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159.

Colombo, L., & Burani, C. (2002). The influence of age of acquisition, root frequency, and context availability in processing nouns and verbs. *Brain and Language, 81*(1-3), 398-411.

Coltheart, M. (2005). Modeling reading: The dual-route approach. In M. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 6-23). Oxford: Blackwell Publishing.

Conklin, K., & Schmitt, N. (2012). The processing of formulaic language. *Annual Review of Applied Linguistics, 32*, 45-61.

Conroy, P., Sage, K., & Lambon Ralph, M. (2006). Towards theory-driven therapies for aphasic verb impairments: A review of current theory and practice. *Aphasiology, 20*(12), 1159-1185.

Conroy, P., Sage, K., & Lambon Ralph, M. (2009). Improved vocabulary producttion after naming therapy in aphasia: can gains in picture naming generalise to connected speech? *International Journal of Language and Communication Disorders, 44*(6), 1036-1062.

Cook, V. J., & Newson, M. (1996). *Chomsky's Universal Grammar: An Introduction* (2nd ed.). Oxford, UK: Blackwell Publishers.

Cortese, M., & Fugett, A. (2004). Imageability ratings for 3,000 monosyllabic words. *Behavior and Research Methods, Instruments, & Computers, 36*(3), 384-387.

Coutanche, M. N. (2013). Distinguishing multi-voxel patterns and mean activation: why, how, and what does it tell us? *Cognitive, Affective, & Behavioral Neuroscience, 13*, 667-673.

Crary, M., Wertz, R., & Deal, J. (1992). Classifying aphasias: Cluster analysis of Western Aphasia Battery and Boston Diagnostic Aphasia Examination results. *Aphasiology, 6*(1), 29-36.

Crawford, J., & Garthwaite, P. (2002). Investigation of the single case in neuropsychology: confidence limits on the abnormality of test scores and test score differences. *Neuropsychologia, 40*, 1196-1208.

Crawford, J., & Garthwaite, P. (2007). Comparison of a single case to a control or normative sample in neuropsychology: Development of a Bayesian approach. *Cognitive Neuropsychology, 24*(4), 343-372.

Crawford, J., & Howell, D. (1998). Comparing an individual's test score against norms derived from small samples. *The Clinical Neuropsychologist, 12*(4), 482-486.

Crawford, J., Garthwaite, P., & Porter, S. (2010). Point and interval estimates of effect sizes for the case-controls design in neuropsychology: Rationale, methods, implementations, and proposed reporting standards. *Cognitive Neuropsychology, 27*(3), 245-260.

Crawford, J., Garthwaite, P., Azzalini, A., Howell, D., & Laws, K. (2006). Testing for a deficit in single-case studies: Effects of departures from normality. *Neuropsychologia 44*, 666-677.

Croft, W. (2001). *Radical construction grammar: syntactic theory in typological perspective*. Oxford: Oxford University Press.

Cunningham, S. (2017). Automatic extraction of speech measurements from audio data. Department of Human Communication Sciences, University of Sheffield.

Dąbrowska, E. (2009). Words as constructions. In V. Evans & S. Pourcel (Eds.), *New Directions in Cognitive Linguistics*. Amsterdam: John Benjamins Publishing Company.

Dahan, D., Manguson, J., & Tanenhaus, M. (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology, 42*, 317-367.

Damasio, A. (1998). Signs of aphasia. In M. Sarno (Ed.), *Acquired Aphasia* (3rd ed., pp. 25-41). San Diego, California: Academic Press.

Damasio, A., & Geshwind, N. (1984). The neural basis of language. *Annual Review of Neuroscience, 7*(1), 127-147.

Davies, M. (2004-). BYU-BNC. (Based on the British National Corpus from Oxford University Press.), from http://corpus.byu.edu/bnc

Davis, G. A. (1993). *A survey of adult aphasia and related language disorders* (2nd ed.). Englewood Cliffs, N.J.: Prentice Hall.

Dax, M. (1836-1865). Lesions de la moitie gauche de l'encephale coincidant avec l'oubli des signes de la pensee - Lu au congres meridional tenu a Montpellier en 1836. *Gazette Hebdomadaire de Medecine et de Chirurgie 2*(2nd ser.), 259-260.

Deary, I., Liewald, D., & Nissan, J. (2011). A free, easy-to-use, computer-based simple and four-choice reaction time programme: The Deary-Liewald reaction time task. *Behavior Research Methods, 43*, 258-268.

DeDe, G. (2012). Effects of word frequency and modality on sentence comprehension impairments in people with aphasia. *American Journal of Speech-Language Pathology, 21*, S103-S114.

DeDe, G. (2013a). Verb transitivity bias affects on-line sentence reading in people with aphasia. *Aphasiology, 27*(3), 326-343.

DeDe, G. (2013b). Effects of verb bias and syntactic ambiguity on reading in people with aphasia. *Aphasiology, 27*(12), 1408-1425.

Dell, G. & Chang, F. (2014). The P-chain: relating sentence production and its disorders to comprehension and acqusition. *Philosophical Transactions of the Royal Society, 369*(20120394), 1-9.

Denes, G., & Dalla Barba, G. (1998). G.B. Vico, precursor of cognitive neuropsychology? The first reported case of noun-verb dissociation following brain damage. *Brain and Language, 62*(1), 29-33.

Diessel, H. (2015). Usage-based construction grammar. In E. Dabrowska & D. Divjak (Eds.), *Handbook of Cognitive Linguistics* (pp. 296-321). Berlin/Boston: De Gruyter Mouton.

DiLallo, J., Mettler, H., & DeDe, G. (2017). Corpus-based transitivity biases in individuals with aphasia. *Aphasiology, 31*(4), 447-464.

Divjak, D., & Caldwell-Harris, C. (2015). Frequency and entrenchment. In E. Dabrowska & D. Divjak (Eds.), *Handbook of Cognitive Linguistics* (pp. 53-74). Berlin/Boston: De Gruyter Mouton.

Dominey, P. (2002). A model of learning syntactic comprehension for natural and artificial grammars. In E. Witruk, A. Friederici & T. Lachmann (Eds.), *Basic functions of language, reading and reading disability* (pp. 61-77). Dordrecht, the Netherlands: Kluwer Academic Publishers.

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language, 67*(3), 547-619.

Druks, J. (2002). Verbs and nouns -- a review of the literature. *Journal of Neurolinguistics, 15*, 289-315.

Druks, J., & Marshall, J. (1995). When passives are easier than actives: two case studies of aphasic comprehension. *Cognition, 55*, 311-331.

Druks, J., & Masterson, J. (2000). *An Object and Action Naming Battery*. Hove: Psychology Press.

Edwards, S. (2005). *Fluent aphasia*. Cambridge: Cambridge University Press.

Ellis, N., & O'Donnell, M. (2012). Statistical construction learning: Does a Zipfian problem space ensure robust langauge learning? In J. Rebuschat & J. Williams (Eds.), *Statistical Learning and Language Acquisition*. Berlin: Mouton de Gruyter.
Ellis, N., O'Donnell, M., & Römer, U. (2014). The processing of verb-argument constructions is sensitive to form, function, frequency, contingency and prototypicality. *Cognitive Linguistics, 25*(1), 55-98.
Faroqi-Shah, Y., & Thompson, C. (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.
Federmeier, K., & Kutas, M. (2005). Aging in context: Age-related changes in context use during sentence comprehension. *Psychophysiology, 42*, 133-141.
Ferguson, A. (1998). Conversational turn-taking and repair in fluent aphasia. *Aphasiology, 12*(11), 1007-1031.
Ferreira, V., & Schotter, E. (2013). Do verb bias effects on sentence production reflect sensitivity to comprehension or production factors? . *The Quarterly Journal of Experimental Psychology, 66*(8), 1548-1571.
Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). Los Angeles, California: Sage.
Fonollosa, J., Neftci, E., & Rabinovich, M. (2015). Learning of chunking sequences in cognition and behavior. *PLOS Computational Biology*.
Forster, K., & Chambers, S. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior, 12*, 627-635.
Francis, W. N., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.
Frederici, A., & Graetz, P. (1987). Processing passive sentences in aphasia: deficits and strategies. *Brain and Language, 30*(1), 93-105.
Freed, E., Hamilton, S., & Long, D. (2017). Comprehension in proficient readers: The nature of individual variation. *Journal of Memory and Language, 97*, 135-153.
Friederici, A., Schönle, P., & Garrett, M. (1982). Syntactically and semantically based computations: Processing of prepositions in agrammatism. *Cortex, 18*(4), 525-534.
Friedmann, N., & Grodzinsky, Y. (1997). Tense and agreement in agrammatic production: Pruning the syntactic tree. *Brain and Language, 56*, 397-425.
Friedmann, N., & Gvion, A. (2007). As far as individuals with conduction aphasia understood these sentences were ungrammatical: Garden path in conduction aphasia. *Aphasiology, 21*(6/7/8), 570-586.
Friedrich, F., Martin, R., & Kemper, S. (1985). Consequences of a phonological coding deficit on sentence processing. *Cognitive Neuropsychology, 2*(4), 385-412.
Gahl, S. (2002). Lexical biases in aphasic sentence comprehension: An experimental and corpus linguistic study. *Aphasiology, 16*(12), 1173-1198.
Gahl, S. (2008). TIME and THYME are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language, 84*(3), 474-496.
Gahl, S., & Garnsey, S. (2004). Knowledge of grammar, knowledge of usage: Syntactic probabilities affect pronunciation variation. *Language, 80*(4), 748-775.
Gahl, S., & Menn, L. (2016). Usage-based approaches to aphasia. *Aphasiology, 30*(11), 1361-1377.
Gahl, S., Menn, L., Ramsberger, G., Jurafsky, D., Elder, E., Rewega, M., & Audrey, L. (2003). Syntactic frame and verb bias in aphasia: Plausibility judgments of undergoer-subject sentences. *Brain and Cognition, 53*, 223-228.
Gardner, M., Rothkopf, E., Lapan, R., & Lafferty, T. (1987). The word frequency effect in lexical decision: Finding a frequency-based component. *Memory and Cognition, 15*(1), 24-28.
Garnsey, S., Pearlmutter, N., Myers, E., & Lotocky, M. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language, 37*, 58-93.
Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155-170.

Gerratt, B., & Jones, D. (1987). Aphasic performance on a lexical decision task: Multiple meanings and word frequency. *Brain and Language, 30*, 160-115.

Gick, M., & Holyoak, K. (1980). Analogical problem solving. *Cognitive Psychology, 12*, 306-355.

Gilner, L., & Morales, F. (2005). TheSage English Dictionary and Thesaurus. Retrieved from www.sequencepublishing.com

Goldberg, A. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: The University of Chicago Press.

Goldberg, A. (2003). Constructions: a new theoretical approach to language. *Trends in Cognitive Sciences, 7*(5), 219-224.

Goldberg, A. (2006). *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.

Goldberg, A. (2013). Argument structure constructions versus lexical rules or derivational verb templates. *Mind and Language, 28*(4), 435-465.

Goldwater, M. B., & Markman, A. B. (2009). Constructional sources of implicit agents in sentence comprehension. *Cognitive Linguistics, 20*(4), 675-702.

Goodenough, C., Zurif, E., & Weintraub, S. (1977). Aphasics' attention to grammatical morphemes. *Language and Speech, 20*(1), 11-19.

Goodglass, H. (1993). *Understanding Aphasia*. San Diego, California: Academic Press, Inc.

Goodglass, H. (1997). Agrammatism in aphasiology. *Clinical Neuroscience, 4*, 51-56.

Goodglass, H., & Berko, J. (1960). Agrammatism and inflectional morphology in English. *Journal of Speech and Hearing Research, 3*(3), 257-267.

Goodglass, H., & Kaplan, E. (1972). *Boston Diagnostic Aphasia Examination*. Philadelphia: Lea & Febiger.

Goodglass, H., & Kaplan, E. (1983). *Boston diagnostic examination for aphasia*. Philadelphia: Lea & Febiger.

Goodglass, H., & Wingfield, A. (1997). Word-finding deficits in aphasia. In H. Goodglass & A. Wingfield (Eds.), *Anomia: Neuroanatomical and Cognitive Correlates*. San Diego, California: Academic Press.

Goodin, D., Aminoff, M., Ortiz, T., & Chequer, R. (1996). Response times and handedness in simple reaction-time tasks. *Experimental Brain Research, 109*, 117-126.

Goral, M., Levy, E., Swann-Sternberg, T., & Obler, L. (2010). Frequency and Word-Length Factors and Lexical Retrieval in Sentence Production in Aphasia. *Procedia - Social and Behavioral Sciences, 6*, 107-108.

Gordon, B., & Caramazza, A. (1982). Lexical decision for open- and closed-class words: Failure to replicate differential frequency sensitivity. *Brain and Language, 15*, 143-160.

Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory and Language, 29*, 228-244.

Graybiel, A. (1998). The basal ganglia and chunking of action repetoires. *Neurobiology of Learning and Memory, 70*(1-2), 119-136.

Gregory, E., Varley, R., & Herbert, R. (2012). Determiner primes as facilitators of lexical retrieval in English. *Journal of Psycholinguistic Research, 41*(6), 439-453.

Gregory, M., Raymond, W., Bell, A., Fosler-Lussier, E., & Jurafsky, D. (1999). The effects of collocational strength and contextual predictability in lexical production *Chicago Linguistic Society* (Vol. 35, pp. 151-166).

Griffin, Z., & Bock, K. (1998). Constraint, word frequency, and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language, 38*, 313-338.

Grodzinsky, Y. (1986). Language deficits and the theory of syntax. *Brain and Language, 27*, 135-159.

Gruenewald, P., & Lockhead, G. (1980). The free rexcall of category examples. *Journal of Experimental Psychology: Human Learning and Memory, 6*(3), 225-240.

Guastavino, C. (2007). Categorization of environmental sounds. *Canadian Journal of Experimental Psychology, 61*(1), 54-63.

Gvion, A., & Friedmann, N. (2012). Phonological short-term memory in conduction aphasia. *Aphasiology, 26*(3-4), 579-614.

Haarman, H., Just, M., & Carpenter, P. (1997). Aphasic sentence comprehension as a resrouce deficit: A computational approach *Brain and Language, 59*(1), 76-120.

Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In Proceedings of the second meeting of the north american chapter of the association for computational linguistics on language technologies (pp. 1–8). Stroudsburg, PA, USA: Association for Computational Linguistics.

Hallgren, K. (2012). Computing inter-rater reliability for observational data: an overview and tutorial. *Tutorials in quantitative methods for psychology, 8*(1), 23-34.

Hare, M., & Goldberg, A. (1999). Sturctural priming: purely syntactic? *Proceedings of the Cognitive Science Society Meeting*.

Hartsuiker, R., & Kolk, H. (1998). Syntactic facilitation in agrammatic sentence production. *Brain and Language, 62*(2), 221-254.

Haynes, J.-D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience, 7*, 523-534.

Henley, N. (1969). A psychological study of the semantics of animal terms. *Journal of Verbal Learning and Verbal Behavior, 8*, 176-184.

Hennecke, M., & Freund, A. (2010). Staying on and getting back on the wagon: Age-related improvement in self-regulation during a low-calorie diet. *Pscyhology and Aging, 25*(4), 876-885.

Herbert, R., Anderson, E., Best, W., & Gregory, E. (2014). Activation of syntax in lexical production in healthy speakers and in aphasia. *Cortex, 57*, 212-226.

Herbert, R., Haw, C., Brown, C., Gregory, E., & Brumfitt, S. (2012). Accessible Information Guidelines. London: Stroke Association.

Herbert, R., Webster, D., & Dyson, L. (2012). Effects of syntactic cueing therapy on picture naming and connected speech in acquired aphasia. *Neuropsychological Rehabilitation, 22*(4), 609-633.

Hoffman, P., Rogers, T., & Lambon Ralph, M. (2011). Semantic diversity accounts for the 'missing' word frequency effect in stroke aphasia: insights using a novel method to quantify contextual variability in meaning. *Journal of Cognitive Neuroscience, 23*(9), 2432-2446.

Howard, D., Patterson, K., Wise, R., Brown, W., Friston, K., Weiller, C., & Frackowiak, R. (1992). The cortical localization of the lexicons. *Brain, 115*, 1769-1782.

Huck, A., Thompson, R., Cruice, M., & Marshall, J. (2017). Effects of word frequency and contextual predictability on sentence reading in aphasia: an eye movement analysis. *Aphasiology*, 1-26.

Hund, A., & Plumert, J. (2005). The stability and flexibility of spatial categories. *Cognitive Psychology, 51*(1), 1-44.

Huttenlocher, J., & Kubicek, L. (1983). The source of relatedness effects on naming latency. *Journal of Experimental Psychology: Learning, Memory and Cognition, 9*(3), 486-496.

Inhoff, A., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception and Psychophysics, 40*(6), 431-439.

Jacobs, C., Dell, G., & Bannard, C. (2017). Phrase frequency effects in free recall: Evidence for redintegration. *Journal of Memory and Language, 97*, 1-16.

Jacobs, C., Dell, G., Benjamin, A., & Bannard, C. (2016). Part and whole linguistic experience affect recognition memory for multiword sequences. *Journal of Memory and Language, 87*, 38-58.

Jacobs, J., Hammerman-Rozenberg, R., Cohen, A., & Stessman, J. (2008). Reading daily predicts reduced mortality among men from a cohort of community-dwelling 70-year-olds. *Journal of Gerontology, 63*(2), S73-S80.

Jaeger, T. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology, 61*, 23-62.

Janata, P., & Grafton, S. (2003). Swinging in the brain: shared neural substrates for behaviors relating to sequencing and music. *Nature Neuroscience, 6*(7), 682-687.

Janssen, N., & Barber, H. A. (2012). Phrase frequency effects in language production. *PLoS One, 7*(3), e33202.

Jensen, L. (2000). Canonical structure without access to verbs? *Aphasiology, 14*(8), 827-850.

Jescheniak, J., & Levelt, W. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory and Cognition, 20*(4), 824-843.

Johnson, M., & Goldberg, A. (2013). Evidence for automatic accessing of constructional meaning: Jabberwocky sentences prime associated verbs. *Language and Cognitive Processes, 28*(10), 1439-1452.

Johnson, M., Turke-Browne, N., & Goldberg, A. (2013). Prediction plays a key role in language development as well as processing. *Behavioral and Brain Sciences, 36*(4), 360-361.

Johnson, M., Turke-Browne, N., & Goldberg, A. (2016). Neural systems involved in processing novel linguistic constructions and their visual referents. *Language Cognition and Neuroscience, 31*(1), 129-144.

Juhasz, B. (2005). Age-of-acquisition effects in word and picture identification. *Psychological Bulletin, 131*(5), 684-712.

Juhasz, B., & Rayner, K. (2003). Investigating the effects of a set of intercorrelated variables on eye fixation durations in reading. *Journal of Experimental Psychology: Learning, Memory and Cognition, 29*(6), 1312.

Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 229-254). Amsterdam/Philadelphia: John Benjamins

Just, M., & Carpenter, P. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*(4), 329-354.

Kahlaoui, K., Di Sante, G., Barbeau, J., Maheux, M., Lesage, F., Ska, B., & Joanette, Y. (2012). Contribution of NIRS to the study of prefrontal cortex for verbal flueny in aging. *Brain and Language, 121*, 164-173.

Kail, R., & Nippold, M. (1984). Unconstrained retrieval from semantic memory. *Child Development, 55*, 944-951.

Kako, E. (2006). Thematic role properties of subjects and objects. *Cognition, 101*, 1-42.

Kamide, Y., Altmann, G., & Haywood, S. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language, 49*(1), 133-156.

