An investigation into the feasibility of a novel, low dose clinician-delivered verb and sentence treatment programme, supplemented by self-managed home practice conducted via computer, including preliminary efficacy testing.

A thesis submitted to City, University of London for the degree of Doctor of Philosophy in the School of Health Sciences.

April 1, 2022

Julie Hickin

Division of Language and Communication Science.

Acknowledgements.

My sincere thanks to my supervisors Dr Lucy Dipper and Dr Madeline Cruice for all your hard work – your detailed and thoughtful feedback has been absolutely invaluable. I couldn't have done it without you. My sincere gratitude also to the Worshipful Company of Saddlers for the Doctoral Studentship which enabled me to achieve a lifelong ambition – to study for my PhD. It seems fitting that a journey through Higher Education that started with a National Union of Miners Welfare Fund Grant should end this way.

I am deeply indebted to the participants who so generously gave of their time and their energy to take part in the research – it was an absolute privilege to work with them. Thank you also to colleagues in the Division of Language and Communication Science who generously put me in touch with potential participants and who offered timely words of support and encouragement.

A special thank you to Professor David Howard who gave me my first research post and who has been an inspiration throughout my research career.

Very, very special thanks go to my family for their support and putting up with Mum being in PhD world – to my sons Laurie and Sam, and especially to my daughters Grace and Saskia who have been so self-sufficient – I am so proud of them. Finally, to my partner Ian - you have been my technical support and rock throughout – thank you for your faith in me. My words cannot do justice to the support you have given me.

Declaration.

The author of this thesis grants powers of discretion to the Director of Library Services to allow the thesis to be copied in whole or in part without further reference to the author. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgement.

Abstract.
This PhD study set out firstly to synthesize verb treatment literature in aphasia and secondly to explore the feasibility, acceptability, compliance, fidelity, and preliminary efficacy of a novel verb and sentence production treatment using a pre-post design with six single cases. Two reviews synthesized 37 studies of verb-in-isolation treatments (182 participants) and 33 studies of verb-in-sentence treatments (126 participants) and found these comprised primarily Level 4 evidence for treatment effect on trained items in at least 80% of participants, with varied generalization (15-59% participants depending on target), and clear preference for verb-in-sentence treatments. Dose varied widely, and fidelity was rarely assessed. The subsequent novel Sentence Production Treatment (SPT) was low dose (8 hours) and clinician delivered, supplemented by a minimum set level (16hrs) of self-managed computer-based treatment. Six participants (three males and three females aged 49 – 81 years) took part. Each chose 20 personally relevant verbs and worked on these in a series of exercises based on single verb and sentence treatments from the reviews. The SPT was found to be feasible and acceptable to participants although illness and other factors required accommodation during the study. Five were independent in using the SPT, and four complied with the minimum amount of self-delivered treatment requested. Treatment effects were noted on trained verb production and sentence production for five participants each. Generalisation to untrained verb and sentence targets and discourse was more limited, however four participants perceived functional communication improvements. The study represents the first preliminary evidence that treatment for verb and sentence production deficits self-delivered by computer can be effective. Given these overall positive findings of feasibility and benefit, further feasibility testing is warranted, exploring intervention refinement, candidacy, and a stronger research design.

Introduction.

This PhD research study is a feasibility study, with some preliminary efficacy information arising from pre-post studies of individual cases. It investigates a novel, low dose, clinician-delivered verb and sentence production treatment, which was supplemented by self-managed home practice conducted via computer, and which targeted personally relevant (PR) verbs. The design used was a pre-post treatment design, with a series of single cases. The data were collected between January 2018 and October 2019.

The study was motivated by an awareness of the limited amount of publicly funded treatment available to people with aphasia (PwA) (e.g., Palmer, Witts & Chater, 2018), and the need to address this (for example, using computer-based treatments which can be self-delivered by PwA). However, it was also motivated by an awareness of the potential that technology has to enable PwA to receive *more* treatment (e.g., Des Roches, Balachandran, Ascenso, Tripodis, & Kiran, 2015; Kurland, Liu & Stokes, 2018) which may be a key active ingredient of treatment, potentially allowing PwA to achieve greater improvement (e.g., Kleim and Jones, 2008). Indeed, developing computer-based treatments which can be self-administered at home has become even more important due to the COVID19 pandemic, with recent studies demonstrating that treatment for acquired communication disorders can be effective even when delivered entirely remotely (Braley, Sims Pierce, Saxena, De Oliveira, Taraboanta et al., 2021), and that age does not represent a barrier to complying with such treatments (Munsell, De Oliveira, Saxena, Godlove & Kiran, 2020). An additional benefit of self-delivered treatments is that they allow PwA *more autonomy* in their treatment which PwA see as important (e.g., Kurland, 2014; Kearns, Kelly & Pitt, 2019).

Whilst computer-based treatments have demonstrated efficacy in improving language skills in PwA (e.g., Mortley, Wade & Enderby, 2004; Des Roches et al, 2015; Palmer et al., 2012), this evidence is very largely based on improved production of single *nouns* which most computer-based treatments target (e.g., Hickin, Cruice & Dipper, 2020; Lavoie, Bier & Macoir, 2017), and there is limited evidence that computer-based treatments have an impact on *functional* communication (e.g. Lavoie et al., 2017; Kurland et al., 2018; Palmer et al, 2019). This is partly because studies of computer-based treatments have not routinely assessed functional communication (an omission addressed in this study), but it is also possible that treatments which target *verb* and *sentence* production (rather than noun production) may have greater potential to impact functional communication than those which target nouns. This is because verbs perform a key role in sentence production since they encode syntactic information (e.g., Garrett, 1988; Levelt, 1989). This means that verbs determine both *lexical selection* within a sentence and its *syntactic structure* (e.g Conroy, Sage & Lambon-Ralph, 2006; Webster & Whitworth, 2012). Hence, improving verb retrieval may have more impact on sentence production and consequently functional communication in comparison to improving noun retrieval. (See Chapter 1 for a more detailed discussion of this issue).

This study investigates a novel intervention for verb and sentence production deficits in aphasia: Sentence Production Treatment (SPT). The SPT includes verb-based exercises and exercises which target verb production in sentences. The use of personally relevant verbs together with generalization exercises aims to facilitate verb and sentence production in real life communication. The SPT exercises in both clinician-delivered sessions and in self-managed home practice are computer-based, with the aim of increasing the dose of treatment received. This study investigates the feasibility, acceptability, compliance with and fidelity of such a treatment, together with preliminary efficacy testing.

The SPT investigated in this study is a complex intervention as defined by the Medical Research Council (MRC) Guidance on Developing Complex Interventions (2019: https://mrc.ukri.org/documents/pdf/complex-interventions-guidance/) because it has several interacting components. The development of the SPT was, therefore, also a complex process as it drew on several different fields of aphasia research to ensure that the treatment was based on the best available evidence (also in line with the MRC Guidance on Developing Complex Interventions (ibid)). These fields were:

verb-in-isolation treatment research: a systematic scoping review of verb-in-isolation treatments was carried out to inform the first Verb Treatment phase of the SPT. It was subsequently published in February 2020 (Hickin et al., 2020) and comprises Chapter 1.

sentence treatment research: a systematic scoping review of sentence treatments was carried out to inform the second Sentence Treatment phase of the SPT. It was accepted in September 2021 (Hickin, Cruice & Dipper, 2022) and comprises Chapter 2.

a narrative review of the literature relating to the *selection of personally relevant words* for aphasia treatment was carried out to inform the process by which a set of personally relevant verbs was chosen for the SPT. This is reported in the first half of Chapter 3.

a narrative review of the literature relating to the *self-delivery of computer-based treatments* is reported in the second half of Chapter 3. This literature informed the development of the SPT in terms of both the *content* included (e.g., it confirmed the importance of personalising treatment: Kearns et al., 2019) and it informed the development of the PowerPoint slides used to deliver the SPT, ensuring that they included as many aphasia-friendly features as possible (e.g., Brandenburg, Worrall, Rodriguez & Copland, 2013).