Kapatsinski, V., & Radicke, J. (2009). Frequency and the emergence of prefabs: Evidence from monitoring. In R. Corrigan, E. Moravcsik, H. Ouali & K. Wheatley (Eds.), *Formulaic language* (pp. 499-522). Amsterdam: John Benjamins.

Kaschak, M. P., & Glenberg, A. M. (2000). Constructing meaning: the role of affordances and grammatical constructions in sentence comprehension. *Journal of Memory and Language, 43*, 508-529.

Kay, J., & Ellis, A. W. (1987). A cognitive neuropsychological case study of anomia. *Brain, 110*, 613-629.

Kay, J., Lesser, R., & Coltheart, M. (1992). *PALPA: Psycholinguistic assessments of language performance in aphasia*: Lawrence Erlbaum.

Kay, P., & Fillmore, C. J. (1999). Grammatical constructions and linguistic generalizations: the *What's X doing Y?* construction. *Language, 75*(1), 1-33.

Kemmerer, D., & Tranel, D. (2000). Verb retrieval in brain-damaged subjects: 1. Analysis of stimulus, lexical and conceptual factors *Brain and Language, 73*, 347-392.

Kertesz, A. (1982). *The Western Aphasia Battery*. New York: Grune and Stratton.

Kern, S. (2013). Inferential statistics, power estimates, and study design formalities continue to surpress biomedical innovation. Retrieved from https://arxiv.org/abs/1411.0919v1. 1-18.

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods, 42*(3), 643-650.

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods, 42*(3), 643-650.

Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2012). The British Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods, 44*, 287-304.

Kielar, A., Meltzer-Asscher, A., & Thompson, C. (2012). Electrophysiological responses to argument structure violations in healthly adults and individuals with agrammatic aphasia. *Neuropsychologia, 50*, 3320-3337.

Kim, H., Kim, J., Kim, D. Y., & Heo, J. (2011). Differentiating between aphasic and nonaphasic stroke patients using semantic verbal fluency measures with administration time of 30 seconds. *European Neurology, 65*(2), 113-117.

Kim, M., & Thompson, C. (2000). Patterns of comprehension and production of nouns and verbs in agrammatism: implications for lexical organization. *Brain and Language, 74*(1), 1-25.

Kiran, S. (2007). Complexity in the treatment of naming deficits. *American Journal of Speech-Language Pathology, 16*, 18-29.

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.

Kohen, F., Milsark, G., & Martin, N. (2011). Effects of syntactic and semantic argument structure on sentence repetition in agrammatism: things we can learn from particles and prepositions. *Aphasiology, 25*(6-7), 736-747.

Komar, J., Chow, J., Chollet, D., & Seifert, L. (2014). Effect of analogy instructions with an internal focus on learning a complex motor skill. *Journal of Applied Sport Psychology, 26*(1), 17-32.

Kuperman, V., & Bresnan, J. (2012). The effects of construction probability on word durations during spontaneous incremental sentence production. *Journal of Memory and Language, 66*, 588-611.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods, 44*(4), 978-990.

Lachaud, C., & Renaud, O. (2011). A tutorial for analyzing human reaction times: How to filter data, manage missing values, and choose a statistical model. *Applied Psycholinguistics, 32*, 389-416.

Lachman, R. (1973). Uncertainty effects on time to access the internal lexicon. *Journal of Experimental Psychology, 99*(2), 199-208.

Lachman, R., Shaffer, J., & Hennrikus, D. (1974). Language and cognition: Effects of stimulus codability, name-word frequency, and age of acquistion on lexical reaction time. *Journal of Verbal Learning and Verbal Behavior, 13*, 613-625.

Lam, W., Maxwell, J., & Masters, R. (2009). Analogy versus explicit learning of a modified basketball shooting task: Performance and kinematic outcomes. *Journal of Sport Sciences, 27*(2), 179-191.

Landis, J., & Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174.

Langacker, R. (1988). A usage-based model. In B. Rudzka-Ostyn (Ed.), *Topics in Cognitive Linguistics* (Vol. 50, pp. 127-161). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Langacker, R. W. (1987). *Foundations of Cognitive Grammar* (Vol. 1). Stanford: Stanford University Press.

Lapata, M., Keller, F., & Schulte im Walde, S. (2001). Verb frame frequency as a predictor of verb bias. *Journal of Psycholinguistic Research, 30*(4), 419-435.

Lee, C., & Federmeier, K. (2012). Ambiguity's aftermath: How age differences in resolving lexical ambiguity affect subsequent comprehension. *Neuropsychologia, 50*(5), 869-879.

Leech, G., Rayson, P., & Wilson, A. (2001). *Word frequencies in written and spoken English: based on the British National Corpus*. Abingdon: Routedge

Leon, A., Davis, L., & Kraemer, H. (2011). The role and interpretation of pilot studies in clinical research. *Journal of Psychiatric Research, 45*(5), 626-629.

Levelt, W. (1983). Monitoring and self-repair in speech. *Cognition, 14*, 41-104.

Levelt, W., & Kelter, S. (1982). Surface from and memory in question answering. *Cognitive Psychology, 14*(1), 78-106.

Levitt, A., & Healy, A. (1985). The roles of phoneme frequency, similarity, and availability in the experimental elicitation of speech errors. *Journal of Memory and Language, 24*, 717-733.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition, 106*(3), 1126–1177.

Lowder, M., & Gordon, P. (2017). Print exposure modulates the effects of repetition priming during sentence reading. *Psychonomic Bulletin and Review*, 1-8.

Lubbers, M., & Torreira, F. (2016). Praatalign: an interactive Praat plug-in for performing forced alignment. Retrieved from https://github.com/dopefishh/praatalign, last accessed 22 July 2017.

Luzzatti, C., Raggi, R., Zonca, G., C., P., Contardi, A., & Pinna, G. (2002). Verb-noun double dissociation in aphasic lexical impairments: the role of word frequency and imageability. *Brain and Language, 81*, 432-444.

MacWhinney, B. (2015). Emergentism. In E. Dabrowska & D. Divjak (Eds.), *Handbook of Cognitive Linguistics* (pp. 689-706). Berlin/Boston: De Gruyter Mouton.

Manikandan, S. (2010). Data transformation. *Journal of Pharmacology and Pharmacotherapeutics, 1*(2), 126.

Markman, A. B., & Ross, B. (2003). Category use and category learning. *Psychological Bulletin, 129*(4), 592-613.

Marshall, J., Pring, T., Chiat, S., & Robson, J. (2001). When ottoman is easier than chair: An inverse frequency effect in jargon aphasia. *Cortex, 37*, 33-53.

Marsolais, Y., Methqal, I., & Joanette, Y. (2015). Marginal neurofunctional changes in high-performing older adults in a verbal fluency task. *Brain and Language, 140*, 13-23.

Mätzig, S., Druks, J., Masterson, J., & Vigliocco, G. (2009). Noun and verb differences in picture naming: past studies and new evidence. *Cortex, 45*(6), 738-758.

Mauner, G., & Koenig, J.-P. (2000). Linguistic vs. conceptual sources of implicit agents in sentence comprehension. *Journal of Memory and Language, 43*, 110-134.

McCarrey, A., An, Y., Kitner-Triolo, M., Ferrucci, L., & Resnick, S. (2016). Sex differences in cognitive trajectories in clinically normal adults. *Pscyhology and Aging, 31*(2), 166-175.

McCarthy, R., & Warrington, E. K. (1985). Category specificity in an agrammatic patient: the relative impairment of verb retrieval and comprehension. *Neuropsychologia, 23*, 709-727.

McDonald, S., & Shillcock, R. (2003). Eye movements reveal the on-line computation of lexical probabilities during reading. *Psychological Science, 14*(6), 648-652.

McRae, K., Hare, M., Elman, J., & Ferretti, T. (2005). A basis for generating expectancies for verbs from nouns. *Memory and Cognition, 33*(7), 1174-1184.

Menn, L., Gahl, S., Holland, A., Ramsberger, G., & Jurafsky, D. (2003). Beyond canonical form: Verb-frame frequency affects verb production and comprehension. *Brain and Language, 87*(1), 23-24.

Miceli, G., Silveri, M. C., Nocentini, U., & Caramazza, A. (1988). Patterns of dissociations in comprehension and production of nouns and verbs. *Aphasiology, 2*, 351-358.

Miceli, G., Silveri, M. C., Villa, G., & Caramazza, A. (1984). On the basis for the agrammatic's difficulty in producing main verbs. *Cortex, 20*, 207-240.
Michaelis, L. (2010). Sign-Based Construction Grammar. In B. Heine & H. Narrog (Eds.), *The Oxford Handbook of Linguistic Analysis* (pp. 139-158). Oxford: Oxford University Press.
Michaelis, L., & Lambrecht, K. (1996). Toward a construction-based theory of language function: the case of nominal extraposition. *Language, 72*(2), 215-247.
Miyake, A., Carpenter, P., & Just, M. (1994). A capacity approach to syntactic comprehension disorders: making normal adults perform like aphasic patients. *Cognitive Neuropsychology, 11*(6), 671-717.
Morrison, C., & Ellis, A. W. (1995). Roles of word frequency and age of acquistion in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*(1), 116-133.
Morrison, C., Ellis, A. W., & Quinlan, P. (1992). Age of acquisition, not word frequency, affects object naming, not object recognition. *Memory and Cognition, 20*(6), 705-714.
Morrison, C., Hirsh, K., & Duggan, G. (2003). Age of acquisition, ageing, and verb production: Normative and experimental data. *The Quarterly Journal of Experimental Psychology, 56*(4), 705-730.
Mumby, P. (2002). Statistical power of non-parametric tests: A quick guide for designing sampling strategies. *Marine Pollution Bulletin, 44*, 85-87.
Newcombe, F., Oldfield, R., & Wingfield, A. (1965). Object-naming by dysphasic patients. *Nature 207*, 1217-1218.
Nickels, L., & Howard, D. (1995). Aphasic naming: what matters? *Neuropsychologia, 33*(10), 1281-1303.
Nozari, N., Kittredge, A. K., Dell, G. S., & Schwartz, M. F. (2010). Naming and repetition in aphasia: Steps, routes, and frequency effects. *Journal of Memory Language, 63*(4), 541-559.
Nuyts, J. (1995). Autonomous vs. non-autonomous syntax In J. Verschueren, J. Ostman & J. Blommaert (Eds.), *Handbook of Pragmatics* (pp. 80-85). Amsterdam: John Benjamins Publishing Company.
Oldfield, R., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology, 17*(4), 273-281.
O'Malley, S., & Besner, D. (2012). Reading aloud and the question of intent. *Consciousness and Cognition, 21*, 1298-1310.
Pashek, G. (1998). Gestural facilitation of noun and verb retrieval in aphasia: A case study. *Brain and Language, 65*, 177-180.
Petruck, M. (1996). Frame semantics. In J. Verschueren & J. Ostman (Eds.), *Handbook of Pragmatics* (pp. 1-12). Amsterday: John Benjamins Publishing Company.
Pettit, L., Tönsing, K., & Dada, S. (2016). The persepectives of adults and their team members regarding the importance of nine life areas for rehabilitation: a pilot investigation. *Topics in Stroke Rehabilitation*.
Piantadosi, S. T. (2014). Zipf's word frequency law in language language: A critical review and future directions. *Psychonomic Bulletin and Review, 21*(5), 1112-1130.
Pickering, M., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *TRENDS in Cognitive Sciences, 11*(3), 105-110.
Pickering, M., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36*, 329-392.
Pickering, M., McLean, J., & Branigan, H. (2013). Persistent structural priming and frequency effects during comprehension. *Journal of Experimental Psychology: Learning, Memory and Cognition, 39*(3), 890-897.
Pierce, J. (2007). PsychoPy - Psychophysics software in Python. *Journal of Neuroscience Methods, 162*, 8-13.
Plaut, D. (1996). Relearning after damage in connectionist networks: Toward a theory of rehabilitation. *Brain and Language, 52*, 25-82.

Porta, M. (2014). *A Dictionary of Epidemiology* (6th ed.). New York: Oxford University Press.

Price, C., Wise, R., Ramsay, S., Friston, K., Howard, D., Patterson, K., & Frackowiak, R. (1992). Regional response differences within the human auditory cortex when listening to words. *Neuroscience Letters, 146*, 179-182.

Rahardja, D., Zhao, Y., & Qu, Y. (2009). Sample size determinations for the Wilcoxon-Mann-Whitney test: A comprehensive review. *Statistics in Biopharmaceutical Research, 1*(3), 317-322.

Ramscar, M., Hendrix, P., Love, B., & Baayen, H. (2013). Learning is not decline: The mental lexicon as a window into cognition across the lifespan. *Mental Lexicon, 8*(3), 450-481.

Ramscar, M., Hendrix, P., Shaoul, C., Milin, P., & Baayen, H. (2014). The myth of cognitive decline: Non-linear dynamics of lifelong learning. *Topics in Cognitive Science, 6*(1), 5-42.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin, 114*(3), 510-532.

Raymer, A., Foundas, A., Maher, L., Greenwald, M., Morris, M., Rothi, L., & Heilman, K. (1997). Cognitive neuropsychological analysis and neuroanatomic correlates in a case of acute anomia. *Brain and Language, 58*, 137-156.

Rayner, K., & Duffy, S. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory and Cognition, 14*(3), 191-201.

Reali, F., & Christiansen, M. H. (2007). Processing of relative clauses is made easier by frequency of occurrence. *Journal of Memory and Language, 57*, 1-23.

Reynolds, M., & Besner, D. (2006). Reading aloud is not automatic: Processing capacity is required to generate a phonological code from print. *Journal of Experimental Psychology: Human Perception and Performance, 32*(6), 1303-1323.

Risko, E., Stolz, J., & Besner, D. (2005). Basic processes in reading: Is visual word recognition obligatory? *Psychonomic Bulletin and Review, 12*(1), 119-124.

Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of Basic English Grammatical Structures: A Corpus Analysis. *Journal of Memory and Language, 57*(3), 348-379.

Rosch, E., & Mervis, C. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7*(4), 573-605.

Saffran, E., & Martin, N. (1997). Effects of structural priming on sentence production in aphasics. *Language and Cognitive Processes, 12*(5/6), 877-882.

Schiel, F. (1999). *Automatic phonetic transcription of non-prompted speech.* Proceedings of the ICPhS 1999, San Francisco, pp. 607-610.

Schwartz, M., Linebarger, M., Saffran, E., & Pate, D. (1987). Syntactic transparency and sentence interpretation in aphasia. *Language and Cognitive Processes, 2*(2), 85-113.

Sedley, D. (2013, Fall 2013). Plato's *Cratylus*. *The Stanford Encyclopedia of Philosophy* Retrieved 27 January 2014, 2014, from http://plato.stanford.edu/archives/fall2013/entries/plato-cratylus/

Seidenberg, M., & McClelland, J. (1989). A distributed, developmental model of word recognition and naming. *Pscyhological Review, 96*, 523-568.

Shapiro, L., & Levine, B. (1990). Verb processing during sentence comprehension in aphasia. *Brain and Language, 38*, 21-47.

Sherratt, S., Worrall, L., Pearson, C., Howe, T., Hersh, D., & Davidson, B. (2011). "Well it has to be language related". Speech-language pathologists' goals for people with aphasia and their families. *International Journal of Speech-Language Pathology, 13*(4), 317-328.

Shieh, G., Jan, S., & Randles, R. (2006). On power and sample size determinations for the Wilcoxon-Mann-Whitney test. *Journal of Nonparametric Statistics, 18*(1), 33-43.

Shieh, G., Lan, S., & Randles, R. (2007). Power and sample size determinations for the Wilcoxon signed-rank test. *Journal of Statistical Computation and Simulation, 77*(8), 717-724.

Siyanova-Chanturia, A., & Spina, S. (2015). Investigation of native speaker and second language learner intuition of collocation frequency. *Language Learning, 65*(3), 533-562.
Siyanova-Chanturia, A., Conklin, K., & van Heuven, W. (2011). Seeing a phrase "time and time" again matters: The role of phrasal frequency in the processing of multiword sequences. *Journal of Experimental Psychology: Learning, Memory and Cognition, 37*(3), 776-784.
Smiler, A., Gagne, D., & Stine-Morrow, E. (2003). Aging, memory load, and resource allocation during reading. *Psychology and Aging, 18*(2), 203-209.
Solopchuk, O., Alamia, A., Olivier, E., & Zenon, A. (2016). Chunking improves symbolic sequence processing and relies on working memory gating mechanisms. *Learning and Memory, 23*, 108-112.
Sonbul, S. (2015). Fatal mistake, awful mistake, or extreme mistake? Frequency effects on off-line/on-line collocational processing. *Bilingualism: Language and Cognition, 18*(3), 419-437.
Steels, L. (2012). *Design patterns in fluid construction grammar*. Amsterdam: John Benjamins Publishing Company.
Steinthal, H. (1871). *Einleitung in die Psychologie und Sprachwissenschaft*. Berlin, Germany: Dummler's Verlagsbuchhandlung.
Stine, E., & Hindman, J. (1994). Age differences in reading time allocation for propositionally dense sentences. *Aging, Neuropsychology, and Cognition, 1*(1), 2-16.
Stine-Morrow, E., Miller, L., & Hertzog, C. (2006). Aging and self-regulated language processing. *Pscyhological Bulletin, 132*(4), 582-606.
Stine-Morrow, E., Shake, M., Miles, J., & Noh, S. R. (2006). Adult age differences in the effects of goals on self-regulated sentence processing. *Psychology and Aging, 21*(4), 790-803.
Stites, M., & Federmeier, K. (2015). Subsequent to suppression: Downstream comprehension consequences of noun/verb ambiguity in natural reading. *Journal of Experimental Psychology: Learning, Memory and Cognition, 41*(5), 1497-1515.
Street, J., & Dąbrowska, E. (2014). Lexically specific knowledge and individual differences in adult native speakers' processing of the English passive. *Applied Psycholinguistics, 35*, 97-118.
Stroke Association (2017). State of the Nation: Stroke Statistics. 38pp.
Studtmann, P. (2013). Aristotle's *Categories*. *The Stanford Encyclopedia of Philosophy* Winter 2013. Retrieved 27 January 2014, 2014, from http://plato.stanford.edu/archives/win2013/entries/aristotle-categories/
Swinburn, K., Porter, G., & Howard, D. (2004). *Comprehensive Aphasia Test*. Hove: Psychology Press.
Szekely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., . . . Bates, E. (2005). Timed action and object naming. *Cortex, 41*, 7-25.
Tabachnick, B., & Fidell, L. (2007). *Using multivariate statistics* (5th ed.). Boston: Pearson.
Taylor, J. (2015). Prototype effects in grammar. In E. Dabrowska & D. Divjak (Eds.), *Handbook of Cognitive Linguistics* (pp. 562-578). Berlin/Boston: De Gruyter Mouton.
Thabane, L., Ma, J., Chu, R., Cheng, J., Ismaila, A., Rios, L., . . . Goldsmith, C. (2010). A tutorial on pilot studies: the what, why and how. *BMC Medical Research Methodology, 10*(1), 1-10.
Thompson, C. (2003). Unaccusative verb production in agrammatic aphasia: the argument structure complexity hypothesis. *Journal of Neurolinguistics, 16*(2-3), 151-167.
Thompson, C., & Shapiro, L. (2005). Treating agrammatic aphasia within a linguistic framework: Treatment of Underlying Forms. *Aphasiology, 19*(10-11), 1021-1036.
Thompson, C., Lange, K., Schneider, S., & Shapiro, L. (1997). Agrammatic and non-brain-damaged subjects' verb and verb argument structure production. *Aphasiology, 11*(4-5), 473-490.

Thompson, C., Shapiro, L., Kiran, S., & Sobecks, J. (2003). The role of syntactic complexity in treatment of sentence deficits in agrammatic aphasia: The complexity account of treatment efficacy (CATE). *Journal of Speech, Language, and Hearing Research, 46*(3), 591-607.

Tily, H., Gahl, S., Arnon, I., Snider, N., Kothari, A., & Bresnan, J. (2009). Syntactic probabilities affect pronunciation variation in spontaneous speech. *Language and Cognition, 1-2*, 147-165.

Tooley, K., & Traxler, M. (2010). Syntactic priming effects in comprehension: A critical review. *Language and Linguistics Compass, 4*(10), 925-937.

Tremblay, A., & Baayen, H. (2010). Holistic processing of regular four-word sequences: A behavioural and ERP study of the effects of structure, frequency, and probability on immediate free recall. In D. Wood (Ed.), *Perspectives on formulaic language: Acquisition and communication* (pp. 151-173). London: Continuum.

Tremblay, A., & Tucker, B. (2011). The effects of N-gram probabilistic measures on the recognition and production of four-word sequences. *The Mental Lexicon, 6*(2), 302-324.

Trueswell, J., & Kim, A. (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*, 102-123.

Trueswell, J., Tanenhaus, M., & Kello, C. (1993). Verb-specific constraints in sentence processing: separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory and Cognition, 19*, 528-553.

Tse, A., Fong, S., Wong, T., & Masters, R. (2017). Analogy of motor learning by young children: a study of rope skipping. *European Journal of Sport Science, 17*(2), 152-159.

Tse, A., Wong, A., Whitehall, T., Ma, E., & Masters, R. (2016). Examining the cognitive demands of analogy instructions compared to explicit instructions. *International Journal of Speech-Language Pathology, 18*(5), 465-472.

Tse, A., Wong, T., & Masters, R. (2017). Examining motor learning in older adults using analogy instructions. *Psychology of Sports and Exercise, 28*, 78-84.

UCREL. (1993). UCREL CLAWS7 Tagset, from http://ucrel.lancs.ac.uk/claws7tag.html

van der Auwera, J., Dobrushina, N., & Goussev, V. (2013). Imperative-Hortative Systems. In M. Dryer & M. Haspelmath (Eds.), The World Atlas of Language Structures Online. Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved from http://wals.info/chapter/72.

Varley, R., Whiteside, S., & Luff, H. (1999). Apraxia of speech as a disruption of word-level schemata: Some durational evidence. *Journal of Medical Speech-Language Pathology, 7*(2), 127-132.

Verhaeghen, P. (2003). Aging and vocabulary scores: A meta-analysis. *Psychology and Aging, 18*(2), 332-339.

Viera, A., & Garrett, J. (2005). Understanding interobserver agreement: the kappa statistic. *Family Medicine, 37*(5), 360-363.

Vigliocco, G., Vinson, D. P., Druks, J., Barber, H., & Cappa, S. F. (2011). Nouns and verbs in the brain: a review of behavioural, electrophysiological, neuropsychological and imaging studies. *Neuroscience and Biobehavioral Reviews, 35*(3), 407-426.

Vogel Sosa, A., & MacFarlane, J. (2002). Evidence for frequency-based constituents in the mental lexicon: collocations involving the word *of*. *Brain and Language, 83*, 227-236.

Wallace, S., Worrall, L., Rose, T., Le Dorze, G., Cruice, M., Isaksen, J., . . . Gauvreau, C. (2016). Which outcomes are most important to people with aphasia and their families? an international nominal group technique study framed within the ICF. *Disability and Rehabilitation*.

Wambaugh, J. L., Doyle, P. J., Martinez, A. L., & Kalinyak-Fliszar, M. (2002). Effects of two lexical retrieval cueing treatments on action naming in aphasia. *Journal of Rehabilitation Research and Development, 39*(4), 455-466.

Wambaugh, J. L., Linebaugh, C. W., Doyle, P. J., Martinez, A. L., Kalinyak-Fliszar, M., & Spencer, K. A. (2001). Effects of two cueing treatments on lexical retrieval in aphasic speakers with different levels of deficit. *Aphasiology, 15*(10/11), 933-950.

Wang, J., Baucom, L. B., & Shinkareva, S. V. (2013). Decoding abstract and concrete concept representations based on single-trial fMRI data. *Human Brain Mapping, 34*(5), 1133-1147.

Webster, J., Franklin, S., & Howard, D. (2004). Investigating the sub-processes involved in the production of thematic structure: An analysis of four people with aphasia. *Aphasiology, 18*(1), 47-68.

Wernicke, C. (1874). *Der Aphasische Symptomencomplex: Eine psychologische Studie auf anatomischer Basis*. Breslau, Poland: Cohn & Weigert.

Whitworth, A., Webster, J., & Howard, D. (2015). Argument structure deficit in aphasia: it's not all about verbs. *Aphasiology, 29*(12), 1426-1447.

Williams, S., & Canter, G. (1987). Action naming performance in four syndromes of aphasia. *Brain and Language, 32*, 124-136.

Wilson VanVoorhis, C., & Morgan, B. (2007). Understanding power and rules of thumb for determining sample sizes. *Tutorials in Quantitative Methods for Psychology, 3*(2), 43-50.

Wilson, M., & Garnsey, S. (2009). Making simple sentences hard: verb bias effects in simple direct object sentences. *Journal of Memory and Language, 60*, 368-392.

Wingfield, A. (1968). Effects of frequency on identification and naming of objects. *The Americal Journal of Psychology, 81*(2), 226-234.

Woods, D., Wyma, J., Yund, E. W., Herron, T., & Reed, B. (2015). Age-related slowing of response selection and production in a visual choice reaction time task. *Frontiers in Human Neuroscience, 9*(193), 1-12.

Woolams, A., Lambon Ralph, M., Madrid, G., & Patterson, K. (2016). Do you read how I read? Systematic individual differences in semantic reliance amongst normal readers. *Frontiers in Psychology, 7*, Article 1757.

Wright, C. (1979). Duration differences between rare and common words and their implications for the interpretation of word frequency effects. *Memory and Cognition, 7*(6), 411-419.

Yamaguchi, M., & Logan, G. (2016). Pushing typists back on the learning curve: Memory chunking in the hierarchical control of skilled typewriting. *Journal of Experimental Psychology: Learning, Memory and Cognition, 42*(12), 1919-1936.

Yap, M., Tse, C.-S., & Balota, D. (2009). Individual differences in the joint effects of semantic priming and word frequency revealed by RT distributional analyses: The role of lexical integrity. *Journal of Memory and Language, 61*(3), 303-325.

Young, J. S., Evermann, G., Gales, M. J. F., Hain, T., Kershaw, D., Moore, G., . . . Woodland, P. T. (2006). The HTK Book, version 3.4. Cambridge University Engineering Department, Cambridge, UK.

Zimmerer, V., Dąbrowska, E., Romanowski, C., Blank, C., & Varley, R. (2014). Preservation of passive constructions in a patient with primary progressive aphasia. *Cortex, 50*, 7-18.

Zingeser, L., & Berndt, R. (1990). Retrieval of nouns and verbs in agrammatism and anomia. *Brain and Language, 39*(1), 14-32.

Zipf, G. (1935). *The psychobiology of language*. Boston: Houghton Mifflin Co.

Zou, G. (2007). Toward using confidence intervals to compare correlations. *Psychological Methods, 12*(4), 399-413.

Appendix A

This appendix contains the ethics documents relating to Phase 1, including the project approval letter; information sheets for typical participants, participants with aphasia and the family and/or carers of participants with aphasia; and consent forms for typical participants and participants with aphasia.

Project approval

The University Of Sheffield.

Downloaded: 16/07/2017
Approved: 17/09/2014

Elizabeth Anderson
Registration number: 130116703
Human Communication Sciences
Programme: Full-time PhD

Dear Elizabeth

PROJECT TITLE: Relating verbs and sentence types
APPLICATION: Reference Number 000205

On behalf of the University ethics reviewers who reviewed your project, I am pleased to inform you that on 17/09/2014 the above-named project was **approved** on ethics grounds, on the basis that you will adhere to the following documentation that you submitted for ethics review:

- University research ethics application form 000205 (dated 01/08/2014).
- Participant information sheet 000631 version 1 (05/06/2014).
- Participant information sheet 000656 version 1 (05/06/2014).
- Participant information sheet 002330 version 1 (01/08/2014).
- Participant information sheet 002342 version 1 (01/08/2014).
- Participant information sheet 002343 version 1 (01/08/2014).
- Participant consent form 000633 version 1 (05/06/2014).
- Participant consent form 000638 version 1 (05/06/2014).
- Participant consent form 002331 version 1 (01/08/2014).
- Participant consent form 002344 version 1 (01/08/2014).

If during the course of the project you need to deviate significantly from the above-approved documentation please inform me since written approval will be required.

Yours sincerely

Thomas Muskett
Ethics Administrator
Human Communication Sciences

Information sheet for typical participants

The University Of Sheffield.

Department of Human Communication Sciences
362 Mushroom Lane
Sheffield S10 2TS UK

Head of Department
Professor Shelagh Brumfitt Senate Award Fellow
PhD, M.Phil, Dip CST, Cert MRCSLT (Hons)

Telephone: +44 (0) 114 222 2418/ 2402/ 2405
International: +44 (0) 114 222 2418
Fax: +44 (0) 114 222 2439
Email: hcs-support@lists.sheffield.ac.uk
http://www.shef.ac.uk/hcs

Information sheet for healthy participants

Relating verbs and sentence types

This Information Sheet describes a research project at the University of Sheffield, which is investigating the knowledge that speakers have of different types of sentences and words that occur in those types.

You are invited to take part in this research project. This Information Sheet gives you information about the study. You can decide whether you would like to take part after reading this sheet. You can ask the researcher any questions you have about your involvement.

What is this project about?
The researchers are interested in how speakers produce sentences. When people speak, they must choose the words they want to say and the order that the words should be said. The sentences *Sam sent Ashley a letter* and *Sam sent a letter to Ashley* describe the same event. Because they have words in different orders, they are different sentence types. Each sentence type can be used with many different verbs, or action words. The sentences *Sam sent Ashley a letter* and *Alex gave Chris a present* are the same type but have different verbs.

We are interested in the verbs that speakers use in different sentence types. Results from this project will contribute to further research on this relationship, which may result in a clinical tool to help adults with language impairment.

What would I be asked to do?
Participants will be asked to attend one session that will last no more than 45 minutes. They will see and hear a sentence with a missing word, such as *we like to* _____, and have thirty seconds to name as many words as possible that can fit in the sentence. There are sixteen sentences.

Afterwards, participants will be asked to take part in three simple language tasks that involve reading words, naming pictures and listening to sentences.

Who can take part?
Participants will:

- be between 50 and 80 years old;
- be native speakers of British English;
- not have any current or past speech and language difficulties; and
- not have any current or past psychiatric or mental health disorders.

Will I be paid?
No. Participation is voluntary. No rewards or compensation will be given for for taking part. No reimbursment for time or travel will be granted.

Can I stop at any time?
Yes. Participants can withdraw from participating at any time, without giving a reason. There will be no negative consequences for doing so.

Where will the study take place?
The study takes place at the Department of Human Communication Sciences at the University of Sheffield, 362 Mushroom Lane, S10 2TS.

What information will be collected?
Participants will be asked background information about their age, the languages they know, their education and profession. They will then provide vocal responses to four language tasks. The researcher will audio record the responses so they can be written down and checked later.

What happens to my data?
Electronic data will be kept securely in password protected storage locations. Paper materials will be kept in a locked filing cabinet in a secure office.

Results from this study will be used to create materials for another study on the same topic. Participants' responses will contribute to Elizabeth Anderson's PhD research. This research will be presented in conferences, published papers and a thesis, but participants will be anonymous in the presentation of results.

Will I remain anonymous? Will my participation be confidential?
Yes. All the information that is collected about participants will be kept confidential. Participants will not be identifiable in any reports, publications or presentations of the data.

What are the potential risks or disadvantages of taking part?
There are no known risks to taking part in this study.

What are the potential benefits of taking part?
There are no direct benefits to you. However, the data collected in this study will provide valuable information on how adults produce sentences and may contribute to the development of a clinical tool for adults with language impairment.

Has the project obtained ethical approval?
This project has received ethical approval from the Department of Human Communication Sciences at the University of Sheffield.

Who is funding the project?
The Department of Human Communication Sciences is funding this project.

Who is on the research team?
Elizabeth Anderson is a PhD student in the Department of Human Communication Sciences. Her project is supervised by Dr Ruth Herbert and Prof Patricia Cowell, who are senior academics in the department.

Researcher:	Elizabeth Anderson
	Department of Human Communication Sciences University of Sheffield 362 Mushroom Lane Sheffield S10 2TS
	ecanderson1@sheffield.ac.uk 0114 22 22 412
Supervisors:	Dr Ruth Herbert r.herbert@sheffield.ac.uk 0114 22 22 403
	Prof Patricia Cowell p.e.cowell@sheffield.ac.uk 0114 22 22 426

How can I get more information or sign up to take part?
Please contact Elizabeth Anderson for more information or to arrange a time to take part in the study.

What if there is a problem or I need to make a complaint?
You can speak with Elizabeth Anderson or her supervisors under any circumstances.

If you would like to speak to someone unrelated to the research team, you can contact the Ethics Lead for the Department of Human Communication Sciences:

Prof Ray Wilkinson
ray.wilkinson@sheffield.ac.uk
0114 22 22 449

If you are not satisfied with the responses from these individuals, you can contact the Registrar and Secretary of the University of Sheffield:

Dr Philip Harvey
registrar@sheffield.ac.uk

Thank you for considering taking part in this study!

Information sheet for participants with aphasia

The University Of Sheffield.

Department of Human Communication Sciences
362 Mushroom Lane
Sheffield S10 2TS UK

Head of Department
Professor Shelagh Brumfitt Senate Award Fellow
PhD, M.Phil, Dip CST, Cert MRCSLT (Hons)

Telephone: +44 (0) 114 222 2418/ 2402/ 2405
International: +44 (0) 114 222 2418
Fax: +44 (0) 114 222 2439
Email: hcs-support@lists.sheffield.ac.uk
http://www.shef.ac.uk/hcs

Information sheet for participants with aphasia

Relating verbs and sentence types

PhD student

Elizabeth Anderson  ☎ 0114 22 22 412

Supervisors

Ruth Herbert  ☎ 0114 22 22 403

Patricia Cowell  ☎ 0114 22 22 426

Address

Department of Human Communication Sciences
University of Sheffield
362 Mushroom Lane
Sheffield
S10 2TS

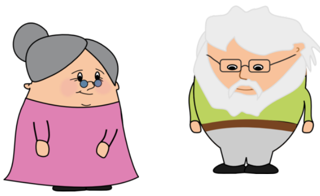
Elizabeth Anderson. Relating verbs and sentence types. August 2014.

The study

We are looking at **language** in **aphasia**

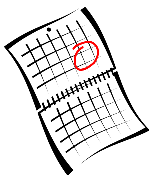
The study

We are looking for **volunteers** with **aphasia**



Taking part

Elizabeth will see you **once**

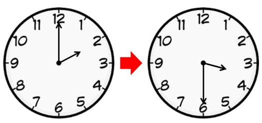


Taking part

Elizabeth can visit you at **home**

Taking part

The session will last for **90 minutes**



Taking part

You can come to the **clinic**

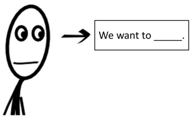
You choose

Elizabeth Anderson. Relating verbs and sentence types. August 2014.

Elizabeth Anderson. Relating verbs and sentence types. August 2014.

Taking part

You will **see** a **sentence**



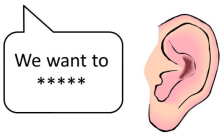
Taking part

The sentence will have a word **missing**

We want to ______.

Taking part

You will **hear** the **sentence**



Taking part

You will **say** words that can complete the sentence

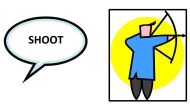
Elizabeth Anderson, Relating verbs and sentence types, August 2014

Taking part

You will then do three **activities**

You will **read** words

You will **name** pictures



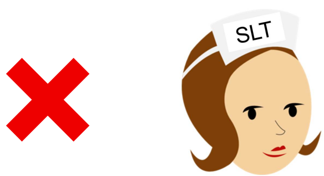
You will **hear** sentences

Taking part

You can still attend your **communication group**

Taking part

This is **NOT** speech and language **therapy**



Elizabeth Anderson, Relating verbs and sentence types, August 2014

Taking part

You can **rest** when you need

Taking part

You can **stop** at any time

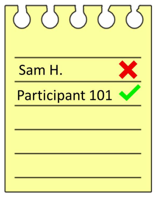
You do **NOT** need to give a **reason**

Data

You can allow us to **record** your responses

Data

We will **NOT** use your **name**



Data

We will store your information on a **secure computer**

Data

We will store your information in a **locked cabinet**

Research

We will write **reports** about this research

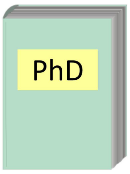
Research

We will present the research at **conferences**

Elizabeth Anderson, Relating verbs and sentence types, August 2014

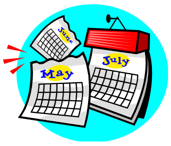
Research

Elizabeth will write about the research in her **thesis**



Research

Elizabeth will use the results in **a later study**



Elizabeth Anderson. Relating verbs and sentence types. August 2014.

Questions

You can ask Elizabeth **questions** about this project

Elizabeth Anderson

☎ 0114 22 22 412

ecanderson1@sheffield.ac.uk

You can contact her to **sign up**

Complaints

You can speak to the Ethics Lead to make a **complaint**

Ray Wilkinson

☎ 0114 22 22 449

ray.wilkinson @sheffield.ac.uk

Elizabeth Anderson. Relating verbs and sentence types. August 2014

Information sheet for family and/or carers of participants with aphasia

The University Of Sheffield.

Department of Human Communication Sciences
362 Mushroom Lane
Sheffield S10 2TS UK

Head of Department
Professor Shelagh Brumfitt Senate Award Fellow
PhD, M.Phil, Dip CST, Cert MRCSLT (Hons)

Telephone: +44 (0) 114 222 2418/ 2402/ 2405
International: +44 (0) 114 222 2418
Fax: +44 (0) 114 222 2439
Email: hcs-support@lists.sheffield.ac.uk
http://www.shef.ac.uk/hcs

Information sheet for family and/or carers of participants with aphasia

Relating verbs and sentence types

This Information Sheet describes a research project at the University of Sheffield, which is investigating the knowledge that speakers have of different types of sentence and words that occur in those sentence types.

Adults with aphasia are invited to take part in this research project. This Information Sheet gives you information about the study. You can ask the researcher any questions you may have.

What is this project about?
The researchers are interested in how speakers with and without aphasia produce sentences. When people speak, they must choose the words they want to say and the order that the words should be said. The sentences *Sam sent Ashley a letter* and *Sam sent a letter to Ashley* describe the same event. Because they have words in different orders, they are different sentence types. Each sentence type can be used with many different verbs, or action words. The sentences *Sam sent Ashley a letter* and *Alex gave Chris a present* are the same type but have different verbs.

We are interested in the verbs that speakers use in different sentence types. Results from this project will contribute to further research on this relationship, which may result in a clinical tool to help adults with language impairment.

What would participants be asked to do?
Participants will attend one session that will last no more than 90 minutes. Participants can meet the researcher at the Department of Human Communication Sciences at the University of Sheffield, or they can request Elizabeth Anderson to meet them in a quiet room in their own home.

In the study, participants will see and hear a sentence with a missing word and have thirty seconds to name words that can fit in the sentence. There are sixteen sentences. Participants will then be asked to take part in three short language tasks that involve reading words, naming pictures and listening to sentences.

This is **NOT** speech and language therapy. No counselling or therapy will be provided at the session.