The study set out to answer the following research questions (RQs):

1. Is it *feasible* to deliver and self-manage personalised sentence production treatment (SPT) by computer? Specifically, is it feasible to a) recruit and retain suitable participants to the SPT, b) deliver and self-manage the SPT using a computer, and c) is it feasible to select a set of PR verbs?
2. Is the SPT *acceptable* to the participants with aphasia and their significant others?

3. What factors influence *compliance* with self-managed computer treatment?
4. Were the treatment procedures that were carried out during the SPT administered with acceptable *fidelity*?
5. Preliminary efficacy testing of the effect of the SPT on the production of i) trained and untrained verbs, ii) untrained nouns, iii) sentence production using trained and untrained verbs, iv) verb and sentence production in discourse and v) in functional communication as perceived by i) the participants with aphasia themselves and ii) their significant others.

The underpinning scoping reviews and literature reviews are presented in Chapters 1-3. The design and methodology used to answer the RQs is described in Chapter 4. The results of the SPT programme in relation to each of the research questions are reported in Chapters 5 (RQs 1 -4 relating to feasibility) and 6 (RQ 5 relating to preliminary efficacy testing), and discussed in Chapter 7, with the emphasis on the feasibility aspects of the study (RQs 1 – 4).

Chapter 1. Hickin, J., Cruice, M., & Dipper, L. (2020). A systematically conducted scoping review of the evidence and fidelity of treatments for verb deficits in aphasia: verb in isolation treatments. *American Journal of Speech Language Pathology, 29,* 530-559. https://doi.org/10.1044/2019_AJSLP-CAC48-18-0234

A systematically conducted scoping review of the evidence and fidelity of treatments for verb deficits in aphasia: verb in isolation treatments.

Authors: Julie Hickin,[1] Madeline Cruice[1] and Lucy Dipper[1]

[1] Division of Language and Communication Science, City, University London, UK

Corresponding Author: Julie Hickin, Division of Language and Communication Science, City, University of London, Northampton Square, London EC1V 0HB. Email: jules.hickin.1@gmail.com

Telephone: +447496079850

Conflict of Interest: There are no relevant conflicts of interest.

Funding Statement: This research was supported by a doctoral studentship to the first author jointly funded by the Worshipful Company of Saddlers and the School of Health Sciences, City, University of London.

Purpose. Aphasia research demonstrates increasing interest in the treatment of verb retrieval deficits. This systematically conducted scoping review reports on the level and fidelity of the current evidence for verb treatments, on its effectiveness regarding the production of trained and untrained verbs, functional communication, sentences and discourse, and on the potential active ingredients. Recommendations to guide clinical decision-making and future research are made.

Method. The computerized database search included studies January 1980 to September 2018. The level of evidence of each study was documented, as was fidelity in terms of treatment delivery, enactment and receipt. Studies were also categorised according to the treatment methods used.

Results. Thirty seven studies were accepted into the review and all but one constituted a low level of evidence. Thirty three studies (89%) described treatment in sufficient detail to allow replication, dosage was poorly reported, and the fidelity of treatment was rarely assessed. The most commonly reported treatment techniques were phonological and semantic cueing in 25 (67.5%) and 20 (54%) of studies respectively. Retrieval of trained verbs improved for 80% of participants, and improvements generalised to untrained verbs for 15% of participants. There was not sufficient detail to evaluate the impact of treatment on sentence production, functional communication and discourse.

Conclusions. The evidence for verb treatments is predominantly of a low level. There are encouraging findings in terms of treatments being replicable, however

this is tempered by poor monitoring of treatment fidelity. The quality of verb treatment research would be improved by researchers reaching consensus regarding outcome measures (including generalisation to e.g. sentences and discourse), by manualising treatment to facilitate implementation and exploring the opinions of participants. Finally, whilst treatment is largely effective in improving production of trained verbs, lack of generalisation to untrained items leads to the recommendation that personally relevant verbs are prioritised.

Introduction.

Aphasia research demonstrates increasing interest in the treatment of verb production (e.g. Webster & Whitworth, 2012). This likely reflects a greater awareness of the centrality of verbs in sentence processing (e.g. Edmonds, 2016) and thus of the potential for improved verb retrieval to impact real life discourse. This review aimed to review verb in isolation treatments. Studies that explore the treatment of verbs in the context of a sentence such as Verb Network Strengthening Treatment (VNeST: e.g. Edmonds, 2016; Edmonds, Obermeyer & Kernan, 2015) and mapping treatments (e.g. Marshall, 1997; Rochon, Laird, Bose, & Scofield, 2005) are the subject of a future manuscript in preparation. The review reports on a) the level of evidence for verb in isolation treatments, b) the fidelity of the reviewed verb treatment studies in terms of treatment delivery, treatment enactment and treatment receipt as defined by Hinckley and Douglas (2013); c) the evidence for the effectiveness of treatment on the production of trained and untrained verbs, sentences, functional communication and discourse; and d) the potential active ingredients of verb treatments.

Existing reviews of verb treatments.

The existing reviews of verb treatments will be discussed briefly to demonstrate the additional contributions made by the study reported here. Conroy, Sage & Ralph (2006) reviewed the relationship between theory relating to verbs and verb treatments in aphasia. They present a very detailed and insightful overview of the relationship between theory and practice but do not consider the level of evidence for verb treatments, the fidelity of treatment or its overall effectiveness. All of these are addressed in this review.

Boyle (2017) conducted a review of semantic treatments for word and sentence production in aphasia i.e. her review was not exclusive to verb treatments as it included studies investigating semantic treatments for nouns as well. The current review adds to the existing evidence because it is restricted to verb treatments only thus allowing a more detailed consideration of the impact of treatment on verbs per se. Additionally, whilst Boyle included studies that explored semantic treatments for nouns, she excluded studies which *combined* a semantic treatment with another treatment technique (such as phonological cueing). This review encompasses not only studies which explore *single* treatment techniques (e.g. semantic treatments) but also those which *combine* treatment techniques (e.g. semantic + gesture treatment), the latter of which constitute most verb treatment studies.

Webster and Whitworth (2012) conducted a review of verb treatments which did include studies that used techniques other than semantic, those which used a combination of techniques and those which treated verbs as single words in the context of a sentence. They concluded that there is insufficient evidence to establish what type of treatment is most effective, but they did note that treatment that targeted verbs *and their arguments* appeared to result in better generalisation to sentence production than treatment that targeted single verbs. They found no clear relationship between the treatment given to participants and the deficit/s underlying their verb/sentence level difficulties and thus no evidence to advocate particular treatments for specific deficits. They state that there is a need for a more systematic approach to evaluation of the outcomes of therapy for spoken verb deficits, including evaluation of sentence production before and after treatment and evaluation of the impact on connected speech/communication in real life. Our review aims to address this need and updates Webster and Whitworth's review which only included studies published up to March 2011. In an attempt to elucidate the active ingredients of treatment there is more detailed reporting of the techniques used in treatment and of the impact of treatment (in that effect sizes are reported for individual participants whenever possible) in the current review than was reported in Webster and Whitworth. Finally, this review uniquely includes an evaluation of the level of evidence of verb treatment studies and their fidelity which has not been carried out before.

As well as the qualitative reviews discussed above, three systematic reviews of verb treatments have been conducted since 2013 (Efstratiadou, Papathanasiou, Holland, Archonti, & Hilari, 2018; Maddy, Capilouto & McComas, 2014; Rose, Raymer, Lanyon & Attard, 2013) and all of these contribute to our understanding of the effectiveness of verb treatments. However, none of these encompassed *all* researched treatments for verbs or included an evaluation of treatment fidelity. Maddy et al. (2014) and Efstratiadou et al. (2018) both conducted systematic reviews restricted to *Semantic Feature Analysis* (*SFA) treatment*, including studies that used SFA to treat nouns or discourse, as well as those which targeted verbs. The methodological quality of the reviewed studies was evaluated using the Single Case Experimental Design (SCED) Scale (Tate et al., 2008) and effect sizes were calculated (using Cohen's d and benchmarks specified by Beeson and Robey (2006)) to evaluate the effectiveness of SFA. Efstratiadou et al., and Maddy et al., rated reviewed studies highly on the SCED with an average of 9.55 and 9.3/11 respectively (range 8-11), but only small or negligible treatment effect sizes were found for the majority of participants. Rose and colleagues (2013) also conducted a systematic review this time restricted to treatments which used *gesture* either in isolation or in combination with verbal techniques. They included both group and single case studies in their review and found that whilst the quality of single case studies was high (as rated on the SCED scale), group studies were of low methodological quality (as rated on the Physiotherapy Evidence Database scale (PEDro-P) scale, see Verhagen et al.,1998). They found positive effects of treatment on verb production in over 50% of participants with indications that combined gesture and verbal treatment were more effective than verbal treatments alone. However they highlighted the need for further research in this neglected field in order to reach firmer conclusions about the effectiveness of verb treatments and hence to make clinical recommendations about the treatments of choice. The current review evaluates the evidence for *all verb in isolation* treatments with the aim of making such recommendations to guide clinical practice.

Finally, de Aguiar, Bastiaanse & Miceli (2016) conducted a meta-analysis of single case studies of verb treatments with the specific aim of identifying factors that predicted response to treatment including generalization of treatment to untrained verbs. **Predictive factors were grouped into demographic (e.g. age, gender), clinical (e.g. severity of aphasia, size of lesion) and treatment related factors (e.g. amount, intensity, type of treatment),** and Random Forests were used to assess the contribution of each factor (see de Aguiar et al., for a detailed discussion). Whilst the review provides a very valuable insight into the factors which influence response to verb treatments, it excludes case series studies (which comprise the majority of verb treatment studies). In addition, de Aguiar et al., do not evaluate the fidelity of the studies included in the meta-analysis (not an aim for their study).

In summary, the review reported here included all verb treatments that treated verbs in isolation that is, semantic treatments including but not restricted to SFA, and studies that used phonological cues, gestural cues, orthographic cues, video cues or a combination thereof. The review included studies which used group, case series or single case designs and studies which delivered verb in isolation treatments via computer, as these have not been included in any review of verb treatments to date. The degree to which verb treatment studies have attempted to evaluate the impact of treatment beyond the naming of treated and untreated verbs is systematically reviewed as recommended by Webster and Whitworth (2012). Finally this review is unique in considering the level of evidence and the fidelity of verb treatment studies.

Verb retrieval deficits in aphasia: theory and therapy.

It is beyond the scope of this paper to present a detailed review of the relationship between theories of verb retrieval deficits in aphasia and their treatment (see Conroy et al., 2006 for such a review). However, this will be discussed briefly to highlight the contributions made by current theories to verb in isolation treatments and the challenges that remain.

Verbs have been the subject of much discussion in the aphasia research literature for a number of years. For example, Berndt, Haediges, Mitchum & Sandson (1997) investigated the ability of eleven participants with aphasia to produce words of different grammatical classes and found that whilst verbs were harder to retrieve than nouns for five participants, there was no straight forward relationship between verb retrieval and aphasia type. This finding contradicted the generally held *belief* that a double dissociation exists between noun and verb impairments in aphasia: i.e. that people with nonfluent (and particularly agrammatic) aphasia show greater impairment of verbs than nouns, and the reverse is true for fluent aphasia. In contrast, a review of the *evidence* suggests that the pattern of noun and verb impairment in people with aphasia is more complicated than this. For example, Luzzatti, Raggi, Zonca, Pistarini, Contardi & Pinna (2001) compared verb and noun retrieval in 58 participants with aphasia and found that whilst people with non-fluent aphasia showed a strong tendency to have poorer verb than noun retrieval, this was not always the case, with some people with nonfluent aphasia not exhibiting a difference and some showing the reverse pattern. In fluent aphasia the pattern is more variable with the naming of verbs and of nouns almost equally liable to selective impairment.

The theoretical standpoint taken regarding the factors that underlie the differing patterns of impairment of nouns and verbs in aphasia is split in terms of whether this reflects that nouns and verbs are stored and/or processed *separately* or, that the differing patterns of impairment are an inevitable consequence of their differing psycholinguistic properties.

The psycholinguistic properties of verbs have given rise to differences in verb treatment design, in the studies reviewed here, and so merit further discussion. Firstly, verbs are thought to be both less richly represented *semantically* than nouns (i.e. they are more abstract) and to have "looser" connections with their semantic networks than nouns. This is because whilst nouns represent relatively unchanging physical entities, verbs represent actions and therefore have a temporal component as well as attributes which can vary according to context (e.g. consider “drive” in the following sentences: “The man drives a car,” “The sheep dog drove the sheep into the pen” and “The crying child drove his mother to distraction”). Thus their relationship with their semantic networks is more fluid than that of nouns. (See Black and Chiat (2003) for a detailed review.)

Verbs can also be categorized according to their *syntactic* properties, in terms of their transitivity. The fact that verbs encode syntactic information means that they have a pivotal role in sentence production. Garrett (1988) and Levelt (1989) proposed that information contained within the semantic representation of verbs was essential to sentence production because this encoded *syntactically relevant* information about predicate argument structure as well as core semantic information pertaining to a verb’s meaning. Thus being unable to retrieve the semantic representations of “buy” and “sell” for example, would entail not only difficulty retrieving information about the verbs’ core meanings (that they are change of possession verbs) but also difficulty retrieving information about their *argument structures,* and how to map these arguments onto the syntax of the sentence according to their *thematic roles.* For “buy/sell”, the lexicon contains information about argument structure and thematic roles such that both have two obligatory arguments and an optional one, and both have arguments carrying the thematic roles of theme, source and goal, but they differ as to the role that the optional argument takes (optional source for “buy” and optional goal for “sell”). The rules for mapping this information onto syntax are similarly distinct: “goal” is mapped onto the subject of the sentence for “buy” but onto a prepositional phrase for “sell”. Several studies have indeed attested to the destructive effect of impaired semantic representations of verbs on sentence production (e.g. Jones 1986; Marshall, Chiat & Pring, 1997). That verbs encode additional information (in comparison to nouns) which is essential for sentence production makes them potentially very fruitful targets for treatment in that improving the retrieval of a verb should have a greater impact on improving sentence production than improving the retrieval of a noun (or indeed other word classes such as adjectives). In other words, verb retrieval treatments have greater potential to *generalize* beyond improving single word production to sentence production (across level generalization).

Whilst verbs store information at the lexical level which is essential to sentence production, perhaps surprisingly this can also affect their production *in isolation*. For example, Thompson, Lange, Schneider & Shapiro (1997) found that the complexity of a verb's argument structure influenced verb production in both a single word and sentence context for ten participants with agrammatic aphasia. This finding was replicated by Kim and Thompson (2000) whose seven agrammatic participants demonstrated a hierarchy of difficulty between one, two and three place verbs in a single word context for both production (naming) and comprehension (categorization) tasks.

Verbs are also more complex *morphologically* (e.g. being marked for person and time – for instance, look/looks/looked). There are also *phonological* features that could make English verbs vulnerable to impairment: verbs tend to be shorter in duration than nouns, have fewer syllables and tend to carry their stress on the second rather than the first syllable making them less salient. And finally, there are other psycholinguistic properties that distinguish verbs from nouns, including relative imageability, age-of-acquisition and frequency. The combination of these differences between nouns and verbs may account for their differential impairment in aphasia.