Who can take part?
Participants will:

- be native speakers of British English;
- be between 50 and 80 years old;
- have been told by a speech and language therapist that they have aphasia;
- have been living with aphasia for at least six months;
- have no other speech or language difficulty, in addition to aphasia;
- have no history of psychiatric or mental health disorders;
- have had only one brain injury resulting in aphasia. For example, participants cannot have had two strokes.

Will participants receive payment?
No. Participation is voluntary. No rewards or compensation will be given for taking part. No reimbursement for time or travel will be provided.

Can participants stop at any time?
Yes. Participants can withdraw at any time, without giving a reason. There will be no negative consequences for doing so.

Will participation in this study impact the support that participants recieve?
No. Participation in this study will not affect any support that participants receive, such as their attendance at communication groups or speech and language therapy.

What information will be collected?
Participants will be asked background information about their age, language background, education, profession and language impairment. They will provide vocal responses in the four tasks of the study. Participants can agree to have their responses audio recorded and transcribed later.

What happens to the data?
Electronic data will be kept securely in password protected storage locations. Paper materials will be

The results from
topic. The results
project will be pr

Will participants remain anonymous? Will their participation be confidential?
Yes. All the information that is collected about participants will be kept confidential. They will not be identifiable in any reports, publications or presentations of the data.

What are the potential risks or disadvantages of taking part?
Participants may become mentally fatigued during the study. They can rest as much as they need. They can also stop at any time, without giving a reason.

What are the potential benefits of taking part?
There are no direct benefits to participants. However, the data collected in this study will provide valuable information on the language abilities of adults with aphasia. The results may contribute to the development of a clinical tool for adults with language impairment.

Has the project obtained ethical approval?
This project has received ethical approval from the Department of Human Communication Sciences at the University of Sheffield.

Who is funding the project?
The Department of Human Communication Sciences is funding this project.

Who is on the research team?
Elizabeth Anderson is a PhD student in the Department of Human Communication Sciences. Her project is supervised by Dr Ruth Herbert and Prof Patricia Cowell, who are senior academics in the department.

Researcher: Elizabeth Anderson

Department of Human Communication Sciences
University of Sheffield
362 Mushroom Lane
Sheffield
S10 2TS

ecanderson1@sheffield.ac.uk
0114 22 22 412

Supervisors: Dr Ruth Herbert
r.herbert@sheffield.ac.uk
0114 22 22 403

Prof Patricia Cowell
p.e.cowell@sheffield.ac.uk
0114 22 22 426

How can I get more information or sign up?
Please contact Elizabeth Anderson for more information or to arrange a time to take part in the study.

What if there is a problem or I need to make a complaint?
You can speak with Elizabeth Anderson or her supervisors under any circumstances.

If you would like to speak to someone unrelated to the research team, you can contact the Ethics Lead for the Department of Human Communication Sciences:

Prof Ray Wilkinson
ray.wilkinson@sheffield.ac.uk
0114 22 22 449

If you are not satisfied with the responses from these individuals, you can contact the Registrar and Secretary of the University of Sheffield:

Dr Philip Harvey
registrar@sheffield.ac.uk

Thank you for your interest in this study!

Consent form for typical participants

The University Of Sheffield.

Department of Human Communication Sciences
362 Mushroom Lane
Sheffield S10 2TS UK

Head of Department
Professor Shelagh Brumfitt Senate Award Fellow
PhD, M.Phil, Dip CST, Cert MRCSLT (Hons)

Telephone: +44 (0) 114 222 2418/ 2402/ 2405
International: +44 (0) 114 222 2418
Fax: +44 (0) 114 222 2439
Email: hcs-support@lists.sheffield.ac.uk
http://www.shef.ac.uk/hcs

Consent form for healthy participants

Relating verbs and sentence types

Researcher: Elizabeth Anderson

Participant number: _____ **Please initial.**

1. I have read and understood the information sheet. ☐
2. My questions have been answered. ☐
3. I understand that my participation is voluntary. I am free to withdraw at any time, without giving a reason. There are no negative consequences for doing this. ☐
4. I understand that my responses will be confidential. I will not be identifiable in any report, publication or presentation of the data from this project. ☐
5. I agree that my responses can be audio recorded. I give permission for members of the research team to access the recordings. ☐
6. I understand that the results of this study will be used in future research. ☐
7. I agree to take part in the above research. ☐

____
Name of participant Date Signature

____
Name of researcher Date Signature

Consent form for participants with aphasia

The University Of Sheffield.

Department of Human Communication Sciences
362 Mushroom Lane
Sheffield S10 2TS UK

Head of Department
Professor Shelagh Brumfitt Senate Award Fellow
PhD, M.Phil, Dip CST, Cert MRCSLT (Hons)

Telephone: +44 (0) 114 222 2418/ 2402/ 2405
International: +44 (0) 114 222 2418
Fax: +44 (0) 114 222 2439
Email: hcs-support@lists.sheffield.ac.uk
http://www.shef.ac.uk/hcs

Consent form for participants with aphasia

Relating verbs and sentence types

Researcher: Elizabeth Anderson

Participant number: _____

Information sheet **Please tick ✓**

I understand the **information sheet** ☐

I had a chance to ask **questions** ☐

I am **volunteering** to take part ☐

Doing this will **NOT** affect any **support** I receive ☐

I can **stop** at any time ☐

The results will be used to make **another study** ☐

Elizabeth Anderson, Relating verbs and sentence types, August 2014

Audio recordings

I agree that my voice can be **recorded** ☐

The research team will **listen** to the recordings ☐

Confidentiality

The study will **NOT** use my **name** ☐

Telling people about results

I understand the researchers will **talk** about the results to **other researchers** ☐

I understand my **name** will **NOT** be used in presentations ☐

I **agree** to take part in this research project

____
Name of participant Date Signature

____
Name of researcher Date Signature

Elizabeth Anderson, Relating verbs and sentence types, August 2014

Appendix B

This appendix contains the items included in the language assessments in Phase 1.

Verbal fluency task

Example items

I like to ___	fish, garden, hike
They might ___	slip, apologise, scream
We could ___	rest, escape, dream
You have to ___	beg, shout, apply
It should ___	emerge, expand, ascend

Practice items

They want to ___

We can ___

I try to ___

You need to ___

It must ___

Test items

It was ___ by them	You ___ at us
I ___ you something	We ___ our way there
I ___ it from you	They ___ us some things
They ___ at it	We were ___ by them
They ___ their way to it	You ___ it from there
It ___ through there	We ___ them
You ___ it	You ___ it to us
I ___ it over there	We ___ through there

Action-naming task

Practice items

- building
- opening
- sitting
- carrying
- reading

Test items

bleeding	tying
licking	crawling
dripping	pulling
touching	sneezing
waving	walking
dropping	combing
writing	pointing
sinking	cutting
cooking	folding
playing	running

Function word processing task

Practice items

she

rhese

thip

fome

will

Test items

ag	I
o	shey
through	ik
wak	ut
ghere	from
they	bu
there	phrough
it	us
te	hrom
way	you
thek	them
by	we
at	mou
de	to

Grammaticality judgement task

Practice items

They might lose

It start should

You have to share

I smile to like

We wait must

Test items

You drove through there

You were called by us

She sent them something

She caught us it from

It you pushed

You bought some us things

She pointed at them

They passed to me it

You made your way there

They watched me

She met by was you

We looked it at

She walked there way her

We took it from them

We sold it to him

It dropped there through

Appendix C

This appendix details the method for extracting values of construction frequency for verbs in Phase 1. Table A1 shows the search queries that were used to derive values of construction frequency from the British National Corpus (Davies, 2004-).

Table A1 Search queries to derive values of construction frequency from British National Corpus

<table>
<tr><th>Construction</th><th>Corpus query</th><th>Gloss</th></tr>
<tr><td>Caused motion</td><td>[vv*] [pp*] to|over -[v*]</td><td>lexical verb - pronoun - *to* or *over* - not a verb</td></tr>
<tr><td>Conative</td><td>[vv*] at [pp*]</td><td>lexical verb - *at* - pronoun</td></tr>
<tr><td>Ditransitive</td><td>[vv*] [pp*] [d*] [nn*]</td><td>lexical verb - pronoun - determiner - noun</td></tr>
<tr><td>Intransitive motion</td><td>[vv*] through</td><td>lexical verb - *through*</td></tr>
<tr><td>Passive</td><td>[vb*] [vvn] by</td><td>form of *be* - past participle of lexical verb - *by*</td></tr>
<tr><td>Removal</td><td>[vv*] [pp*] from</td><td>lexical verb - pronoun - *from*</td></tr>
<tr><td>Transitive</td><td>[vv*] [pp*] .|,|;</td><td>lexical verb - pronoun - clause-final punctuation</td></tr>
<tr><td>Way</td><td>[vv*] [appge] way [i*]</td><td>lexical verb - possessive pronoun - *way* - preposition</td></tr>
</table>

Note. Alternating rows shown in grey for ease of reference.

For each verb in the dataset, the verb lemma was specified in the query for the construction to which the verb was produced. For example, the verb *leave* was produced in response to the caused motion construction. In order to ascertain the construction frequency of the verb *leave* in the caused motion construction, Brigham Young University's interface to the British National Corpus (Davies, 2004-) was searched with the query shown below.

[leave].[vv*] [pp*] to|over -[v*]

To summarise, this search returned text strings in the corpus that contained any form of the verb *leave* tagged as a lexical verb, followed by a pronoun, followed by the word *to* or *over* and finally by any word that was not tagged as a verb. Results to each query were reviewed to ensure that they were genuine instances of the caused motion construction. This process was repeated for all 105 verbs that participants produced in response to the caused motion construction.

A similar process was undertaken to derive construction frequency values for all verbs produced in response to the remaining seven argument structure constructions.

Appendix D

Tables A2 and A3 contains a breakdown of responses from participants with aphasia, including DS, EF, LM and TP, in terms of the coding schemes described in Chapter 3 for the verbal fluency task and action-naming task in Phase 1.

Responses from participants in Phase 1 verbal fluency task

Table A2 Participant responses in Phase 1 verbal fluency task

	Typical participants				DS	EF	LM	TP
	Mean	SD	Min	Max				
Total number of responses	112	27	65	158	38	72	21	30
Verb types	65	16	35	93	21	11	12	11
Real verb	0.94	0.04	0.86	1.00	0.63	0.40	0.62	0.94
Phonological error	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00
Non-verb	0.01	0.02	0.00	0.09	0.11	0.04	0.00	0.00
Multiple words	0.00	0.01	0.00	0.02	0.16	0.14	0.19	0.00
Repetition	0.04	0.03	0.00	0.14	0.05	0.38	0.14	0.07
Rejection	0.00	0.01	0.00	0.02	0.00	0.04	0.05	0.00
Non-word	0.00	0.00	0.00	0.01	0.03	0.00	0.00	0.00

Note. Total number of responses produced by typical participants and individuals with aphasia in entire task; number of verb types produced across all responses; proportion of responses of each response type for group of typical participants and individuals with aphasia.

Responses from participants with aphasia in Phase 1 action-naming task

Table A3 Responses from participants with aphasia in Phase 1 action-naming task

	DS	EF	LM	TP
Target responses	0.95	**0.90**	**0.60**	**0.80**
Semantic	1	0	0	2
Phonological	0	0	4	0
Unrelated	0	0	0	1
Non-word	0	0	0	0
Other	0	2	4	1
No response	0	0	0	0

Note. Proportion of target responses and number of each error type from participants with aphasia. Scores outside normal range in bold. N = 20.

Appendix E

This appendix contains documents relating to Phase 2, including the project approval letter; information sheets for typical participants, participants with aphasia and the family and/or carers of participants with aphasia; and consent forms for typical participants and participants with aphasia.

Project approval

The University Of Sheffield.

Downloaded: 16/07/2017
Approved: 01/12/2015

Elizabeth Anderson
Registration number: 130116703
Human Communication Sciences
Programme: Full-time PhD

Dear Elizabeth

PROJECT TITLE: Language after stroke
APPLICATION: Reference Number 006821

On behalf of the University ethics reviewers who reviewed your project, I am pleased to inform you that on 01/12/2015 the above-named project was **approved** on ethics grounds, on the basis that you will adhere to the following documentation that you submitted for ethics review:

- University research ethics application form 006821 (dated 05/11/2015).
- Participant information sheet 1013317 version 1 (05/11/2015).
- Participant information sheet 1013316 version 1 (05/11/2015).
- Participant information sheet 1013315 version 1 (05/11/2015).
- Participant information sheet 1013314 version 1 (05/11/2015).
- Participant information sheet 1013313 version 1 (05/11/2015).
- Participant consent form 1013319 version 1 (05/11/2015).
- Participant consent form 1013318 version 1 (05/11/2015).

The following optional amendments were suggested:

none

If during the course of the project you need to deviate significantly from the above-approved documentation please inform me since written approval will be required.

Yours sincerely

John Mason
Ethics Administrator
Human Communication Sciences

Department Of Human Communication Sciences

Head of Department
Professor Patricia E. Cowell, BA, MS, PhD

362 Mushroom Lane
Sheffield
S10 2TS
United Kingdom

Telephone: +44 (0) 114 222 2418/ 2402/ 2405
International: +44 (0) 114 222 2418
Fax: +44 (0) 114 2222439
Email: hcs-support@sheffield.ac.uk
http://www.shef.ac.uk/hcs

Information sheet for participants without aphasia

Language after stroke

This Information Sheet describes a research project at the University of Sheffield, which is investigating language after stroke.

You are invited to take part in this research project. This Information Sheet gives you information about the study. You can decide whether you would like to take part after reading this sheet. You can ask the researcher any questions you have about your involvement.

What is this project about?
The researchers are interested in how language is affected by stroke. In order to study this, we need to collect information both from people who have had a stroke, and also people who have not had a stroke.

We are interested in how speakers use words in sentences. Information from this project will contribute to further research on this topic, and it may inform clinical approaches to language after stroke.

What would I be asked to do?
You would attend one session that will last no more than 60 minutes. You would be asked some information about yourself, such as your age, gender and language background.

You would then be asked to complete two different tasks on the computer. In one task, you would be asked to replace a blank space in a sentence with a word. For example, you might see the sentence *you can* _____ and the word *dream*. You would then say the entire sentence, *you can dream*. In the other task, you would be asked to decide whether the word can be used in the sentence. In this example, you would press a button to indicate that *dream* can be used in the sentence.

After that, you would be asked to complete one task off the computer. This may involve deciding whether a list of words are real words in English, constructing complete sentences from phrases of words, naming pictures or producing sentences.

Who can take part?
You must:
- be between 50 and 80 years old;
- be a native speaker of British English;
- not have any current or past problems with your speech or language; and
- not have any current or past psychiatric or mental health disorder.

Will I be paid?
No. Your participation is voluntary. You will not receive any rewards or compensation for taking part. You will not be reimbursed for your time or travel.

Can I stop at any time?
Yes. You can withdraw from participating at any time, without giving a reason. There will be no negative consequences for doing so.

Where will the study take place?
You can meet the researcher at the Department of Human Communication Sciences at the University of Sheffield, 362 Mushroom Lane, S10 2TS, or in a quiet room of your own home.

What information would be collected?
You would be asked background information about your age, language, education and profession. You would then provide responses to the two language tasks on the computer and the one task off the computer. Your vocal responses would be recorded.

What happens to my data?
Electronic data will be kept securely in password protected storage locations. Paper materials will be kept in a locked filing cabinet in a secure office.

Information from this study will contribute to Elizabeth Anderson's PhD research.

Will I remain anonymous? Will my participation be confidential?
Yes. All the information that is collected about you will be kept confidential. You will not be identifiable in any reports, publications or presentations of the data.

What are the potential risks or disadvantages of taking part?
There are no known risks to taking part in this study. If you become mentally fatigued during the experiment, you can take a break.

What are the potential benefits of taking part?
There are no direct benefits to you. However, results from this study will provide valuable information on how adults use language and may contribute to the development of clinical resources for adults with language difficulties after stroke.

Has the project obtained ethical approval?
This project has received ethical approval from the University Research Ethics Committee at the University of Sheffield.

Who is funding the project?
The Faculty of Medicine, Dentistry and Health at the University of Sheffield is funding this project.

Who is on the research team?
Elizabeth Anderson is a PhD student in the Department of Human Communication Sciences. Her project is supervised by Dr Ruth Herbert and Prof Patricia Cowell.

Researcher: Elizabeth Anderson

Department of Human Communication Sciences
University of Sheffield
362 Mushroom Lane
Sheffield
S10 2TS

ecanderson1@sheffield.ac.uk
0114 22 22 412

Supervisors: Dr Ruth Herbert
r.herbert@sheffield.ac.uk
0114 22 22 403

Prof Patricia Cowell
p.e.cowell@sheffield.ac.uk
0114 22 22 426

How can I get more information or sign up to take part?
Please contact Elizabeth Anderson via the email address or telephone number listed above for more information, or to arrange a time to take part in the study.

What if there is a problem or I need to make a complaint?
You can speak with Elizabeth Anderson or her supervisors under any circumstances.

If you would like to speak to someone unrelated to the research team, you can contact the Ethics Administrator of the Department of Human Communication Sciences:

Dr Traci Walker
traci.walker@sheffield.ac.uk
0114 22 22 420

If you are not satisfied with the responses from these individuals, you can contact the Office of the Registrar and Secretary at the University of Sheffield at registrar@sheffield.ac.uk.

Thank you for considering taking part in this study!

Department Of Human Communication Sciences

Head of Department
Professor Patricia E. Cowell, BA, MS, PhD

362 Mushroom Lane
Sheffield
S10 2TS
United Kingdom

Telephone: +44 (0) 114 222 2418/ 2402/ 2405
International: +44 (0) 114 222 2418
Fax: +44 (0) 114 2222439
Email: hcs-support@sheffield.ac.uk
http://www.shef.ac.uk/hcs

Information sheet for participants with aphasia

Language after stroke

PhD student

Elizabeth Anderson
☎ 0114 22 22 412

Supervisors

Ruth Herbert
☎ 0114 22 22 403

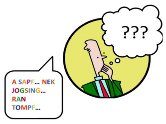
Patricia Cowell
☎ 0114 22 22 426

Address

Department of Human Communication Sciences
University of Sheffield
362 Mushroom Lane
Sheffield S10 2TS

The study

We are looking at **language** in **aphasia**



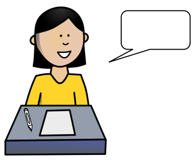
The study

We are looking for **volunteers** with **aphasia**

Elizabeth Anderson. Language after stroke. January 2016.

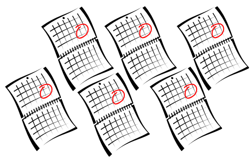
Taking part

You would be asked to do **three activities**



Taking part

Elizabeth would see you **four to six times**



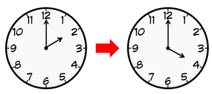
Taking part

Some people would **start** the study

Some people would **not start** the study

Taking part

Each session would last for **two hours**



Elizabeth Anderson. Language after stroke. November 2015.

Taking part

Elizabeth can visit you at **home**

Taking part

You would **talk** to the **researcher**

Taking part

You can come to the **clinic**

You choose

Taking part

You would use a **computer**

Elizabeth Anderson. Language after stroke. November 2015.

Elizabeth Anderson. Language after stroke. January 2016.

Taking part

You would **read** words and sentences

Taking part

You can still attend your **communication group**

Taking part

You would **listen** to speech

Taking part

This is **NOT** speech and language **therapy**

Elizabeth Anderson. Language after stroke. November 2015.

Elizabeth Anderson. Language after stroke. November 2015.