The differences between nouns and verbs may mean that verbs are harder to process both cognitively and linguistically than nouns (a theory borne out by verbs being harder to acquire developmentally) and this may also account for their relative neglect in the treatment literature because it may have fostered the belief that they may be both *harder to treat* and *less responsive* to treatment than nouns. (e.g. Conroy, Sage and Lambon-Ralph, 2009b and c). However despite this, the treatment of verb retrieval deficits has attracted increased attention in the aphasia treatment literature in recent years as demonstrated by a notable increase in the number of studies published (for example, all but three of the thirty-seven studies included in the current review were published after the millennium). As well as trying to establish the effectiveness of verb retrieval treatments, studies have tried to elucidate how verb treatments work – the active ingredients and mechanisms of treatment – in order to develop a "theory of (verb) therapy" (Byng, Nickels & Black, 1994). Regarding so-called active ingredients of treatment these have been defined as:

> "a behavior-influencing procedure shown through experimental analysis to affect a specific behavior and that is indivisible in the sense that removing any of its components would render it inert" (Embry & Biglan, 2008, p. 1573).

Whilst active ingredients may be easy to identify in pharmacological treatments (e.g., a particular antibiotic or analgesic), this is often not the case for behavioral treatments such as those applied in the treatment of aphasia. Indeed, most aphasia treatments are likely to be complex interventions as defined by the Medical Research Council in that they "contain several interacting components" (www.mrc.ac.uk/complexinterventionsguidance, p.7). Thus identifying the active ingredients of aphasia treatment is likely to be a difficult process but this has not stopped researchers attempting it. So for example, studies have attempted to adapt SFA treatment (applied initially to nouns) for verbs by adapting the features generated during treatment to be more appropriate (or active) for verbs (e.g. Wambaugh Mauszycki & Wright, 2014; Wambaugh & Ferguson, 2007). In SFA treatment, the likely active ingredient of treatment is the *generation of semantic features* for a target word (see Gravier et al., 2018). Thus SFA for nouns includes generating the semantic feature of category whereas in SFA adapted for verbs, features unique to verbs (and therefore potentially active ingredients of treatment for them) are generated (e.g. the instrument of a verb). Because it is hypothesized that SFA strengthens the semantic network of treated verbs, it is predicted that treatment should generalize to semantically related verbs (within level generalisation) but not to semantically unrelated verbs (in line with the prediction for SFA targeting nouns). However, to date most of the participants in SFA studies have not demonstrated such generalization.

Other verb treatment studies have used phonological and orthographic cues cue to try to improve verb retrieval (e.g. Conroy, Sage & Lambon-Ralph, 2009a, b & c), once again borrowing from the evidence base for anomia treatment. Because it is hypothesized that these treatments will work on a lexical basis (i.e. that they will strengthen the link between semantics and the phonological form of a verb), it is predicted that treatment will not therefore generalize to untrained verbs, and indeed Conroy et al. found word retrieval improvements almost entirely restricted to treated verbs only in their series of studies (Conroy et al., 2009a, b and c).

Finally, some verb treatment studies have investigated techniques which are designed to target unique qualities of (action) verbs i.e. that the lexical representations of an action verb in the brain may be intimately linked with the representation of the sensory motor features which encode its actions. This leads to the prediction that gesture will be an effective treatment for spoken production of verbs and this has been investigated in a number of studies (see Rose et al. (2013) for a review). Very recently, studies have also investigated whether the observation of an action alone (or in combination with gesture) can facilitate the spoken production of the related verb based on the belief that this will also activate the sensory motor representation of a verb and hence facilitate retrieval of its spoken form. Initial results have proved to be promising, including for action observation alone (e.g. Marangolo et al., 2012). Treatments are also starting to exploit emerging research which suggests that using a dynamic depiction of a verb (i.e. a video) rather than a static (picture) representation of a verb as a stimulus may be more effective in evoking action verb production (e.g. Blankestijn-Wilmsen et al., 2017).

In summary, whilst theory relating to verb deficits in aphasia is beginning to explicitly inform the development of verb treatments there remain many gaps in our understanding of how verb treatments work, how to facilitate generalization of treatment effects beyond the production of trained verbs in isolation and what the treatment of choice should be for any one individual with a verb deficit. Indeed as Conroy et al. (2006) concluded, there is disappointingly little evidence that our understanding of how verbs and nouns differ from each other has informed treatment. Thus for example, theory predicts that verbs are more difficult to process than nouns (because they are less imageable and less phonologically salient, but more complex syntactically) and that they are thus likely require to more processing resources during communication and *during treatment*. However, this has not been taken into account in planning or delivering treatments (e.g. by using errorless learning techniques such as those investigated in a series of studies of the treatment of anomia (e.g. Fillingham, Sage & Lambon-Ralph, 2005).

Fidelity of verb treatment studies.

The fidelity of verb treatment studies has not been reported on to date. With regard to the fidelity of aphasia treatment more generally, Hinckley and Douglas (2013) and Kaderavek and Justice (2010) found that this had been poorly monitored. Hinckley and Douglas reviewed aphasia treatment studies published in the previous ten years and found that only 14% of 149 studies reviewed assessed treatment fidelity (usually by rating a sample of videotaped treatment sessions). They recommended that three levels of treatment fidelity needed to be addressed to improve the quality of aphasia therapy research namely, *treatment delivery* (e.g. by the use of treatment manuals and training), *treatment receipt* (e.g. by the use of homework record sheets and establishing the views of recipients regarding their treatment), and *treatment enactment* (e.g. by observation of treatment delivery). Kaderavek and Justice (2010) argued that the neglect of treatment fidelity has undermined the implementation of evidence-based practice because, for example, treatments are not described in enough detail to enable replication in clinical practice. The fidelity of the studies included in this review is reported according to the three levels recommended by Hinckley and Douglas (2013).

Review Methodology.

A scoping review paradigm was used for this study. Scoping reviews are used to categorise existing literature in terms of its nature, features, and volume; and are best used when a body of literature exhibits a large, complex, or heterogeneous nature not amenable to a more precise systematic review (Dijkers, 2015; Peters, Godfrey, Khalil, McInerney, Parker, & Soares, 2015). Verb treatments are complex in that the method and content varies considerably, as do the outcome measures and methods used to analyse the effectiveness of treatment. The review followed the five steps recommended by Dijkers (ibid). These were firstly to identify the research questions, and secondly to search for relevant papers via established methods (a database search). The third step was to select papers pertinent to the research questions and the fourth to chart the collected data. The final, fifth step was to collate, summarize and report the results of the scoping review. The review was conducted using systematic procedures to ensure that these were rigorous, explicit and replicable.

The research questions that the review set out to answer were:

1. What is the level of evidence for verb in isolation treatments?
2. What is the fidelity of the research in terms of treatment delivery, receipt, and enactment (Hinckley & Douglas, 2013)?
3. What is the evidence of positive gains for treatment in terms of improved production of a) trained verbs in isolation and in sentences and b) untrained verbs in isolation and in sentences (within and across level generalization)?
4. What is the evidence of positive gains for verb production treatments in terms of a) improved functional communication and b) improved production of discourse?
5. What are the potential active ingredients of verb treatments?

The evaluation and charting of the reviewed studies in terms of a) the level of evidence of a study and b) treatment fidelity was guided by established hierarchies of evidence for healthcare research (e.g. the highly influential Cochrane hierarchy), and by the small treatment fidelity literature pertaining to aphasia treatment research respectively. With regard to level of evidence, hierarchies of evidence used to guide evidence-based healthcare generally place Systematic Reviews (SRs) of Randomised Controlled Trials (RCTs) at the top of the hierarchy, with single case studies placed at the bottom. Intermediate levels of evidence are placed in the following, descending, order: RCTs, pseudoRCTs, nonrandomized group studies which include a control group alongside the treated group, followed by nonrandomized group studies where treatment effects are determined by comparison with previous performance/historical data (e.g. case series).

Method.

Cinahl Complete and Medline Complete databases were searched using the following terms:

Verb in isolation treatment Searches:

Face-to-face delivery of verb treatments:

verb and *aphasia* and *treatment* or *therapy*

Computer delivery of verb treatments:

verb and *aphasia* and *technology* or *computer*

Sentence treatment Searches:

Face-to-face delivery of sentence treatments:

sentence or *sentence production* and *aphasia* and *treatment* or *therapy*

Computer delivery of sentence treatments:

sentence and *aphasia* and *technology* or *computer*

Studies which were original research and which were published in peer-reviewed journals, in English from 1980 up to September 2018 were considered for inclusion. The titles of all papers identified in the database search were read. If the purpose of the study was not clear from the title, the abstract was read to determine whether the study should be included. All papers accepted into the review were then read in full and the contents charted including the aim of the study, type of design, number of participants, the type, amount, intensity and duration of treatment, the results of treatment and any limitations of the study. The reference lists of all papers were scrutinized to identify any additional relevant studies. Where possible higher-level evidence was included, but the literature identified almost exclusively represented lower levels of evidence, that is single cases and case series designs.

As there is no pre-established protocol for evaluating fidelity, a tailored approach was developed to evaluate the studies in this review. This was based on the elements of aphasia treatment needed for a 'theory of therapy' (e.g. Byng et al., 1994), to determine the optimal dose of treatment (e.g. Baker, 2012a; 2012b), and the active ingredients of treatment or 'kernels of knowledge' (Embry & Biglan, 2008). The following data were extracted from the reviewed studies, and used to judge fidelity: 1) the amount of treatment given (at least the number of sessions); 2) the number of verbs treated; 3) the type of stimulus (e.g. picture vs video); 4) the treatment hierarchy; 5) the modality of cues; and 6) the contingency under which cues were given and how many times they were given. The impact of treatments on trained and untrained verb production in isolation and in sentences was charted, as was the impact of treatment on functional communication and discourse. The results of treatment for individual participants are described whenever these are reported. When a study reported results for the group only, this is reported separately i.e. the participants in these studies were not added to the totals for individual participants because it was not clear how many participants in the group had (or had not) benefitted from treatment.

Results.

The results of the four searches carried out for the systematically conducted scoping review are summarized in Figure 1, and are reported according to PRISMA guidelines (Moher, Liberati, Tetzlaff & Altman, 2009). The combined searches resulted in 331 papers, with another nine papers identified from additional sources. Once duplications were removed, the titles and abstracts of 178 papers were screened. As a result of this screening, 141 papers were excluded (see Figure 1). Studies that were excluded investigated other types of aphasia treatment (e.g. anomia treatment, dysgraphia treatment, conversation training) or investigated spoken verb production treatment but were:

a) published in a language other than English,

b) had participants with another form of aphasia (e.g. progressive aphasia),

c) were without at least one pre and post treatment measure of verb retrieval (e.g. when the main aim of the study was to investigate the neurophysiological response to treatment),

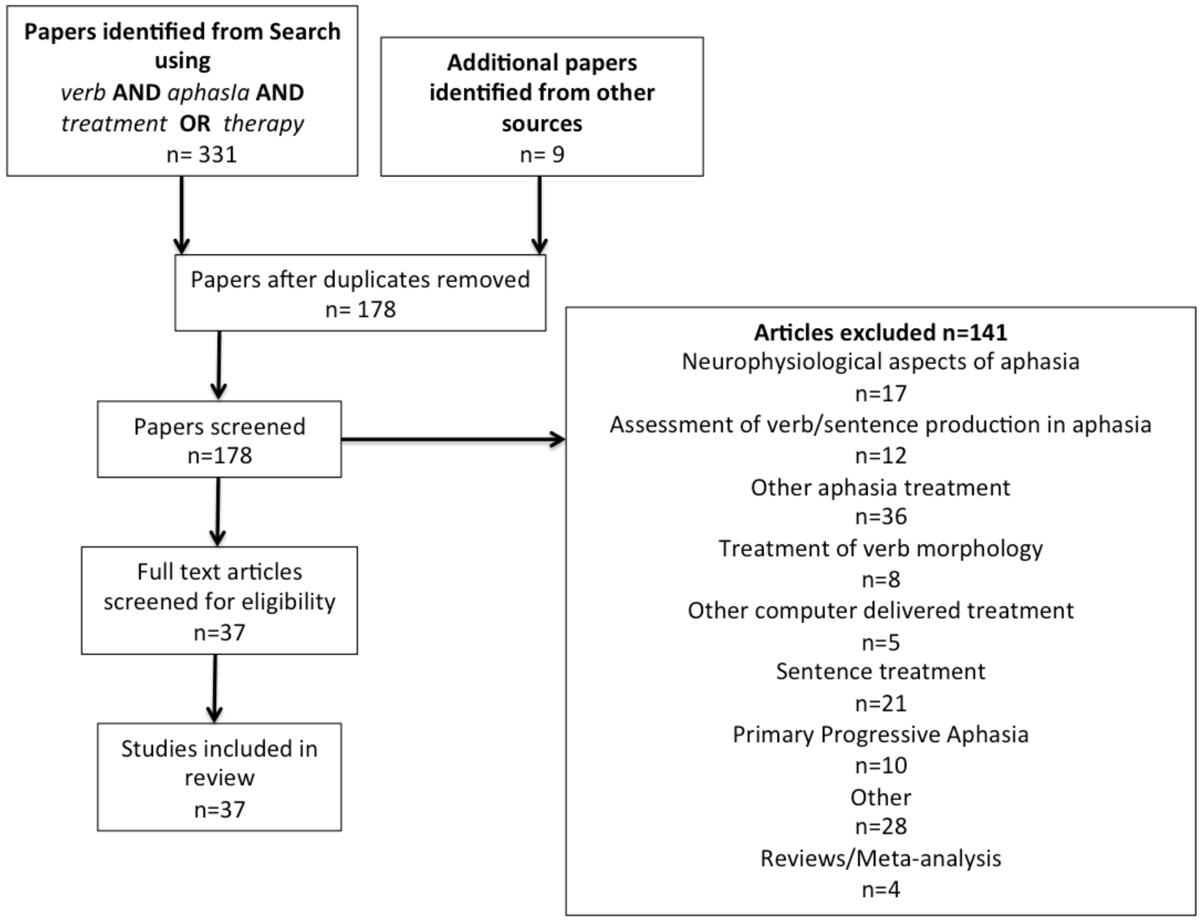
d) used techniques which focused on verbs in sentences (i.e. verbs with their arguments) rather than verb in isolation treatments,

e) targeted correct production of verb morphology rather than correct production of the verb itself,

f) studies whose primary focus was assessment of aphasia, and

g) studies which were reviews or meta-analyses rather than original research.

Figure 1. Results of the systematically conducted scoping review.



Following the exclusion of 141 papers the full text of 37 papers was screened and all papers were included in the subsequent review. The details of the included papers are summarized in Table 1. Thirty two of the reviewed papers reported verbs treatment delivered face-to-face, whilst five reported treatments delivered via computer. Results are reported according to research questions.

1. **What is the level of evidence for verb production treatments?**

The majority of studies comprised lower levels of evidence because they were case series or single case studies. Specifically, verb treatment studies comprised 22 case series, seven case series which also reported group results, six single-case studies, one group study (Marangolo et al., 2012) and one pilot RCT (Palmer et al., 2012). The maximum number of participants in a study was 15 (Carragher et al., 2013; Conroy et al., 2009c; Palmer et al., 2012; Raymer et al., 2006), and the total number of participants in the reviewed studies was 182 (see Table 1, column 2).

Table 1. Studies of verb treatments accepted into the scoping review.