Taking part

You can **rest** when you need

Taking part

You can **stop** at any time

You do **NOT** need to give a **reason**

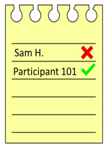
Elizabeth Anderson. Language after stroke. November 2015.

Data

You can let us **record** your responses

Data

We will **NOT** use your **name**



Elizabeth Anderson. Language after stroke. November 2015.

Data

We will keep your information on a **secure computer**

Data

We will keep your information in a **locked cabinet**

Elizabeth Anderson. Language after stroke. November 2015.

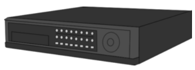
Data

We may put information on a secure site **online**

We will **NOT** put your **name** or **voice** online

Data

We may **keep** your **audio recordings** for future



You do **NOT** have to agree

Elizabeth Anderson. Language after stroke. January 2016.

Research

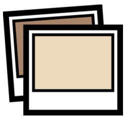
We will talk about this project at **conferences**

Research

We will write **reports** about this project

Research

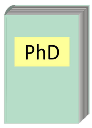
We can use your **photograph**



You do **NOT** have to agree

Research

Elizabeth will write about the project in her **thesis**

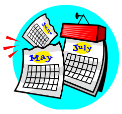


Elizabeth Anderson. Language after stroke. January 2016.

Elizabeth Anderson. Language after stroke. January 2016.

Research

Elizabeth will use the results **in future**



Questions

You can ask Elizabeth **questions**

Elizabeth Anderson
☎ 0114 22 22 412
ecanderson1@sheffield.ac.uk

You can contact her to **sign up**

Research

We may **contact** you **in future**

You do **NOT** have to agree

Complaints

If you wish to make a complaint, please contact:

Traci Walker
☎ 0114 22 22 420
traci.walker@sheffield.ac.uk

Elizabeth Anderson. Language after stroke. January 2016.

Elizabeth Anderson. Language after stroke. January 2016.

Information sheet for family and/or carers of participants with aphasia

The University Of Sheffield.

Department Of Human Communication Sciences

Head of Department
Professor Patricia E. Cowell, BA, MS, PhD

362 Mushroom Lane
Sheffield
S10 2TS
United Kingdom

Telephone: +44 (0) 114 222 2418/ 2402/ 2405
International: +44 (0) 114 222 2418
Fax: +44 (0) 114 2222439
Email: hcs-support@sheffield.ac.uk
http://www.shef.ac.uk/hcs

Information sheet for family or carers of participants with aphasia

Language after stroke

This Information Sheet describes a research project at the University of Sheffield, which is investigating language after stroke.

Adults with aphasia are invited to take part in this research project. This Information Sheet gives you information about the study. You can ask the researcher any questions you may have.

What is this project about?
The researchers are interested in how language is affected by stroke. We are interested in how speakers use words in sentences. Information from this project will contribute to further research on this topic, and it may inform clinical approaches to language after stroke.

What would participants be asked to do?
Participants would first be asked to complete three screening tasks in order to determine whether this study suits their abilities. If participants are comfortable with the screening tasks, they would attend four to six sessions that will each last no more than two hours. Participants can meet the researcher at the Department of Human Communication Sciences at the University of Sheffield, or they can request the researcher to meet them in a quiet room of their own home.

This study involves a variety of language tasks. Two tasks will be performed on the computer. In one task, participants would be asked to replace a blank space in a sentence with a word. For example, they might see the sentence *you can* _____ and the word *dream*. They would be asked to say the entire sentence, *you can dream*. In the other task, participants would be asked to decide whether the word can be used in the sentence. In this example, they would press a button to indicate that *dream* can be used in the sentence.

Participants would also take part in language tasks off the computer. They would be asked to do activities with words and pictures, talk to the researcher and read and produce sentences.

This is **NOT** speech and language therapy. No counselling or therapy will be provided at the sessions.

Who can take part?
Participants must:

- be native speakers of British English;
- have been told by a speech and language therapist that they have aphasia;
- have been living with aphasia for at least six months;
- have no other speech or language difficulty, in addition to aphasia;
- have no history of psychiatric or mental health disorders;
- have had only one brain injury resulting in aphasia. For example, participants cannot have had two strokes.

Participants must complete the screening tasks at the first meeting to attend future sessions.

Will participants receive payment?
No. Participation is voluntary. No rewards or compensation will be given for taking part. No reimbursment for time or travel will be provided.

Can participants stop at any time?
Yes. Participants can withdraw at any time, without giving a reason. There will be no negative consequences for doing so.

Will participation in this study impact the support that participants recieve?
No. Participation in this study will not affect any support that participants receive, such as their attendance at communication groups or speech and language therapy.

What information would be collected?
Participants would be asked background information about their age, language background, education, profession and language abilities. They would provide vocal responses to the tasks in the study. Participants may agree to have their responses audio recorded.

What happens to the data?
Electronic data will be kept securely in password protected storage locations. Paper materials will be kept in a locked filing cabinet in a secure office. The results from this study will contribute to Elizabeth Anderson's PhD research.

Will participants remain anonymous? Will their participation be confidential?
Yes. All the information that is collected about participants will be kept confidential. They will not be identifiable in any reports, publications or presentations of the data.

What are the potential risks or disadvantages of taking part?
Participants may become mentally fatigued during the study. They can rest as much as they need. They can also stop at any time, without giving a reason.

What are the potential benefits of taking part?
There are no direct benefits to participants. However, the data collected in this study will provide valuable information on the language abilities of adults with aphasia. The results may contribute to the development of a clinical tool for adults with aphasia.

Has the project obtained ethical approval?
This project has received ethical approval from the University Research Ethics Committee at the University of Sheffield.

Who is funding the project?
The Faculty of Medicine, Dentistry and Health at the University of Sheffield is funding this project.

Who is on the research team?
Elizabeth Anderson is a PhD student in the Department of Human Communication Sciences. Her project is supervised by Dr Ruth Herbert and Prof Patricia Cowell.

Researcher: Elizabeth Anderson

Department of Human Communication Sciences
University of Sheffield
362 Mushroom Lane
Sheffield
S10 2TS

ecanderson1@sheffield.ac.uk
0114 22 22 412

Supervisors: Dr Ruth Herbert
r.herbert@sheffield.ac.uk
0114 22 22 403

Prof Patricia Cowell
p.e.cowell@sheffield.ac.uk
0114 22 22 426

How can I get more information or sign up?
You can contact Elizabeth Anderson for more information, or to arrange a time to meet. Or, you can let the organiser of your communication group know that you would like to take part, and Elizabeth will contact you directly.

What if there is a problem or I need to make a complaint?
You can speak with Elizabeth Anderson or her supervisors under any circumstances.

If you would like to speak to someone unrelated to the research team, you can contact the Ethics Administrator of the Department of Human Communication Sciences:

Dr Traci Walker
traci.walker@sheffield.ac.uk
0114 22 22 420

If you are not satisfied with the responses from these individuals, you can contact the Office of the Registrar and Secretary at the University of Sheffield at registrar@sheffield.ac.uk.

Thank you for considering this study!

Consent form for typical participants

The University Of Sheffield.

Department Of Human Communication Sciences

Head of Department
Professor Patricia E. Cowell, BA, MS, PhD

362 Mushroom Lane
Sheffield
S10 2TS
United Kingdom

Telephone: +44 (0) 114 222 2418/ 2402/ 2405
International: +44 (0) 114 222 2418
Fax: +44 (0) 114 2222439
Email: hcs-support@sheffield.ac.uk
http://www.shef.ac.uk/hcs

Consent form for participants without aphasia

Language after stroke

Researcher: Elizabeth Anderson

Participant number: __________ **Please initial.**

1. I have read and understood the information sheet.
2. My questions have been answered.
3. I understand that my participation is voluntary. I am free to withdraw at any time, without giving a reason. There are no negative consequences for doing this.
4. I understand that my responses will be confidential. I will not be identifiable in any report, publication or presentation of the data from this project.
5. I agree that my responses can be audio recorded. I give permission for members of the research team to access the recordings.
6. I agree that the data collected from me can be used in future research.
7. I agree to take part in the above research.

Name of participant	Date	Signature
Name of researcher	Date	Signature

Elizabeth Anderson. Language after stroke. November 2015.

Consent form for participants with aphasia

The University Of Sheffield.

Department Of Human Communication Sciences

Head of Department
Professor Patricia E. Cowell, BA, MS, PhD

362 Mushroom Lane
Sheffield
S10 2TS
United Kingdom

Telephone: +44 (0) 114 222 2418/ 2402/ 2405
International: +44 (0) 114 222 2418
Fax: +44 (0) 114 2222439
Email: hcs-support@sheffield.ac.uk
http://www.shef.ac.uk/hcs

Consent form for participants with aphasia

Language after stroke

Researcher: Elizabeth Anderson

Participant number: __________

Information sheet

Please tick ✓

I understand the **information sheet**

I had a chance to ask **questions**

I am **volunteering** to take part

Doing this will **NOT** affect any **support** I receive

I can **stop** at any time

Elizabeth Anderson. Language after stroke. November 2015.

Recordings

I agree that my voice can be **recorded**

The research team will **listen** to the recordings

Confidentiality

The study will **NOT** use my **name**

Future research

I understand that information will be used in **future research**

I **agree** to take part in this research project

Name of participant	Date	Signature
Name of researcher	Date	Signature

Elizabeth Anderson. Language after stroke. November 2015.

Appendix F

This appendix contains the materials employed in the Phase 2 computerised grammaticality judgement task and sentence completion task for typical participants and participants with aphasia.

Grammaticality judgment task – Typical participants

Practice items

You ___ him	NEED	yes
We ___ it from them	TRAIN	no
I ___ through there	SAIL	yes
You ___ it to her	MISS	no
I ___ at him	SMILE	yes
They ___ you something	FIGHT	no
We ___ through there	GAIN	no
They ___ it to you	HIT	yes
We ___ at them	SEEK	no
You ___ us something	MAKE	yes
I ___ it from him	KEEP	yes
They ___ me	YELL	no

Test items

We ___ it to you	TRIP	no
I ___ it from there	LIFT	yes
You ___ her	SIGH	no
We ___ at him	THINK	no
I ___ you something	SMELL	no
They ___ it from me	HEAR	yes
I ___ at them	FIT	no
You ___ them	TRY	yes
We ___ through there	FETCH	no
You ___ at her	SLEEP	no

We ___ him	HATE	yes
I ___ her something	SHOW	yes
They ___ it from him	MEET	no
We ___ at her	LAUGH	yes
I ___ it to him	POST	yes
You ___ through there	HAND	no
I ___ at you	LOOK	yes
We ___ it from her	STEAL	yes
They ___ me	FEAR	yes
I ___ through there	PLACE	no
We ___ her something	GLANCE	no
You ___ it from us	STARE	no
I ___ through there	JUMP	yes
They ___ at him	FLY	yes
I ___ it to her	GROW	no
We ___ him something	BUILD	yes
They ___ at you	SAY	no
I ___ you	LIKE	yes
You ___ it to me	BOIL	no
We ___ through there	LEAD	yes
They ___ him something	SPEAK	no
We ___ it from you	CLIMB	no
You ___ through there	RUN	yes
They ___ it to me	CHUCK	yes
I ___ him	NOD	no
You ___ at them	COME	yes
I ___ it from there	BREAK	no
They ___ through there	ROLL	yes

You ___ me something	SEEM	no
They ___ us	FALL	no
You ___ it to us	SELL	yes
They ___ me something	COOK	yes
You ___ it from him	WIN	yes
They ___ through there	PUT	no
We ___ it to them	SEND	yes
You ___ them something	PACK	yes
We ___ you	LIVE	no
They ___ it to him	CAUSE	no

Sentence completion task – Typical participants

Practice items

I ___ them	WANT
They ___ at us	POINT
You ___ him something	READ
They ___ through there	SLIDE
You ___ it from him	CATCH
We ___ it to her	PUSH

Test items

You ___ at us	SHOUT
I ___ her	PICK
We ___ it to him	DRIVE
I ___ through there	GO
You ___ her something	PAY
They ___ it from there	REACH
I ___ at her	SING
We ___ them	USE
They ___ us something	GIVE

We ___ it from her	BUY
They ___ through there	GET
We ___ at you	TALK
You ___ it from there	FEEL
They ___ it to her	LEAVE
We ___ through there	CUT
They ___ at me	MOVE
I ___ him something	SLIP
You ___ through there	WALK
I ___ it from them	TAKE
You ___ him	LOVE
I ___ it to you	TELL
We ___ them something	OWE
They ___ you	SEE
You ___ it to them	PASS

Grammaticality judgement task – Participants with aphasia

Practice items, Set 1

You ___ him	NEED	yes
You ___ him	YELL	no
You ___ him	MISS	yes
You ___ him	TRAIN	yes
You ___ him	MAKE	no

Test items, Set 1

I ___ them something	STROLL	no
I ___ them something	FEED	yes
I ___ them something	COOK	yes
I ___ them something	SMELL	no
I ___ them something	BUILD	yes

I ___ them something	SEEM	no
I ___ them something	GLANCE	no
I ___ them something	SHOW	yes
I ___ them something	PACK	yes
I ___ them something	SPEAK	no
They ___ through there	CRACK	no
They ___ through there	SNEAK	yes
They ___ through there	FETCH	no
They ___ through there	ROLL	yes
They ___ through there	HAND	no
They ___ through there	PLACE	no
They ___ through there	LEAD	yes
They ___ through there	RUN	yes
They ___ through there	PUT	no
They ___ through there	JUMP	yes
We ___ at her	NAP	no
We ___ at her	SNEER	yes
We ___ at her	THINK	no
We ___ at her	FIT	no
We ___ at her	SAY	no
We ___ at her	LAUGH	yes
We ___ at her	SLEEP	no
We ___ at her	FLY	yes
We ___ at her	LOOK	yes
We ___ at her	COME	yes

Practice items, Set 2

We ___ through there	GAIN	no
We ___ through there	SAIL	yes

We ___ through there	KEEP	no
We ___ through there	SMILE	yes
We ___ through there	HIT	no

Test items, Set 2

You ___ them	TROT	no
You ___ them	VIEW	yes
You ___ them	FALL	no
You ___ them	LIKE	yes
You ___ them	HATE	yes
You ___ them	FEAR	yes
You ___ them	NOD	no
You ___ them	TRY	yes
You ___ them	LIVE	no
You ___ them	SIGH	no
They ___ it to her	LEND	yes
They ___ it to her	HIDE	no
They ___ it to her	TRIP	no
They ___ it to her	POST	yes
They ___ it to her	SEND	yes
They ___ it to her	BOIL	no
They ___ it to her	CHUCK	yes
They ___ it to her	SELL	yes
They ___ it to her	GROW	no
They ___ it to her	CAUSE	no
I ___ it from there	GLARE	no
I ___ it from there	SENSE	yes
I ___ it from there	LIFT	yes
I ___ it from there	STARE	no

I ___ it from there	MEET	no
I ___ it from there	STEAL	yes
I ___ it from there	HEAR	yes
I ___ it from there	CLIMB	no
I ___ it from there	BREAK	no
I ___ it from there	WIN	yes

Sentence completion tasks – Participants with aphasia

Practice items, Set 1

They ___ it from us	CATCH
They ___ it from us	PINCH
They ___ it from us	CHASE
They ___ it from us	FETCH
They ___ it from us	LEARN

Test items, Set 1

I ___ though there	RIDE
I ___ though there	HOP
I ___ though there	GO
I ___ though there	CUT
I ___ though there	WALK
I ___ though there	GET
You ___ at him	POINT
You ___ at him	WAVE
You ___ at him	SING
You ___ at him	TALK
You ___ at him	SHOUT
You ___ at him	MOVE
We ___ her something	DRAW
We ___ her something	THROW

We ___ her something	PAY
We ___ her something	GIVE
We ___ her something	SLIP
We ___ her something	OWE

Practice items, Set 2

I ___her something	FIND
I ___her something	ASK
I ___her something	READ
I ___her something	TEACH
I ___her something	WRITE

Test items, Set 2

They ___ it from there	PULL
They ___ it from there	SNATCH
They ___ it from there	BUY
They ___ it from there	REACH
They ___ it from there	FEEL
They ___ it from there	TAKE
We ___ them	CALL
We ___ them	BEAT
We ___ them	PICK
We ___ them	USE
We ___ them	LOVE
We ___ them	SEE
You ___ it to him	DRAG
You ___ it to him	HAND
You ___ it to him	DRIVE
You ___ it to him	LEAVE
You ___ it to him	TELL

You ___ it to him PASS

Appendix G

Tables A4 and A5 show the construction frequency and lexical frequency of the verbs included in the computerised tasks in Phase 2. Construction frequency refers to the number of times typical participants produced the verb in response to the construction in Phase 2, with a maximum value of 40. Lexical frequency refers to the frequency of the verb lemma in instances per million, based on Leech et al. (2001). In each cell, the value of a verb's construction frequency is shown first, followed by the value of the verb's lexical frequency.

Table A4 Frequency of verbs in Phase 2 grammaticality judgement task

	high cx, high lex		high cx, low lex		low cx, high lex		low cx, low lex	
Caused motion	SEND	9, 250	POST	5, 10	SELL	1, 213	CHUCK	1, 7
	GROW	0, 191	TRIP	0, 8	CAUSE	0, 206	BOIL	0, 12
Conative	LOOK	25, 1151	LAUGH	17, 98	COME	1, 1512	FLY	1, 90
	SAY	0, 6119	FIT	0, 95	THINK	0, 1520	SLEEP	0, 68
Ditransitive	SHOW	9, 598	COOK	10, 37	BUILD	1, 230	PACK	1, 33
	SEEM	0, 624	SMELL	0, 25	SPEAK	0, 261	GLANCE	0, 41
Intransitive motion	RUN	26, 406	JUMP	14, 52	LEAD	1, 334	ROLL	1, 49
	PUT	0, 700	HAND	0, 54	PLACE	0, 150	FETCH	0, 19
Removal	HEAR	16, 367	STEAL	15, 48	WIN	1, 241	LIFT	2, 71
	MEET	0, 339	CLIMB	0, 57	BREAK	0, 193	STARE	0, 84
Transitive	LIKE	18, 424	HATE	23, 50	TRY	1, 552	FEAR	1, 53
	LIVE	0, 329	NOD	0, 60	FALL	0, 273	SIGH	0, 25

Note. Verbs in grey rows included to elicit 'no' judgements. Frequency values for each verb listed as (cx freq, lex freq).

Table A5 Frequency of verbs in Phase 2 sentence completion task

	high cx, high lex		high cx, low lex		low cx, high lex		low cx, low lex	
Caused motion	LEAVE	15, 647	PASS	8, 204	TELL	1, 775	DRIVE	1, 156
Conative	TALK	8, 308	SHOUT	8, 59	MOVE	1, 391	SING	1, 63
Ditransitive	GIVE	34, 1284	OWE	10, 37	PAY	1, 381	SLIP	1, 51
Intransitive motion	GO	25, 2078	WALK	16, 215	GET	2, 2210	CUT	2, 184
Removal	TAKE	25, 1797	BUY	12, 262	FEEL	1, 624	REACH	1, 234
Transitive	SEE	13, 1920	LOVE	25, 150	USE	1, 1071	PICK	1, 150

Note. Frequency values for each verb listed as (cx freq, lex freq). Alternating rows in grey for ease of reference.

Appendix H

This appendix contains the items included in the novel language assessments for participants with aphasia in Phase 2.

Screening tasks

Items in the sentence reading task were taken from PALPA 37 (Kay et al., 1992) and items in the grammaticality judgement task were taken from Kim and Thompson (2000).