1.Paper	2.Study Design	3.Study Aim	4.Amount of Treatment	5.Number of Verbs Treated	6.Type of treatment
Adrian, Gonzalez, Buiza & Sage (2011)	case series + group results n=15	To investigate the effectiveness of a computer delivered treatment program (CARP-2).	30 hours over 4 months	25 Verbs (+175 Nouns)	**C**omputer-assisted **A**nomia **R**ehabilitation **P**rogram (**CARP**): comprehension tasks; semantic, phonological, orthographic & sentence closure cues
Bonifazi, Tomaiuolo, Altoè, Ceravolo, Provinciali & Marangolo (2013)	case series n=6 + group results	To assess the effectiveness of action observation treatment in improving verb retrieval.	30-45 min sessions, 4 times a day, for 4 days a week, over 2 weeks (i.e. 16-24 hours in total).	52 - 120 in 4 equal size sets.	Tx Step 1: Observation of action being performed; 2: observation of action being performed + execution of action 3: observation of action on video 4: (control condition) observation of action being performed + execution of meaningless gesture
Boo & Rose (2011)	case series n=2	To compare the relative effectiveness of semantic only, repetition or gesture only and semantic plus gesture treatment.	10 sessions of 1-2 hours delivered x2-3 weekly (min 20 hours - max 60 hours)	100 (4 sets of 20) + 20 control)	Tx 1: repetition only, 2: semantic cues only, 3: gesture cues only, 4: semantic + gesture
Carragher, Sage & Conroy (2013)	case series + group results n=9	To assess the efficacy of a multi component verb retrieval treatment on a) retrieval of trained, untrained verbs and semantically light verbs, b) sentence production and c) verb retrieval in natural conversation.	1 hour week x 8 weeks plus homework	40 (including 5 light & 5 personally relevant verbs)	repetition + semantic cues (SFA) + gesture cues
Conroy & Scowcroft (2012)	case series n=4	To assess the impact of a decreasing cue treatment on noun and verb retrieval using a dynamic number of verbs (i.e. the number of verbs treated was	10 weekly sessions plus homework	Dynamic (potential variation from 10 -100; 50 verbs and 50 nouns)	orthographic and phonological decreasing cue hierarchy

		determined by how they responded during treatment).			
Conroy, Sage & Lambon-Ralph (2009a)	case series + group results n=7	To compare the effectiveness of treating verb retrieval using a single word cue versus in a sentence context.	5 weeks of Tx, x2 weekly for 40-50 minute sessions. Specified that participants produced each target verb at least 100 times during the Tx program (min 7 hours – max 8 hours)	40 (in 2 sets of 20) plus 20 controls	decreasing cue hierarchy, using phonological and orthographic cues in either a single word or sentence context
Conroy, Sage & Lambon-Ralph (2009b)	case series + group results n=7	To compare the effectiveness of increasing vs decreasing cues in word retrieval therapy for nouns and verbs.	5 weeks of Tx, x2 weekly for 40-50 minute sessions). Specified that the participants produced each target verb at least 100 times during the Tx program (min 7 hrs – max 8 hrs)	40 Verbs and 40 Nouns (in 4 sets of 20) plus 20 controls	decreasing or increasing cue hierarchy, using phonological and orthographic cues
Conroy, Sage & Lambon-Ralph (2009c)	case series + group results n=9	To compare the effectiveness of errorless and errorful learning on noun and verb retrieval.	5 weeks of Tx, x2 weekly for 40-50 minute sessions). Specified that the participants produced each target verb at least 100 times during the Tx program (min 7 hrs – max 8 hrs)	40 Verbs and 40 Nouns (4 sets of 20) plus 20 controls	errorless condition used decreasing phonological and orthographic cues; errorful used increasing semantic, phonological and orthographic cues
Edwards & Tucker (2006)	case series n=3	To establish the effectiveness of a verb retrieval treatment with 3 people with fluent aphasia.	2 months (JR) or 4 months (JD; CB) x2 weekly sessions of 45 minutes (min 12- max 19 hours of treatment)	50	sentence completion, naming to definition, picture naming, semantic then progressive phonemic cues as required
Faroqi-Shah & Graham (2011)	case series n=2	To explore the effect of verb class relatedness on the effectiveness of treatment (i.e. generalization).	not specified	14 (plus 14 controls)	naming from video, semantic feature generation, semantic feature judgement task, sentence construction

Fink, Schwartz, Sobel & Myers (1997)	case series n=5	To investigate the effectiveness of multimodal treatment for action naming.	6 sessions	5 (+5 exposed for naming only) 20 controls	sentence prompt, comprehension task, gesture cue
Knoph, Lind & Simonsen (2015)	single case study	To explore the effectiveness of SFA with a quadrilingual speaker.	c10 hours per week over two and a half weeks (c22 hours,)	44 (plus 34 controls)	SFA
Kristensson, Behrns & Saldert (2015)	case series n=3	To explore the impact of SFA on conversation & functional communication.	20 x 1 hour sessions over 5-6 weeks	Not specified but "everyday" nouns & verbs (84 Nouns and Verbs used as untrained controls)	SFA
Kurland, Wilkins & Stokes (2014)	case series n=5	To establish the effectiveness of a home treatment for noun and verb retrieval delivered via iPad .	participants reported practicing for an average 8 hours 26 mins per month for 6 months (range 4 hrs 23 mins- 11 hrs 11 mins; total range 26 hrs 18 mins – 67 hrs 6 mins)	40 verbs (+ 40 nouns)	Video stimulus, video of articulation of target word, semantic, phonological, orthographic cues, comprehension (semantic) tasks
Marangolo, Cipollari, Fiori, Razzano & Caltagirone (2012)	group study n=7	To explore whether action observation is effective for human vs non-human actions	daily sessions over 2 consecutive weeks	115 verbs (78 human, 37 nonhuman)	observation of video clips of verbs, no verbal cues given
Marangolo, Bonifazi, Tomaiuolo, Craighero, Coccia, Altoè et al., (2010)	case series n=6	To assess whether observation of verb action only is an effective treatment	30-45 mins x3 daily for 2 weeks (estimated 15 - 22.5 hours total)	11 -31 verbs in 4 sets (3 Tx sets and 1 control) i.e. 44 -124 verbs in total	observation of therapist carrying out action of a verb, then either 1) say verb 2) gesture verb and say it or 3) produce meaningless gesture and say verb
Marshall, Pring & Chiat (1998)	single case study	To investigate treatment for a selective deficit in retrieving phonological representations of verbs.	24 hrs over 14 weeks	35	comprehension tasks (semantic judgements) paired with reading aloud followed by verb generation task

McCann & Doleman (2011)	case series n=3	To replicate the therapy carried out by Edwards & Tucker (2006) with nonfluent participants	30-60 minute session x2 weekly for 6-9 weeks (until criteria reached) 14 – 18 sessions (min 7 hours - maximum of 18 hours)	100 in two sets of 50	see Edwards & Tucker (2006)
McNeil, Doyle, Spencer, Jackson Goda, Flores & Small (1998)	single case study	To compare the effectiveness of cueing treatment on words of different grammatical classes.	6-8 1 hour sessions per word class	60 in lists of 10 (mixed verbs, nouns, adjectives + prepositions)	L -SAIT (lexical semantic activation inhibition treatment) requiring involves production of a antonym or of a synonym for the target verb
Mortley, Wade & Enderby (2004)	case series n=6	To assess the efficacy of noun and verb retrieval treatment delivered remotely by computer using StepByStep.	6 months duration (2 x 3 month phases) average home practice per month: 12 hrs 23 mins (range 7 hrs 48 mins – 15 hrs 27 mins; total range 46hrs 48 mins – 92hrs 43 mins).	100 verbs (+162 nouns)	StepByStep: video of articulation of target words, semantic, phonological, orthographic cues, comprehension tasks
Palmer et al., (2012)	RCT n=15	To assess the feasibility of an RCT investigating self administered computer treatment for noun and verb retrieval (StepByStep).	20 mins per day, 3 days a week for 5 months: c25 hours in total (10/15 participants completed to recommended dose)	96 nouns/verbs (48 of personal relevance)	as above
Raymer, Ciampitti , Holliway, Singletary, Blonder et al., (2007)	case series n=8	To assess the effectiveness of semantic-phonological treatment on noun and verb retrieval.	2 x 10 sessions 2-4 times a week	20 verbs and 20 nouns	repetition and Y/N questions about semantic and phonological attributes of target words