Sentence reading	Practice item	The man's moving the horse	
	Test items	The girl's washing the dog	
		The horse's chased by the girl	
		The girl's taller than the dog	
		The cat's licking the man	
		The man's following the dog	
Function word processing	Practice items	each	
		fep	
	Test items	she	out
		undem	mho
		who	under
		poth	both
		zhe	oub
Grammaticality judgement	Practice item	The boy is swimming the girl	
	Test items	The man is snoring	
		The boy is bleeding the girl	
		The lady is weeping her baby	
		The boy is leaning the ladder against the wall	
		The girl is drying the man the dishes	

Lexical decision of function words

Practice items

me

ovep

sde

up

thore

Test items

they	qo
ub	we
hrom	lhem
theru	somedhing
you	throuph
at	him
thep	them
ik	her
there	gou
ag	from
I	to
ver	e
through	it
us	hiw
something	ne

Action-naming task

Images from Druks and Masterson's (2000) *An Object and Action Naming Battery* were used to elicit the following targets in the action-naming task.

Practice items

dancing

ringing

Test items

dropping	cutting
roaring	skiing
singing	weaving
lighting	jumping
kissing	sewing
smiling	tickling
pulling	folding
snowing	riding
cooking	shaving
pointing	marching
rocking	crossing
juggling	sitting
sneezing	pouring
watching	skipping
ironing	playing

Object-naming task

Images from Druks and Masterson's (2000) *An Object and Action Naming Battery* were used to elicit the following targets in the object-naming task.

Practice items

pocket

tongue

Test items

bucket	submarine
tree	fork
dog	hat
mushroom	cow
pram	ball
leg	triangle
bath	basket
crack	flag
bridge	candle
road	stamp
hair	horse
camel	foot
door	box
shoe	fruit
tray	shorts

Sentence production task

Practice items

shooting

flying

Test items

placing	breaking
cuddling	snatching
posting	following
leaning	donating
hiking	removing
deserting	handing
putting	choosing
offering	sticking
pointing	arriving
spitting	hopping

Anagram task

Practice items

The bank lent a business some money

Some scouts hiked through the mountains

Test items

An artist needed the paintbrush

The model smiled at the camera

A chef kept the secret from his rival

The cook made the diners some soup

The runners jogged through a park

A footballer kicked the ball to his teammate

The students learned a lesson from their teacher

The couple gazed at the sunset

A child skipped through the playground

The pilot asked a question

A professor taught her class the information

The coach threw a racquet to the tennis play

Appendix I

Tables A6 and A7 show summary statistics for by-item analyses of response time data from typical participants in the Phase 2 grammaticality judgement task. Statistics for the raw data are shown in the upper panel, in milliseconds, and the values from the back transformations are shown below. Note that standard deviations of transformed values cannot be back transformed and so are represented in the tables by (-).

Table A6 Summary statistics for by-item analyses of dataset containing only accurate responses from typical participants in grammaticality judgement task

	M	95% CI		*SD*	min	max
		Lower bound	Upper bound			
Raw data						
high cx, high lex	1641	1405	1878	225	1405	1977
high cx, low lex	1673	1548	1797	119	1515	1871
low cx, high lex	2112	1789	2434	308	1770	2601
low cx, low lex	2039	1766	2311	260	1531	2227
Back transformations						
high cx, high lex	1479	1342	1647	-	1321	1724
high cx, low lex	1524	1447	1610	-	1414	1650
low cx, high lex	1789	1610	2012	-	1590	2075
low cx, low lex	1755	1567	1994	-	1425	1862

Note. Mean (*M*), lower and upper bound of 95% confidence interval (CI), standard deviation (*SD*), minimum value (min) and maximum value (max).

Table A7 Summary statistics for by-item analyses of dataset containing data from all participants

	M	95% CI		*SD*	min	max
		Lower bound	Upper bound			
Raw data						
high cx, high lex	1592	1395	1788	187	1351	1857
high cx, low lex	1679	1543	1816	130	1503	1870
low cx, high lex	2158	1814	2503	328	1731	2620
low cx, low lex	2113	1847	2380	254	1632	2351
Back transformations						
high cx, high lex	1445	1318	1600	-	1272	1634
high cx, low lex	1509	1437	1590	-	1399	1616
low cx, high lex	1811	1614	2062	-	1565	2088
low cx, low lex	1789	1600	2025	-	1464	1957

Note. Mean (*M*), lower and upper bound of 95% confidence interval (CI), standard deviation (*SD*), minimum value (min) and maximum value (max).

Appendix J

Tables in this Appendix show results from post-hoc paired-samples *t*-tests on response time data from typical participants in the Phase 2 grammaticality judgement task for each of the six constructions in the task.

Table A8 Post-hoc paired-samples *t*-tests on typical participants' response times to the caused motion construction in the grammaticality judgement task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
Raw data				
high cx, high lex	-			
high cx, low lex	$t(83) = 0.30$ $p = 0.766$ $r = 0.03$	-		
low cx, high lex	$t(84) = -0.42$ $p = 0.674$ $r = 0.05$	$t(84) = -0.72$ $p = 0.477$ $r = 0.08$	-	
low cx, low lex	$t(83) = -3.50$ $p = 0.001$ $r = 0.36$	$t(83) = -3.64$ $p < 0.001$ $r = 0.37$	$t(84) = -3.61$ $p = 0.001$ $r = 0.37$	-
Transformed data				
high cx, high lex	-			
high cx, low lex	$t(83) = -0.49$ $p = 0.628$ $r = 0.05$	-		
low cx, high lex	$t(84) = 0.87$ $p = 0.389$ $r = 0.09$	$t(84) = 1.18$ $p = 0.243$ $r = 0.13$	-	
low cx, low lex	$t(84) = 5.01$ $p < 0.001$ $r = 0.48$	$t(83) = 6.39$ $p < 0.001$ $r = 0.57$	$t(84) = 4.85$ $p < 0.001$ $r = 0.47$	-

Note. Significance at $p = 0.008$ with Bonferroni correction (0.05 ÷ 6 = 0.008).

Table A9 Post-hoc paired-samples *t*-tests on typical participants' response times to the conative construction in the grammaticality judgement task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
Raw data				
high cx, high lex	-			
high cx, low lex	$t(84) = -0.49$ $p = 0.625$ $r = 0.05$	-		
low cx, high lex	$t(73) = -7.21$ $p < 0.001$ $r = 0.64$	$t(74) = -6.77$ $p < 0.001$ $r = 0.62$	-	
low cx, low lex	$t(81) = -6.27$ $p < 0.001$ $r = 0.57$	$t(82) = -6.64$ $p < 0.001$ $r = 0.59$	$t(72) = -0.32$ $p = 0.752$ $r = 0.04$	-
Transformed data				
high cx, high lex	-			
high cx, low lex	$t(84) = 0.76$ $p = 0.451$ $r = 0.08$	-		
low cx, high lex	$t(73) = 10.44$ $p < 0.001$ $r = 0.77$	$t(74) = 10.96$ $p < 0.001$ $r = 0.79$	-	
low cx, low lex	$t(81) = 8.58$ $p < 0.001$ $r = 0.69$	$t(82) = 11.49$ $p < 0.001$ $r = 0.79$	$t(72) = -1.43$ $p = 0.156$ $r = 0.17$	-

Note. Significance at $p = 0.008$ with Bonferroni correction (0.05 ÷ 6 = 0.008).

Table A10 Post-hoc paired-samples *t*-tests on typical participants' response times to the ditransitive construction in the grammaticality judgement task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
Raw data				
high cx, high lex	-			
high cx, low lex	$t(84) = -3.15$ $p = 0.002$ $r = 0.33$	-		
low cx, high lex	$t(82) = -3.88$ $p < 0.001$ $r = 0.39$	$t(82) = -1.82$ $p = 0.073$ $r = 0.20$	-	
low cx, low lex	$t(79) = -6.19$ $p < 0.001$ $r = 0.57$	$t(78) = -3.08$ $p = 0.003$ $r = 0.33$	$t(76) = -1.95$ $p = 0.055$ $r = 0.22$	-
Transformed data				
high cx, high lex	-			
high cx, low lex	$t(84) = 3.97$ $p < 0.001$ $r = 0.40$	-		
low cx, high lex	$t(82) = 6.33$ $p < 0.001$ $r = 0.57$	$t(82) = 2.11$ $p = 0.038$ $r = 0.23$	-	
low cx, low lex	$t(79) = 7.97$ $p < 0.001$ $r = 0.67$	$t(78) = 4.75$ $p < 0.001$ $r = 0.47$	$t(76) = 3.17$ $p = 0.002$ $r = 0.34$	-

Note. Significance at $p = 0.008$ with Bonferroni correction (0.05 ÷ 6 = 0.008).

Table A11 Post-hoc paired-samples *t*-tests on typical participants' response times to the intransitive motion construction in the grammaticality judgement task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
Raw data				
high cx, high lex	-			
high cx, low lex	$t(84) = -0.01$ $p = 0.995$ $r = 0.00$	-		
low cx, high lex	$t(69) = -6.55$ $p < 0.001$ $r = 0.62$	$t(69) = -7.26$ $p < 0.001$ $r = 0.66$	-	
low cx, low lex	$t(82) = -4.56$ $p < 0.001$ $r = 0.45$	$t(81) = -5.01$ $p < 0.001$ $r = 0.49$	$t(69) = 3.40$ $p = 0.001$ $r = 0.38$	-
Transformed data				
high cx, high lex	-			
high cx, low lex	$t(84) = -0.27$ $p = 0.790$ $r = 0.03$	-		
low cx, high lex	$t(69) = 7.13$ $p < 0.001$ $r = 0.65$	$t(69) = 7.95$ $p < 0.001$ $r = 0.69$	-	
low cx, low lex	$t(82) = 7.14$ $p < 0.001$ $r = 0.62$	$t(81) = 7.83$ $p < 0.001$ $r = 0.66$	$t(69) = -2.70$ $p = 0.009$ $r = 0.31$	-

Note. Significance at $p = 0.008$ with Bonferroni correction (0.05 ÷ 6 = 0.008).

Table A12 Post-hoc paired-samples *t*-tests on typical participants' response times to the removal construction in the grammaticality judgement task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
Raw data				
high cx, high lex	-			
high cx, low lex	$t(81) = -0.03$ $p = 0.975$ $r = 0.00$	-		
low cx, high lex	$t(78) = -3.59$ $p = 0.001$ $r = 0.38$	$t(79) = -4.09$ $p < 0.001$ $r = 0.42$	-	
low cx, low lex	$t(80) = -3.84$ $p < 0.001$ $r = 0.39$	$t(80) = -4.61$ $p < 0.001$ $r = 0.46$	$t(77) = 0.22$ $p = 0.825$ $r = 0.03$	-
Transformed data				
high cx, high lex	-			
high cx, low lex	$t(81) = -0.53$ $p = 0.598$ $r = 0.06$	-		
low cx, high lex	$t(78) = 5.35$ $p < 0.001$ $r = 0.52$	$t(79) = 5.41$ $p < 0.001$ $r = 0.52$	-	
low cx, low lex	$t(80) = 4.89$ $p < 0.001$ $r = 0.48$	$t(80) = 5.36$ $p < 0.001$ $r = 0.51$	$t(77) = 0.04$ $p = 0.972$ $r = 0.00$	-

Note. Significance at $p = 0.008$ with Bonferroni correction (0.05 ÷ 6 = 0.008).

Table A13 Post-hoc paired-samples *t*-tests on typical participants' response times to the transitive construction in the grammaticality judgement task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
Raw data				
high cx, high lex	-			
high cx, low lex	$t(84) = -6.57$ $p < 0.001$ $r = 0.58$	-		
low cx, high lex	$t(80) = -7.20$ $p < 0.001$ $r = 0.63$	$t(81) = -4.94$ $p < 0.001$ $r = 0.48$	-	
low cx, low lex	$t(84) = -3.95$ $p < 0.001$ $r = 0.40$	$t(85) = -0.74$ $p = 0.464$ $r = 0.08$	$t(81) = 5.09$ $p < 0.001$ $r = 0.49$	-
Transformed data				
high cx, high lex	-			
high cx, low lex	$t(84) = 6.17$ $p < 0.001$ $r = 0.56$	-		
low cx, high lex	$t(80) = 9.27$ $p < 0.001$ $r = 0.72$	$t(81) = 5.52$ $p < 0.001$ $r = 0.52$	-	
low cx, low lex	$t(84) = 5.34$ $p < 0.001$ $r = 0.50$	$t(85) = -0.41$ $p = 0.681$ $r = 0.04$	$t(81) = -6.46$ $p < 0.001$ $r = 0.58$	-

Note. Significance at $p = 0.008$ with Bonferroni correction (0.05 ÷ 6 = 0.008).

Appendix K

Tables in this Appendix show results of Wilcoxon signed-rank tests on the number of target responses from participants with aphasia in the Phase 2 grammaticality judgement task between conditions in each of the six constructions in the task.

Table A14 Wilcoxon signed-rank tests on number of target responses from participants with aphasia to the caused motion construction in the grammaticality judgement task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	-			
high cx, low lex	*T* = 0.00 *p* = 0.317 *r* = -0.19	-		
low cx, high lex	*T* = 0.00 *p* = 1.000 *r* = 0.00	*T* = 0.00 *p* = 0.317 *r* = -0.19	-	
low cx, low lex	*T* = 0.00 *p* = 0.005 *r* = -0.53	*T* = 0.00 *p* = 0.008 *r* = -0.50	*T* = 0.00 *p* = 0.005 *r* = -0.53	-

Table A15 Wilcoxon signed-rank tests on number of target responses from participants with aphasia to the conative construction in the grammaticality judgement task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	-			
high cx, low lex	*T* = 0.00 *p* = 0.157 *r* = -0.27	-		
low cx, high lex	*T* = 0.00 *p* = 0.005 *r* = -0.53	*T* = 11.00 *p* = 0.058 *r* = -0.36	-	
low cx, low lex	*T* = 0.00 *p* = 0.003 *r* = -0.57	*T* = 12.00 *p* = 0.035 *r* = -0.40	*T* = 2.00 *p* = 0.564 *r* = -0.11	-

Table A16 Wilcoxon signed-rank tests on number of target responses from participants with aphasia to the ditransitive construction in the grammaticality judgement task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	-			
high cx, low lex	*T* = 0.00 *p* = 0.157 *r* = -0.27	-		
low cx, high lex	*T* = 0.00 *p* = 0.083 *r* = -0.33	*T* = 2.00 *p* = 0.564 *r* = -0.11	-	
low cx, low lex	*T* = 0.00 *p* = 0.046 *r* = -0.38	*T* = 0.00 *p* = 0.157 *r* = -0.27	*T* = 6.00 *p* = 0.655 *r* = -0.08	-

Table A17 Wilcoxon signed-rank tests on number of target responses from participants with aphasia to the intransitive motion construction in the grammaticality judgement task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	-			
high cx, low lex	*T* = 0.00 *p* = 0.317 *r* = -0.19	-		
low cx, high lex	*T* = 0.00 *p* = 0.046 *r* = -0.38	*T* = 0.00 *p* = 0.083 *r* = -0.33	-	
low cx, low lex	*T* = 0.00 *p* = 0.083 *r* = -0.33	*T* = 2.50 *p* = 0.317 *r* = -0.19	*T* = 6.00 *p* = 0.655 *r* = -0.08	-

Table A18 Wilcoxon signed-rank tests on number of target responses from participants with aphasia to the removal construction in the grammaticality judgement task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	-			
high cx, low lex	*T* = 0.00 *p* = 0.025 *r* = -0.42	-		
low cx, high lex	*T* = 0.00 *p* = 0.025 *r* = -0.42	*T* = 5.00 *p* = 1.000 *r* = 0.00	-	
low cx, low lex	*T* = 3.00 *p* = 0.180 *r* = -0.25	*T* = 0.00 *p* = 0.157 *r* = -0.27	*T* = 7.00 *p* = 0.414 *r* = -0.15	-

Table A19 Wilcoxon signed-rank tests on number of target responses from participants with aphasia to the transitive construction in the grammaticality judgement task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	-			
high cx, low lex	*T* = 2.50 *p* = 0.317 *r* = -0.19	-		
low cx, high lex	*T* = 3.00 *p* = 0.180 *r* = -0.25	*T* = 2.00 *p* = 0.564 *r* = -0.11	-	
low cx, low lex	*T* = 2.50 *p* = 0.317 *r* = -0.19	*T* = 0.00 *p* = 1.000 *r* = 0.00	*T* = 2.00 *p* = 0.564 *r* = -0.11	-

Appendix L

Tables in this Appendix show results of Wilcoxon signed-rank tests on response times from participants with aphasia in the Phase 2 grammaticality judgement task between conditions in each of the six constructions in the task.

Table A20 Wilcoxon signed-rank tests on response times from participants with aphasia to the caused motion construction in the grammaticality judgement task

	n	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	14	-			
high cx, low lex	13	*T* = 31.00 *p* = 0.311 *r* = -0.19	-		
low cx, high lex	14	*T* = 52.00 *p* = 0.975 *r* = -0.01	*T* = 41.00 *p* = 0.753 *r* = -0.06	-	
low cx, low lex	6	*T* = 8.00 *p* = 0.600 *r* = -0.12	*T* = 3.00 *p* = 0.116 *r* = -0.36	*T* = 10.00 *p* = 0.917 *r* = -0.03	-

Table A21 Wilcoxon signed-rank tests on response times from participants with aphasia to the conative construction in the grammaticality judgement task

	n	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	14	-			
high cx, low lex	12	*T* = 37.00 *p* = 0.875 *r* = -0.03	-		
low cx, high lex	6	*T* = 3.00 *p* = 0.116 *r* = -0.35	*T* = 1.00 *p* = 0.144 *r* = -0.34	-	
low cx, low lex	5	*T* = 6.00 *p* = 0.686 *r* = -0.09	*T* = 0.00 *p* = 0.109 *r* = -0.39	*T* = 9.00 *p* = 0.144 *r* = -0.44	-

Table A22 Wilcoxon signed-rank tests on response times from participants with aphasia to the ditransitive construction in the grammaticality judgement task

	n	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	14	-			
high cx, low lex	12	*T* = 29.00 *p* = 0.433 *r* = -0.15	-		
low cx, high lex	11	*T* = 23.00 *p* = 0.374 *r* = -0.18	*T* = 22.00 *p* = 0.575 *r* = -0.12	-	
low cx, low lex	10	*T* = 13.00 *p* = 0.139 *r* = -0.30	*T* = 23.00 *p* = 0.646 *r* = -0.10	*T* = 4.00 *p* = 0.050 *r* = -0.43	-

Table A23 Wilcoxon signed-rank tests on response times from participants with aphasia to the intransitive motion construction in the grammaticality judgement task

	n	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	14	-			
high cx, low lex	13	$T = 30.00$ $p = 0.279$ $r = -0.21$	-		
low cx, high lex	10	$T = 14.00$ $p = 0.169$ $r = -0.28$	$T = 13.00$ $p = 0.139$ $r = -0.31$	-	
low cx, low lex	11	$T = 24.00$ $p = 0.424$ $r = -0.16$	$T = 0.00$ $p = 0.005$ $r = -0.57$	$T = 16.00$ $p = 0.779$ $r = -0.06$	-

Table A24 Wilcoxon signed-rank tests on response times from participants with aphasia to the removal construction in the grammaticality judgement task

	n	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	13	-			
high cx, low lex	8	$T = 8.00$ $p = 0.161$ $r = -0.31$	-		
low cx, high lex	8	$T = 12.00$ $p = 0.401$ $r = -0.18$	$T = 10.00$ $p = 0.917$ $r = 0.03$	-	
low cx, low lex	10	$T = 17.00$ $p = 0.515$ $r = -0.14$	$T = 15.00$ $p = 0.674$ $r = -0.10$	$T = 7.00$ $p = 0.463$ $r = -0.17$	-

Table A25 Wilcoxon signed-rank tests on response times from participants with aphasia to the transitive construction in the grammaticality judgement task

	n	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	9	-			
high cx, low lex	11	$T = 14.00$ $p = 0.575$ $r = -0.13$	-		
low cx, high lex	12	$T = 5.00$ $p = 0.069$ $r = -0.40$	$T = 17.00$ $p = 0.285$ $r = -0.22$	-	
low cx, low lex	11	$T = 11.00$ $p = 0.327$ $r = -0.22$	$T = 27.00$ $p = 0.594$ $r = -0.11$	$T = 10.00$ $p = 0.074$ $r = -0.37$	-

Appendix M

This Appendix contains results from Bayesian Standardised Difference Tests (BSDTs) comparing the difference between response times from participants with aphasia to high frequency verbs and low frequency verbs to the difference from typical participants in the grammaticality judgement task. BSDTs were carried out on raw and transformed data.

Results are reported following the recommendations of Crawford et al. (2010). Results are provided in tabular format and include information on the value of response times to high and low frequency verbs entered into the analysis for each participant with aphasia., e.g. median raw response times and mean transformed response times; the *z*-score associated with high and low frequency items; the two-tailed *p*-value of the BSDT; the estimated percentage of the typical population expected to demonstrate a more extreme difference than the individual with aphasia, in the same direction, as both a point and interval estimate; and an estimated effect size. The effect size for a BSDT is expressed as z_{DCC}, referring to the *z* of the difference between a case and controls. This effect size reflects the number of standard deviations that a case's difference deviates from the mean difference in controls. The sign of z_{DCC} depends on the order that tasks are input into the computation; in the present study, a positive value of z_{DCC} indicates that a participant with aphasia demonstrated a higher standardised score on high frequency than low frequency items. A negative value of z_{DCC} indicates that a participant with aphasia demonstrated a higher standardised score on low frequency than high frequency items.

Table A26 Results of Bayesian Standardised Difference Tests comparing response times to verbs with high and low construction frequency on raw data from grammaticality judgement task

	Median response time to high frequency verbs (*n*)	Median response time to low frequency verbs (*n*)	*z*-score for high frequency verbs	*z*-score for low frequency verbs	Significance test *p*-value (two-tailed)	Estimated percentage of the typical population displaying more extreme discrepancy than PWA, in same direction		Estimated effect size (z_{DCC})	
						Point	95% CI	Point	95% CI
BF	2009 (12)	3136 (9)	0.94	1.34	0.464	23%	14% - 34%	-0.74	-1.08 - -0.41
CJ	4129 (10)	4058 (11)	6.24	2.59	<0.001***	0%	0% - 0%	6.80	5.46 - 8.21
GW	2589 (12)	2988 (11)	2.39	1.14	0.024*	1%	0% - 4%	2.33	1.80 - 2.89
HM	1619 (12)	2373 (8)	-0.04	0.30	0.532	27%	19% - 35%	-0.63	-0.86 - -0.40
IC	2500 (12)	3609 (12)	2.17	1.98	0.733	37%	20% - 55%	0.35	-0.13 - 0.83
JF	2047 (9)	2838 (9)	1.03	0.93	0.851	43%	31% - 54%	0.19	-0.10 - 0.48
KT	2120 (10)	4749 (7)	1.22	3.53	<0.001***	0%	0% - 0%	-4.30	-5.16 - -3.49
MJ	5956 (12)	7667 (9)	10.80	7.49	<0.001***	0%	0% - 0%	6.18	4.12 - 8.31
PD	4009 (11)	6438 (10)	5.94	5.82	0.850	42%	7% - 84%	0.22	-1.00 - 1.45
RE	3841 (12)	5531 (8)	5.52	4.59	0.128	6%	0% - 26%	1.73	0.66 - 2.83
SP	5902 (7)	5632 (3)	10.67	4.72	<0.001***	0%	0% - 0%	11.07	8.84 - 13.43
UT	4412 (10)	3489 (6)	6.94	1.82	<0.001***	0%	0% - 0%	9.56	7.91 - 11.31
VH	4104 (9)	3821 (3)	6.17	2.27	<0.001***	0%	0% - 0%	7.28	5.92 - 8.73
WD	7958 (9)	9282 (8)	15.80	9.68	<0.001***	0%	0% - 0%	11.42	8.35 - 14.61

Note. Typical mean RT to high frequency verbs 1633 msec (*SD* 400) and to low frequency verbs 2151 msec (*SD* 737).
* $p < 0.05$. *** $p < 0.001$.

Table A27 Results of Bayesian Standardised Difference Tests comparing response times to verbs with high and low construction frequency on transformed data from grammaticality judgement task

	Mean response time to high frequency verbs (*n*)	Mean response time to low frequency verbs (*n*)	*z*-score for high frequency verbs	*z*-score for low frequency verbs	Significance test *p*-value (two-tailed)	Estimated percentage of the typical population displaying more extreme discrepancy than PWA, in same direction		Estimated effect size (z_{DCC})	
						Point	95% CI	Point	95% CI
BF	2128 (12)	3125 (9)	-1.50	-1.64	0.778	39%	25% - 54%	0.29	-0.10 - 0.68
CJ	3704 (10)	3846 (11)	-2.93	-2.07	0.099	5%	1% - 13%	-1.72	-2.34 - -1.12
GW	2500 (12)	3125 (11)	-2.00	-1.64	0.483	24%	12% - 39%	-0.72	-1.16 - -0.28
HM	1493 (12)	2273 (8)	-0.07	-0.79	0.156	8%	4% - 13%	1.43	1.13 - 1.75
IC	2778 (12)	3226 (12)	-2.29	-1.71	0.264	13%	5% - 25%	-1.15	-1.65 - -0.66
JF	2128 (9)	3226 (9)	-1.50	-1.71	0.673	34%	20% - 49%	0.43	0.03 - 0.83
KT	2381 (10)	3846 (7)	-1.86	-2.07	0.675	34%	19% - 51%	0.43	-0.03 - 0.89
MJ	5882 (12)	7692 (9)	-3.64	-3.00	0.226	11%	2% - 29%	-1.29	-2.04 - -0.56
PD	3226 (11)	5556 (10)	-2.64	-2.64	0.999	50%	28% - 72%	0.00	-0.58 - 0.58
RE	3571 (12)	4348 (8)	-2.86	-2.29	0.271	14%	4% - 29%	-1.15	-1.75 - -0.56
SP	5556 (7)	5000 (3)	-3.57	-2.50	0.043*	2%	0% - 8%	-2.15	-2.90 - -1.43
UT	3704 (10)	3571 (6)	-2.93	-1.93	0.055	3%	0% - 8%	-2.01	-2.64 - -1.40
VH	4000 (9)	3846 (3)	-3.07	-2.07	0.056	3%	0% - 9%	-2.01	-2.66 - -1.37
WD	4762 (9)	6667 (8)	-3.36	-2.86	0.342	17%	4% - 37%	-1.00	-1.70 - -0.32

Note. Typical mean RT to high frequency verbs 1475 msec and to low frequency verbs 1808 msec. * $p < 0.05$.

Table A28 Results of Bayesian Standardised Difference Tests comparing response times to verbs with high and low lexical frequency on raw data from grammaticality judgement task

	Median response time to high frequency verbs (*n*)	Median response time to low frequency verbs (*n*)	*z*-score for high frequency verbs	*z*-score for low frequency verbs	Significance test *p*-value (two-tailed)	Estimated percentage of the typical population displaying more extreme discrepancy than PWA, in same direction		Estimated effect size (z_{DCC})	
						Point	95% CI	Point	95% CI
BF	2910 (10)	2161 (11)	1.98	0.44	0.003**	0%	0% - 1%	3.05	2.50 - 3.62
CJ	4129 (12)	4058 (9)	4.28	3.67	0.267	13%	2% - 36%	1.21	0.35 - 2.08
GW	3002 (11)	2672 (12)	2.16	1.31	0.104	5%	2% - 12%	1.67	1.19 - 2.16
HM	1415 (10)	1916 (10)	-0.83	0.03	0.094	5%	2% - 9%	-1.69	-2.04 - -1.36
IC	2178 (12)	3539 (12)	0.60	2.78	<0.001***	0%	0% - 0%	-4.31	-5.09 - -3.57
JF	2512 (8)	2UT (10)	1.23	1.20	0.948	47%	35% - 60%	0.07	-0.26 - 0.39
KT	2532 (10)	4749 (7)	1.27	4.84	<0.001***	0%	0% - 0%	-7.05	-8.34 - -5.84
MJ	6547 (11)	6459 (10)	8.83	7.75	0.109	5%	0% - 34%	2.14	0.41 - 3.90
PD	3953 (11)	4522 (10)	3.95	4.45	0.360	18%	3% - 46%	-1.00	-1.91 - -0.11
RE	3761 (10)	4661 (10)	3.58	4.69	0.048*	2%	0% - 10%	-2.19	-3.13 - -1.26
SP	5632 (7)	5902 (3)	7.11	6.80	0.625	31%	2% - 80%	0.61	-0.84 - 2.06
UT	3812 (8)	3765 (8)	3.68	3.17	0.343	17%	9% - 39%	1.01	0.27 - 1.77
VH	3914 (8)	4562 (4)	3.87	4.52	0.243	12%	1% - 35%	-1.28	-2.20 - -0.38
WD	8924 (11)	5422 (6)	13.30	5.98	<0.001***	0%	0% - 0%	14.47	11.60 - 17.49

Note. Typical mean RT to high frequency verbs 1857 msec (*SD* 531) and to low frequency verbs 1901 msec (SD 588).
* $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

Table A29 Results of Bayesian Standardised Difference Tests comparing response times to verbs with high and low lexical frequency on transformed data from grammaticality judgment task

	Mean response time to high frequency verbs (*n*)	Mean response time to low frequency verbs (*n*)	*z*-score for high frequency verbs	*z*-score for low frequency verbs	Significance test *p*-value (two-tailed)	Estimated percentage of the typical population displaying more extreme discrepancy than PWA, in same direction		Estimated effect size (z_{DCC})	
						Point	95% CI	Point	95% CI
BF	2564 (10)	2326 (11)	-1.71	-1.29	0.318	16%	8% - 27%	-1.02	-1.43 - -0.62
CJ	3704 (12)	4000 (9)	-2.57	-2.57	0.999	50%	28% - 72%	0.00	-0.58 - 0.57
GW	2857 (11)	2703 (12)	-2.00	-1.71	0.508	25%	13% - 41%	-0.68	-1.13 - -0.23
HM	1587 (10)	1886 (10)	0.00	-0.57	0.179	9%	5% - 14%	1.35	1.06 - 1.66
IC	2703 (12)	3333 (12)	-1.86	-2.21	0.410	21%	9% - 36%	0.85	0.36 - 1.34
JF	2564 (8)	2564 (10)	-1.71	-1.57	0.739	37%	23% - 53%	0.34	-0.74 - 0.06
KT	2564 (10)	3448 (7)	-1.71	-2.29	0.189	9%	3% - 20%	1.35	0.85 - 1.87
MJ	6667 (11)	6250 (10)	-3.43	-3.21	0.633	32%	11% - 58%	-0.51	-1.24 - 0.21
PD	3846 (11)	4167 (10)	-2.64	-2.64	0.999	50%	28% - 72%	0.00	-0.59 - 0.59
RE	3846 (10)	3846 (10)	-2.64	-2.50	0.744	37%	18% - 59%	-0.34	-0.92 - 0.24
SP	5000 (7)	6667 (3)	-3.07	-3.29	0.632	32%	11% - 57%	0.51	-0.19 - 1.21
UT	3333 (8)	4000 (8)	-2.36	-2.57	0.625	31%	14% - 52%	0.51	-0.05 - 1.07
VH	3846 (8)	4000 (4)	-2.64	-2.57	0.870	44%	23% - 66%	-0.17	-0.75 - 0.41
WD	6250 (11)	4762 (6)	-3.36	-2.86	0.263	13%	3% - 31%	-1.19	-1.89 - -0.49

Note. Typical mean RT to high frequency verbs 1595 msec and to low frequency verbs 1639 msec.

Appendix N

Tables A30 and A31 show summary statistics for by-item analyses of verb production latencies from typical participants in the Phase 2 sentence completion task. Statistics for the raw data are shown in the upper panel, in milliseconds, and the values from the back transformations are shown below. Note that standard deviations of transformed values cannot be back transformed and so are represented in the tables by (-).

Table A30 Summary statistics for by-item analyses of verb production latencies from participants with no excluded responses in sentence completion task

	M	95% CI		*SD*	min	max
		Lower bound	Upper bound			
Raw data						
high cx, high lex	1023	970	1075	50	969	1100
high cx, low lex	1108	1017	1200	87	1036	1275
low cx, high lex	1054	1012	1097	41	999	1117
low cx, low lex	1069	1038	1099	29	1028	1093
Back transformations						
high cx, high lex	978	938	1022	-	938	1047
high cx, low lex	1057	992	1132	-	1008	1197
low cx, high lex	1002	961	1046	-	956	1065
low cx, low lex	1027	1002	1054	-	996	1053

Note. Mean (*M*), lower and upper bound of 95% confidence interval (CI), standard deviation (*SD*), minimum value (min) and maximum value (max). n = 36.

Table A31 Summary statistics for by-item analyses of verb production latencies from typical participants in sentence completion task

	M	95% CI		*SD*	min	max
		Lower bound	Upper bound			
Raw data						
high cx, high lex	1039	995	1082	41	982	1082
high cx, low lex	1093	1024	1162	65	1036	1221
low cx, high lex	1074	1045	1103	28	1022	1097
low cx, low lex	1069	1045	1093	23	1048	1109
Back transformations						
high cx, high lex	985	954	1017	-	949	1013
high cx, low lex	1034	974	1102	-	974	1162
low cx, high lex	1005	978	1033	-	957	1031
low cx, low lex	1016	990	1043	-	991	1058

Note. Mean (*M*), lower and upper bound of 95% confidence interval (CI), standard deviation (*SD*), minimum value (min) and maximum value (max). N = 86.

Appendix O

The tables in this Appendix show the results of post-hoc paired-samples *t*-tests on verb production latencies from typical participants in the Phase 2 sentence completion task, which were carried out to explore interactions in individual constructions.

Table A32 Post-hoc paired-samples *t*-tests on typical participants' verb production latencies to the caused motion construction in the sentence completion task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
Raw data				
high cx, high lex	-			
high cx, low lex	$t(73) = -0.19$ $p = 0.851$ $r = 0.02$	-		
low cx, high lex	$t(67) = -0.28$ $p = 0.784$ $r = 0.03$	$t(74) = -0.10$ $p = 0.921$ $r = 0.01$	-	
low cx, low lex	$t(72) = 0.48$ $p = 0.635$ $r = 0.06$	$t(82) = 0.62$ $p = 0.536$ $r = 0.07$	$t(74) = 0.78$ $p = 0.440$ $r = 0.09$	-
Transformed data				
high cx, high lex	-			
high cx, low lex	$t(73) = 0.23$ $p = 0.821$ $r = 0.03$	-		
low cx, high lex	$t(67) = -0.06$ $p = 0.951$ $r = 0.01$	$t(74) = -0.53$ $p = 0.597$ $r = 0.06$	-	
low cx, low lex	$t(72) = -0.33$ $p = 0.745$ $r = 0.04$	$t(82) = -0.69$ $p = 0.494$ $r = 0.08$	$t(74) = -0.63$ $p = 0.531$ $r = 0.07$	-

Note. Significance at $p = 0.008$ with Bonferroni correction ($0.05 \div 6 = 0.008$).

Table A33 Post-hoc paired-samples *t*-tests on typical participants' verb production latencies to the conative construction in the sentence completion task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
Raw data				
high cx, high lex	-			
high cx, low lex	$t(85) = 1.04$ $p = 0.303$ $r = 0.11$	-		
low cx, high lex	$t(78) = -1.39$ $p = 0.168$ $r = 0.16$	$t(78) = -1.49$ $p = 0.139$ $r = 0.17$	-	
low cx, low lex	$t(83) = 0.37$ $p = 0.712$ $r = 0.04$	$t(83) = -0.53$ $p = 0.595$ $r = 0.06$	$t(77) = 1.00$ $p = 0.319$ $r = 0.11$	-
Transformed data				
high cx, high lex	-			
high cx, low lex	$t(85) = -1.40$ $p = 0.166$ $r = 0.15$	-		
low cx, high lex	$t(78) = 0.48$ $p = 0.633$ $r = 0.05$	$t(78) = 1.14$ $p = 0.257$ $r = 0.13$	-	
low cx, low lex	$t(83) = -0.60$ $p = 0.552$ $r = 0.07$	$t(83) = 0.57$ $p = 0.572$ $r = 0.06$	$t(77) = -0.74$ $p = 0.462$ $r = 0.08$	-

Note. Significance at $p = 0.008$ with Bonferroni correction ($0.05 \div 6 = 0.008$).

Table A34 Post-hoc paired-samples *t*-tests on typical participants' verb production latencies to the ditransitive construction in the sentence completion task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
Raw data				
high cx, high lex	-			
high cx, low lex	$t(79) = -6.60$ $p < 0.001$ $r = 0.60$	-		
low cx, high lex	$t(81) = -2.49$ $p = 0.015$ $r = 0.27$	$t(81) = 3.93$ $p < 0.001$ $r = 0.40$	-	
low cx, low lex	$t(82) = -2.37$ $p = 0.020$ $r = 0.25$	$t(82) = 4.66$ $p < 0.001$ $r = 0.46$	$t(84) = 0.15$ $p = 0.885$ $r = 0.02$	-
Transformed data				
high cx, high lex	-			
high cx, low lex	$t(79) = 8.33$ $p < 0.001$ $r = 0.68$	-		
low cx, high lex	$t(81) = 2.56$ $p = 0.012$ $r = 0.27$	$t(81) = -5.07$ $p < 0.001$ $r = 0.49$	-	
low cx, low lex	$t(82) = 3.11$ $p = 0.003$ $r = 0.32$	$t(82) = -6.13$ $p < 0.001$ $r = 0.56$	$t(84) = 0.40$ $p = 0.691$ $r = 0.04$	-

Note. Significance at $p = 0.008$ with Bonferroni correction ($0.05 \div 6 = 0.008$).

Table A35 Post-hoc paired-samples *t*-tests on typical participants' verb production latencies to the intransitive motion construction in the sentence completion task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
Raw data				
high cx, high lex	-			
high cx, low lex	$t(83) = -4.12$ $p < 0.001$ $r = 0.41$	-		
low cx, high lex	$t(78) = -2.93$ $p = 0.004$ $r = 0.31$	$t(80) = 0.58$ $p = 0.565$ $r = 0.06$	-	
low cx, low lex	$t(82) = -5.11$ $p < 0.001$ $r = 0.49$	$t(84) = -1.03$ $p = 0.308$ $r = 0.11$	$t(79) = -1.32$ $p = 0.192$ $r = 0.15$	-
Transformed data				
high cx, high lex	-			
high cx, low lex	$t(83) = 3.77$ $p < 0.001$ $r = 0.38$	-		
low cx, high lex	$t(78) = 2.15$ $p = 0.034$ $r = 0.24$	$t(80) = -1.25$ $p = 0.214$ $r = 0.14$	-	
low cx, low lex	$t(82) = 5.20$ $p < 0.001$ $r = 0.50$	$t(84) = 1.34$ $p = 0.184$ $r = 0.14$	$t(79) = 2.36$ $p = 0.021$ $r = 0.26$	-

Note. Significance at $p = 0.008$ with Bonferroni correction ($0.05 \div 6 = 0.008$). Interaction approached significance in raw data ($p = 0.068$) and did not reach significance in transformed data.