Raymer, Singletary, Rodriguez, Ciampitti, Heilman, & Gonzalez Rothi (2006)	case series n=9	To assess the effectiveness of gesture-verbal (phonological) treatment on noun and verb retrieval.	2 x 10 sessions	20 nouns and 20 verbs	repetition and gesture
Raymer & Kohen (2006)	case series n=2	To compare the effect of a word retrieval treatment in the context of the sentence on noun vs verb retrieval.	max 10 sessions (to criterion)	20 verbs and 20 nouns	written and spoken sentence cue, reading aloud
Raymer , Rodriguez, Ciampitti, Singletary, Fuller et al., (2005)	case series n=7	To investigate the effectiveness of pantomime training on noun and verb retrieval.	10 sessions, 3-4 times per week.	not reported	pantomime training
Raymer & Ellsworth (2002)	single case study	To compare the effectiveness of phonological, semantic and rehearsal (repetition) on verb retrieval and sentence production	unclear	60 (in 3 sets of 20)	Y/N questions about semantic or phonological attributes of target words or repetition and silent rehearsal
Rochon & Reichman (2003)	single case study	To investigate the effectiveness of verb retrieval and sentence treatment on verb vs sentence production.	7 x 1hr sessions delivered twice weekly (verb treatment)	11 plus 12 controls	written then spoken production of verb, plus written cue if required
Rodriguez, Raymer & Gonzalez-Rothi (2006)	case series n=4	To compare the effectiveness of semantic-phonological and gesture-verbal	1 hr session, 2-3 times a week for 1 month (8-12 hours treatment)	40 (in 2 sets of 20) plus 20 controls	repetition followed by 2 Y/N questions re the semantics or phonology of the V respectively , or

		treatments on verb retrieval.			repetition plus gesture of V
Rose & Sussmilch (2008)	case series n=3	To compare the relative effectiveness of semantic-phonological and gesture-verbal treatments.	20 x c1 hour sessions, 3 times per week	80 (in 4 sets of 20) plus 20 controls	repetition; SFA; gesture; SFA + gesture
Routhier, Bier & Macoir (2016)	case series n=2	To assess the effectiveness of semantic-phonological treatment for verb anomia self-delivered by a tablet.	Participants reported practicing 4 x week for 5 weeks; sessions lasted 45-75 mins mean c60 mins (total c20 hours)	25 Verbs for P1 & 31 Verbs for P2	video verb cue, sentence prime, phoneme/syllable cue, whole word written cue, whole word spoken cue
Routhier, Bier & Macoir (2015)	case series n=2	To compare the effectiveness of semantic-phonological cueing and action-observation with action-observation alone on verb retrieval.	9 sessions of c90 mins, over 3 weeks (c13.5 hours)	74 (37 per set)	naming from an action video, + semantic & phonological cues ending in repetition in treatment condition
Schneider & Thompson (2003)	case series n=7	To compare the effectiveness of semantic verb retrieval treatment and verb argument structure retrieval treatment.	12 sessions (verb treatment)	20 (+ matched controls)	semantic cue (definition) plus repetition if required
Takizawa, Nishida, Ikemoto, & Kurauchi (2015)	case series + group results n=6	To assess the relative effectiveness of single word vs sentence therapy.	1-5, 40-minute sessions per week over, 2–8 months	30 (2 sets of 15) (40 in 2 sets of 20 for 1P)	semantic and phonological cueing with repetition if required
Wambaugh Mauszycki & Wright (2014)	case series n=4	To explore the effectiveness of SFA on action naming.	to 90% criterion or max 12 sessions.	20 (2 sets of 10) plus 20 controls	SFA
Wambaugh & Ferguson, (2007)	single case study	To explore the effectiveness of SFA on action naming.	maximum 12 sessions (to criterion), 45-60	40 (4 sets of 10)	SFA plus repetition if required

			minutes, x3 per week (max 12 hours)		
Wambaugh, Cameron, Kalinyak-Fliszar, Nessler & Wright (2004)	case series n=5	To compare the effectiveness of semantic cueing treatment (SCT) and phonological cueing treatment (PCT)	maximum of 15 sessions (to criterion), of 25 -45 minutes (max 11 hours)	40 (4 sets of 10)	semantic or phonological cueing hierarchy, both including sentence closure and repetition if required
Wambaugh, Doyle, Martinez, & Kalinyak-Fliszar (2002)	case series n=3	To compare the effectiveness of PCT and SCT when applied to verbs as opposed to nouns.	maximum of 20 sessions (to criterion) of 60 minutes (max 20 hours)	24 (2 sets of 12) plus 24 controls	as above

Abbreviations: Tx= treatment; SFA = semantic feature analysis; ES = effect size; 1P = 1 participant.
⁺Results reported for treated language only.

2. What is the fidelity of the research in terms of treatment delivery, receipt, and enactment?

In terms of *treatment delivery,* none of the studies reported the existence of a published manual for the treatment investigated[1]. However, most studies (33 or 89%) reported the treatment *procedure* in sufficient detail to enable replication (excluding Fink, Schwartz, Sobel & Myers, 1997; McNeil et al., 1998; Palmer et al., 2012; Raymer et al., 2005). In particular, some studies gave a detailed description of the treatment protocol in an appendix (e.g. Boo & Rose, 2011; Wambaugh, Doyle, Martinez, & Kalinyak-Fliszar, 2002).

[1] However the StepByStep® therapy software programme used in the studies by Mortley et al. (2004) and Palmer et al. (2012) is commercially available and does have an instruction manual.

Regarding *dose,* only eight studies (22%) reported the exact amount of treatment given (see Table 1, column 4): 30 hours over 4 months (Adrian, Gonzalez, Buiza & Sage, 2011), 20 hours over 5-6 weeks (Kristensson, Behrns & Saldert, 2015), 24 hours over 14 weeks (Marshall, Pring, & Chiat, 1998), 25 hours over 5 months (Palmer et al., 2012), 10 hours 2 - 4 times a week (Raymer et al., 2006), and 7 hours twice weekly (Rochon & Reichman, 2003). Kurland et al. (2014) and Mortley et al. (2004) were able to report the amount of treatment given in their studies in precise detail because this was monitored by the computer programs used in their self-delivered treatments (26-67 total hours in Kurland et al.'s study and 46–93 hours in Mortley et al.). Edwards and Tucker (2006) state the exact amount for two of three participants (17 and 19 hours). Carragher, Sage and Conroy (2013) gave their participants 8 hours of treatment plus an unspecified amount of homework. Three other studies reported an approximate amount: between 13.5 hours and 22 hours (Knoph, Lind & Simonsen, 2015; Rose & Sussmilch, 2008; Routhier, Bier & Macoir, 2015). Six studies reported the number of treatment sessions given but not their length: 6 (Fink et al., 1997); 10 each (Conroy & Scowcroft, 2012; Raymer & Kohen, 2006; Raymer et al., 2005 and 2007); and 12 (Schneider & Thompson, 2003). Ten studies reported the minimum and maximum length of sessions as well as the number of sessions which allowed the estimation of the minimum and maximum total amount of treatment (Bonifazi et al.,; 2013; Boo & Rose, 2011; Conroy, Sage & Lambon-Ralph, 2009 a, b and c; Marangolo et al., 2010; McCann & Doleman, 2011; McNeil et al., 1998; Rodriguez, Raymer & Gonzalez-Rothi, 2006; Routhier, Bier & Macoir, 2016). Wambaugh and colleagues reported a series of studies where the amount of treatment was dependent on reaching a predetermined performance criterion, meaning amount of treatment given differed (Wambaugh et al., 2014; Wambaugh & Ferguson, 2007; Wambaugh, Cameron, Kalinyak-Fliszar, Nessler & Wright, 2004; Wambaugh, Doyle, Martinez, & Kalinyak-Fliszar, 2002). Takizawa, Nishida, Ikemoto, & Kurauchi (2015) reported huge variation in the duration and intensity of treatment given in their clinical study: one to five 40-minute sessions per week over a span of 2–8 months. Three studies did not report the amount of treatment given (Faroqi-Shah & Graham, 2011; Marangolo et al., 2012; Raymer & Ellsworth, 2002).