Table A36 Post-hoc paired-samples *t*-tests on typical participants' verb production latencies to the removal construction in the sentence completion task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
Raw data				
high cx, high lex	-			
high cx, low lex	$t(66) = -2.43$ $p = 0.018$ $r = 0.29$	-		
low cx, high lex	$t(64) = -2.82$ $p = 0.006$ $r = 0.33$	$t(83) = -1.53$ $p = 0.129$ $r = 0.17$	-	
low cx, low lex	$t(61) = -2.96$ $p = 0.004$ $r = 0.35$	$t(79) = -1.06$ $p = 0.313$ $r = 0.12$	$t(77) = 0.74$ $p = 0.464$ $r = 0.08$	-
Transformed data				
high cx, high lex	-			
high cx, low lex	$t(66) = 2.63$ $p = 0.011$ $r = 0.31$	-		
low cx, high lex	$t(64) = 2.89$ $p = 0.005$ $r = 0.34$	$t(83) = 1.22$ $p = 0.225$ $r = 0.13$	-	
low cx, low lex	$t(61) = 3.26$ $p = 0.002$ $r = 0.39$	$t(79) = -1.08$ $p = 0.282$ $r = 0.12$	$t(77) = -0.42$ $p = 0.676$ $r = 0.05$	-

Note. Significance at $p = 0.008$ with Bonferroni correction (0.05 ÷ 6 = 0.008).

Table A37 Post-hoc paired-samples *t*-tests on typical participants' verb production latencies to the transitive construction in the sentence completion task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
Raw data				
high cx, high lex	-			
high cx, low lex	$t(82) = 0.32$ $p = 0.750$ $r = 0.04$	-		
low cx, high lex	$t(83) = 2.04$ $p = 0.045$ $r = 0.22$	$t(82) = 1.99$ $p = 0.050$ $r = 0.21$	-	
low cx, low lex	$t(81) = 1.22$ $p = 0.225$ $r = 0.13$	$t(80) = 1.15$ $p = 0.253$ $r = 0.13$	$t(81) = -1.27$ $p = 0.207$ $r = 0.14$	-
Transformed data				
high cx, high lex	-			
high cx, low lex	$t(82) = 0.39$ $p = 0.696$ $r = 0.04$	-		
low cx, high lex	$t(83) = -2.20$ $p = 0.030$ $r = 0.23$	$t(82) = -2.66$ $p = 0.009$ $r = 0.28$	-	
low cx, low lex	$t(81) = -0.37$ $p = 0.716$ $r = 0.04$	$t(80) = -0.98$ $p = 0.330$ $r = 0.11$	$t(81) = 2.23$ $p = 0.028$ $r = 0.24$	-

Note. Significance at $p = 0.008$ with Bonferroni correction (0.05 ÷ 6 = 0.008). Interaction did not reach significance in raw data and approached significance in transformed data ($p = 0.099$).

Appendix P

Tables in this Appendix show results of Wilcoxon signed-rank tests on the number of target responses from participants with aphasia in the Phase 2 sentence completion task between conditions in each of the six constructions in the task.

Table A38 Wilcoxon signed-rank tests on number of target responses from participants with aphasia to the caused motion construction in the sentence completion task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	-			
high cx, low lex	T = 2.50 p = 0.317 r = -0.21	-		
low cx, high lex	T = 2.00 p = 0.564 r = -0.12	T = 2.00 p = 0.564 r = -0.12	-	
low cx, low lex	T = 2.00 p = 0.564 r = -0.12	T = 2.00 p = 0.564 r = -0.12	T = 1.50 p = 1.000 r = 0.00	-

Table A39 Wilcoxon signed-rank tests on number of target responses from participants with aphasia to the conative construction in the sentence completion task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	-			
high cx, low lex	T = 2.50 p = 0.317 r = -0.21	-		
low cx, high lex	T = 2.00 p = 0.564 r = -0.12	T = 0.00 p = 0.317 r = -0.21	-	
low cx, low lex	T = 0.00 p = 1.000 r = 0.00	T = 2.50 p = 0.317 r = -0.21	T = 2.00 p = 0.564 r = -0.12	-

Table A40 Wilcoxon signed-rank tests on number of target responses from participants with aphasia to the ditransitive construction in the sentence completion task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	-			
high cx, low lex	T = 1.50 p = 1.000 r = 0.	-		
low cx, high lex	T = 2.00 p = 0.564 r = -0.12	T = 0.00 p = 0.317 r = -0.21	-	
low cx, low lex	T = 1.50 p = 1.000 r = 0.00	T = 0.00 p = 1.000 r = 0.00	T = 0.00 p = 0.317 r = -0.21	-

Table A41 Wilcoxon signed-rank tests on number of target responses from participants with aphasia to the intransitive motion construction in the sentence completion task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	-			
high cx, low lex	*T* = 0.00 *p* = 0.317 *r* = -0.21	-		
low cx, high lex	*T* = 0.00 *p* = 0.317 *r* = -0.21	*T* = 1.50 *p* = 1.000 *r* = 0.00	-	
low cx, low lex	*T* = 0.00 *p* = 0.317 *r* = -0.21	*T* = 1.50 *p* = 1.000 *r* = 0.00	*T* = 0.00 *p* = 1.000 *r* = 0.00	-

Table A42 Wilcoxon signed-rank tests on number of target responses from participants with aphasia to the removal construction in the sentence completion task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	-			
high cx, low lex	*T* = 0.00 *p* = 0.317 *r* = -0.21	-		
low cx, high lex	*T* = 0.00 *p* = 0.317 *r* = -0.21	*T* = 1.50 *p* = 1.000 *r* = 0.00	-	
low cx, low lex	*T* = 0.00 *p* = 0.046 *r* = -0.43	*T* = 3.00 *p* = 0.180 *r* = -0.29	*T* = 1.00 *p* = 0.083 *r* = -0.37	-

Table A43 Wilcoxon signed-rank tests on number of target responses from participants with aphasia to the transitive construction in the sentence completion task

	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	-			
high cx, low lex	*T* = 0.00 *p* = 0.317 *r* = -0.21	-		
low cx, high lex	*T* = 0.00 *p* = 0.317 *r* = -0.21	*T* = 1.50 *p* = 1.000 *r* = 0.00	-	
low cx, low lex	*T* = 0.00 *p* = 0.083 *r* = -0.37	*T* = 0.00 *p* = 0.157 *r* = -0.30	*T* = 0.00 *p* = 0.157 *r* = -0.30	-

Appendix Q

Tables in this Appendix show results of Wilcoxon signed-rank tests on verb production latencies from participants with aphasia in the Phase 2 sentence completion task between conditions for each of the six constructions in the task.

Table A44 Wilcoxon signed-rank tests on verb production latencies from participants with aphasia to the caused motion construction in the sentence completion task

	n	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	10	-			
high cx, low lex	8	*T* = 8.00 *p* = 0.310 *r* = -0.24	-		
low cx, high lex	9	*T* = 8.00 *p* = 0.161 *r* = -0.32	*T* = 12.00 *p* = 0.735 *r* = -0.08	-	
low cx, low lex	9	*T* = 10.00 *p* = 0.263 *r* = -0.26	*T* = 14.00 *p* = 1.000 *r* = 0.00	*T* = 11.00 *p* = 0.612 *r* = -0.12	-

Table A45 Wilcoxon signed-rank tests on verb production latencies from participants with aphasia to the conative construction in the sentence completion task

	n	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	8	-			
high cx, low lex	10	*T* = 6.00 *p* = 0.176 *r* = -0.32	-		
low cx, high lex	9	*T* = 3.00 *p* = 0.063 *r* = -0.45	*T* = 21.50 *p* = 0.906 *r* = -0.03	-	
low cx, low lex	8	*T* =11.00 *p* = 0.327 *r* = -0.25	*T* = 3.00 *p* = 0.063 *r* = -0.44	*T* = 7.00 *p* = 0.237 *r* = -0.29	-

Table A46 Wilcoxon signed-rank tests on verb production latencies from participants with aphasia to the ditransitive construction in the sentence completion task

	n	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	9	-			
high cx, low lex	9	*T* = 15.00 *p* = 0.674 *r* = -0.10	-		
low cx, high lex	10	*T* = 12.00 *p* = 0.401 *r* = -0.19	*T* = 19.00 *p* = 0.678 *r* = -0.10	-	
low cx, low lex	9	*T* = 13.00 *p* = 0.484 *r* = -0.16	*T* = 18.00 *p* = 0.594 *r* = -0.13	*T* = 17.50 *p* = 0.553 *r* = -0.14	-

Table A47 Wilcoxon signed-rank tests on verb production latencies from participants with aphasia to the intransitive motion construction in the sentence completion task

	n	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	11	-			
high cx, low lex	10	*T* = 8.00 *p* = 0.047 *r* = -0.43	-		
low cx, high lex	10	*T* = 15.00 *p* = 0.374 *r* = -0.19	*T* = 19.00 *p* = 0.678 *r* = -0.09	-	
low cx, low lex	10	*T* = 22.00 *p* = 0.575 *r* = -0.12	*T* = 8.00 *p* = 0.086 *r* = -0.38	*T* = 21.00 *p* = 0.508 *r* = -0.15	-

Table A48 Wilcoxon signed-rank tests on verb production latencies from participants with aphasia to the removal construction in the sentence completion task

	n	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	11	-			
high cx, low lex	10	*T* = 9.00 *p* = 0.059 *r* = -0.41	-		
low cx, high lex	10	*T* = 13.50 *p* = 0.153 *r* = -0.31	*T* = 17.00 *p* = 0.889 *r* = -0.03	-	
low cx, low lex	7	*T* = 3.50 *p* = 0.075 *r* = -0.42	*T* = 3.00 *p* = 0.116 *r* = -0.38	*T* = 6.00 *p* = 0.176 *r* = -0.33	-

Table A49 Wilcoxon signed-rank tests on verb production latencies from participants with aphasia to the transitive construction in the sentence completion task

	n	high cx, high lex	high cx, low lex	low cx, high lex	low cx, low lex
high cx, high lex	11	-			
high cx, low lex	10	*T* = 7.00 *p* = 0.575 *r* = -0.13	-		
low cx, high lex	10	*T* = 6.00 *p* = 0.069 *r* = -0.40	*T* = 19.00 *p* = 0.285 *r* = -0.22	-	
low cx, low lex	8	*T* = 0.00 *p* = 0.327 *r* = -0.22	*T* = 3.00 *p* = 0.594 *r* = -0.11	*T* = 16.00 *p* = 0.074 *r* = -0.37	-

Appendix R

This Appendix contains results from Bayesian Standardised Difference Tests (BSDTs) comparing the difference between response times from participants with aphasia to high frequency verbs and low frequency verbs to the difference from typical participants in the sentence completion task. BSDTs were carried out on raw and transformed data.

Results are reported following the recommendations of Crawford et al. (2010). Results are provided in tabular format and include information on the value of response times to high and low frequency verbs entered into the analysis for each participant with aphasia., e.g. median raw response times and mean transformed response times; the *z*-score associated with high and low frequency items; the two-tailed *p*-value of the BSDT; the estimated percentage of the typical population expected to demonstrate a more extreme difference than the individual with aphasia, in the same direction, as both a point and interval estimate; and an estimated effect size. The effect size for a BSDT is expressed as z_{DCC}, referring to the *z* of the difference between a case and controls. This effect size reflects the number of standard deviations that a case's difference deviates from the mean difference in controls. The sign of z_{DCC} depends on the order that tasks are input into the computation; in the present study, a positive value of z_{DCC} indicates that a participant with aphasia demonstrated a higher standardised score on high frequency than low frequency items. A negative value of z_{DCC} indicates that a participant with aphasia demonstrated a higher standardised score on low frequency than high frequency items.

Table A50 Results of Bayesian Standardised Difference Tests comparing verb production latencies to verbs with high and low lexical frequency on raw data from sentence completion task

	Median response time to high frequency verbs (*n*)	Median response time to low frequency verbs (*n*)	*z*-score for high frequency verbs	*z*-score for low frequency verbs	Significance test *p*-value (two-tailed)	Estimated percentage of the typical population displaying more extreme discrepancy than PWA, in same direction		Estimated effect size (z_{DCC})	
						Point	95% CI	Point	95% CI
BF	1027 (12)	995 (12)	-0.14	-0.43	0.474	24%	17% - 32%	0.72	0.48 - 0.97
CJ	1660 (7)	3165 (8)	2.92	10.31	<0.001***	0%	0% - 0%	-18.84	-22.09 - -15.79
GW	1250 (11)	1340 (11)	0.94	1.28	0.383	19%	11% - 29%	-0.88	-1.23 - -0.55
HM	1200 (11)	1230 (10)	0.69	0.74	0.912	46%	36% - 56%	-0.11	-0.37 - 0.15
IC	1920 (11)	1820 (9)	4.18	3.66	0.225	11%	1% - 32%	1.32	0.46 - 2.19
JF	1370 (12)	1460 (11)	1.52	1.88	0.369	18%	9% - 31%	-0.92	-1.35 - -0.49
KT	1000 (12)	1020 (12)	-0.27	-0.30	0.943	47%	39% - 56%	0.07	-0.15 - 0.29
MJ	2360 (11)	2485 (8)	6.30	6.95	0.182	9%	0% - 41%	-1.64	-3.07 - -0.23
PD	4875 (10)	3740 (9)	18.46	13.16	<0.001***	0%	0% - 0%	13.52	9.72 - 17.46
RE	1255 (12)	1437 (12)	0.96	1.76	0.046*	2%	1% - 6%	-2.04	-2.52 - -1.59
SP	1120 (9)	990 (6)	0.31	-0.45	0.057ϙ	3%	1% - 6%	1.93	1.58 - 2.30

*** $p < 0.001$. * $p < 0.05$. ϙ $p < 0.10$.

Table A51 Results of Bayesian Standardised Difference Tests comparing verb production latencies to verbs with high and low lexical frequency on transformed data from sentence completion task

	Mean response time to high frequency verbs (*n*)	Mean response time to low frequency verbs (*n*)	*z*-score for high frequency verbs	*z*-score for low frequency verbs	Significance test *p*-value (two-tailed)	Estimated percentage of the typical population displaying more extreme discrepancy than PWA, in same direction		Estimated effect size (z_{DCC})	
						Point	95% CI	Point	95% CI
BF	1000 (12)	990 (12)	-0.06	0.18	0.543	27%	20% - 35%	-0.61	-0.84 - -0.38
CJ	1493 (7)	2564 (8)	-1.89	-3.47	<0.001***	0%	0% - 1%	4.17	3.32 - 5.05
GW	1282 (11)	1235 (11)	-1.28	-1.00	0.469	23%	14% - 34%	-0.73	-1.07 - -0.40
HM	1250 (11)	1282 (10)	-1.17	-1.18	0.980	49%	36% - 62%	0.03	-0.30 - 0.35
IC	1563 (11)	1852 (9)	-2.06	-2.59	0.177	9%	2% - 20%	1.40	0.84 - 1.98
JF	1316 (12)	1220 (11)	-1.39	-0.94	0.245	12%	6% - 21%	-1.18	-1.55 - -0.82
KT	971 (12)	943 (12)	0.11	0.47	0.346	17%	11% - 25%	-0.95	-1.21 - -0.69
MJ	2439 (11)	2500 (8)	-3.33	-3.41	0.847	42%	17% - 70%	0.21	-0.53 - 0.94
PD	4348 (10)	4167 (9)	-4.33	-4.35	0.963	48%	16% - 81%	0.05	-0.88 - 0.98
RE	1163 (12)	1370 (12)	-0.83	-1.47	0.099°	5%	2% - 10%	1.68	1.28 - 2.09
SP	1124 (9)	962 (6)	-0.67	0.35	0.009**	0%	0% - 1%	-2.69	-3.16 - -2.24

*** $p < 0.001$. ** $p < 0.01$. ° $p < 0.10$.

Table A52 Results of Bayesian Standardised Difference Tests comparing verb production latencies to verbs with high and low construction frequency on raw data from sentence completion task

	Median response time to high frequency verbs (*n*)	Median response time to low frequency verbs (*n*)	*z*-score for high frequency verbs	*z*-score for low frequency verbs	Significance test *p*-value (two-tailed)	Estimated percentage of the typical population displaying more extreme discrepancy than PWA, in same direction		Estimated effect size (z_{DCC})	
						Point	95% CI	Point	95% CI
BF	1017 (12)	995 (12)	-0.24	-0.38	0.716	36%	28% - 44%	0.37	0.14 - 0.59
CJ	2480 (10)	1860 (5)	7.18	3.72	<0.001***	0%	0% - 0%	9.10	7.36 - 10.93
GW	1295 (10)	1290 (12)	1.17	1.02	0.697	35%	24% - 47%	0.39	0.08 - 0.71
HM	1180 (10)	1210 (11)	0.59	0.64	0.887	44%	35% - 54%	-0.14	-0.39 - 0.11
IC	1980 (11)	1820 (9)	4.64	3.53	0.009**	0%	0% - 3%	2.92	1.95 - 3.92
JF	1300 (11)	1585 (12)	1.20	2.42	0.002**	0%	0% - 1%	-3.22	-3.89 - -2.59
KT	1285 (12)	1399 (12)	1.12	1.54	0.279	14%	7% - 24%	-1.10	-1.49 - -0.72
MJ	2220 (11)	2525 (8)	5.86	6.87	0.029*	2%	0% - 10%	-2.67	-4.09 - -1.28
PD	3660 (9)	4875 (10)	13.15	18.01	<0.001***	0%	0% - 0%	-12.79	-16.62 - -9.06
RE	1405 (12)	1311 (12)	1.73	1.12	0.117	6%	2% - 12%	1.60	1.17 - 2.04
SP	1060 (9)	1000 (6)	-0.02	-0.35	0.383	19%	13% - 27%	0.88	0.63 - 1.13

*** $p < 0.001$. ** $p < 0.01$. * $p < 0.05$.

Table A53 Results of Bayesian Standardised Difference Tests comparing verb production latencies to verbs with high and low construction frequency on transformed data from sentence completion task

	Mean response time to high frequency verbs (*n*)	Mean response time to low frequency verbs (*n*)	z-score for high frequency verbs	z-score for low frequency verbs	Significance test *p*-value (two-tailed)	Estimated percentage of the typical population displaying more extreme discrepancy than PWA, in same direction		Estimated effect size (z_{DCC})	
						Point	95% CI	Point	95% CI
BF	990 (12)	1000 (12)	0.12	0.06	0.867	43%	35% - 52%	0.17	-0.05 - 0.38
CJ	1887 (10)	2041 (5)	-2.71	-2.78	0.852	43%	21% - 66%	0.20	-0.42 - 0.81
GW	1266 (10)	1250 (12)	-1.18	-1.06	0.745	37%	26% - 50%	-0.33	-0.65 - -0.01
HM	1299 (10)	1250 (11)	-1.29	-1.06	0.522	26%	16% - 38%	-0.65	-0.99 - -0.31
IC	1754 (11)	1613 (9)	-2.47	-2.06	0.277	14%	5% - 28%	-1.13	-1.68 - -0.59
JF	1111 (11)	1449 (12)	-0.53	-1.67	0.003**	0%	0% - 1%	3.08	2.54 - 3.66
KT	1299 (12)	1316 (12)	-1.29	-1.28	0.965	48%	35% - 62%	-0.04	-0.39 - 0.30
MJ	2381 (11)	2564 (8)	-3.35	-3.33	0.960	48%	22% - 75%	-0.05	-0.78 - 0.68
PD	4000 (9)	4762 (10)	-4.35	-4.33	0.961	48%	26% - 81%	-0.05	-0.99 - 0.88
RE	1316 (12)	1205 (12)	-1.35	-0.89	0.215	11%	5% - 18%	-1.26	-1.63 - -0.90
SP	1075 (9)	1020 (6)	-0.35	-0.06	0.422	21%	15% - 29%	-0.81	-1.06 - -0.56

** $p < 0.01$.