Moving on to *treatment receipt*, the only studies that reported the *views* of participants about their treatment were those investigating self-delivered treatments via the computer program StepByStep. Thus Mortley and colleagues report the views of the participants in their 2004 study in a companion paper (Wade, Mortley and Enderby, 2003), and Palmer and colleagues (2012) report their participants' views in Palmer, Enderby and Paterson (2013). Both studies used structured interviews and thematic framework analysis to investigate the views of the participants with aphasia and their carers, with the focus being on the *acceptability* of self-delivered computer treatment and any perceived *advantages* and *disadvantages* of treatment delivered in this way rather than on the verb treatment itself. Thus there is a significant gap in the evidence base regarding verb treatment receipt. Two studies attempted to monitor the amount of *home practice* carried out in addition to face-to-face treatment: Carragher and colleagues (2013) reported a range in homework hours completed during their study from 0.5 to 23.75 hours, and Conroy and Scowcroft (2012) mentioned home practice, but did not report it because it was not recorded accurately by their participants.

In terms of *treatment enactment*, only studies delivering face-to-face treatments (n=32) were considered. (Treatment fidelity of self-delivered computer treatments is an emerging area of research: for example Ball, de Riesthal & Steele (2018) investigated the degree to which participants complied with recommended treatment procedures during self-administered (computer) anomia treatment, and whether adherence influenced accuracy of performance.) In terms of face-to-face treatment, five studies (16%) reported that the fidelity of treatment enactment was assessed (by rating how closely the treatment protocol was followed using a sample of videotaped treatment sessions). These studies were: Carragher et al., (2013), Faroqi-Shah and Graham (2011), Rose & Sussmilch (2008), Wambaugh & Ferguson (2007) and Schneider & Thompson (2003). Whilst technically not treatment enactment, it was encouraging to see that twelve studies (37.5%) addressed *assessment* fidelity, most commonly by establishing inter-rater reliability using another assessor who was sometimes independent of the study and/or blind to the pre/post-treatment status of the assessment.

3. What is the evidence of positive gains for treatment in terms of improved production of a) trained verbs in isolation and in sentences and b) untrained verbs in isolation and in sentences (within and across level generalization)?

The evidence for treatment effectiveness on the production of i) treated and ii) untreated verbs, and on sentences using iii) treated and iv) untreated verbs is summarized in Table 2 Columns 4 - 7. Significance levels and effect sizes are given for each individual participant when available.

Table 2. Individual participant results per study and treatment type.

1.Study	2.Study Aim	3.Type of Treatment	4.Trained Verbs in Isolation	5.UnTrained Verbs in Isolation	6.Trained Verbs in Sentences	7.UnTrained Verbs in Sentences	8.Functional Communication	9.Discourse
Adrian, Gonzalez, Buiza & Sage (2011) *(n=15)*	*To assess the efficacy of a multi component (**combination**) treatment delivered by **computer**. (CARP-2)*	**C**omputer-assisted **A**nomia **R**ehabilitation **P**rogramme (**CARP**): comprehension tasks; semantic, phonological, orthographic & sentence closure cues (5)	*Unclear because nouns and verbs analysed together. However numerical improvements for verbs (>20%) for 10/15 participants[1]*	*Unclear because nouns and verbs analysed together.*	*not assessed*	*not assessed*	*not assessed*	*not assessed*
Bonifazi, Tomaiuolo, Altoè, Ceravolo, Provinciali & Marangolo (2013) (n=6)	To **compare** the relative effectiveness of 3 treatment conditions (all involving **action observation**).	1) Observation of action being performed; 2) observation of action being performed + execution of action 3) observation of action on video	**the 4 phonologically impaired Ps improved significantly** after all 3 treatments ($\chi 2$ p****); no significant improvement for the 2 semantically impaired participants	not assessed	not assessed	not assessed	not assessed	not assessed

Boo & Rose (2011) (n=2)	To **compare** the relative effectiveness of 4 treatment conditions.	1) semantic only, 2) repetition only 3) gesture only and 4) semantic plus gesture	**Participant GF:** treatments 1, 3 & 4 **significant** with small effect sizes **Participant PF**: treatments 1, 2 & 4 **significant** with small effect sizes for 1 & 4, and medium for 2.	not significant for either participant (Chi Square)	not assessed	**Participant GF:** not assessed **Participant PF**: qualitative changes reported	**Participant GF**: his LCQ indicated a significant improvement (McNemar Test, significance level not reported) but his carer's LCQ showed no change **Participant PF**: his LCQ showed no changes (no data for his carer)	**both participants showed improvement in proportion of verbs.**

Carragher, Sage & Conroy (2013) (n=9)	*To assess the efficacy of a multi component (**combination**) treatment.*	*repetition + semantic cues (SFA) + gesture cues*	***8/9 Participants** had a significant effect of treatment (McNemar p<.05) (not significant for **DC**)*	***Five participants (JH, AT, PM, PG, DM**) showed **significant improvement** (McNemar, p* -**) Four participants (KK, BL, GL, DC), showed no significant improvement*	*not assessed*	***2 participants** showed **significant improvement**: **AT**: Wilcoxon, p**; **KK:** Wilcoxon, p*); 4 participants showed no significant change (JH, BL, GL, DM); 3 participants, showed significant decrease (Wilcoxon: DC: p* PG: p* PM: p**)*	*not assessed*	*no significant improvements reported*
Conroy & Scowcroft (2012) (n=4)	*To assess the impact of a decreasing (**combination**) cue treatment using a dynamic number of verbs.*	*orthographic and phonological decreasing cue hierarchy*	***all 4 participants significant improvement** (McNemar, p***)*	*no participant showed significant improvement (in that Txd verbs improved more than untreated (χ2 p*)*	*not assessed*	*not assessed*	*not assessed*	*not assessed*
Conroy, Sage & Lambon-Ralph (2009a) (n=7)	*To **compare** the effectiveness of two (**combination**) treatment conditions.*	*orthographic and phonological decreasing cue hierarchy in either a 1) single word context or 2) sentence context*	***all 7 participants improved significantly** after each treatment (McNemar p**)*	***1 participant improved significantly (JT** (McNemar p*),*	*not assessed*	*not assessed*	*not assessed*	*not assessed*

Conroy, Sage & Lambon-Ralph (2009b) (n=7)	*To **compare** the effectiveness of two (**combination**) treatment conditions*	*orthographic and phonological cues in either a 1) decreasing or 2) increasing cue hierarchy*	***all 7 participants** improved significantly after each treatment (McNemar p**)*	*No significant improvement for any participant*	*not assessed*	*not assessed*	*not assessed*	*not assessed*
Conroy, Sage & Lambon-Ralph (2009c) (n=9)	*To **compare** the effectiveness of two (**combination**) treatment conditions.*	*1) an errorless condition using a decreasing orthographic and phonological cue hierarchy or 2) an errorful condition using an increasing semantic, orthographic and phonological cue hierarchy*	***all 9 participants improved significantly** after each treatment (McNemar p***)*	*no significant improvement for any participant*	*not assessed*	*not assessed*	*not assessed*	*not assessed*
Edwards & Tucker (2006) (n=3)	*To establish the effectiveness of a verb retrieval (**combination**) treatment with 3 people with fluent aphasia.*	*sentence completion, naming to definition, picture naming, semantic then progressive phonemic cues as required*	***2/3 participants (JD & CB) improved significantly** (McNemar p*)*	***2/3 participants improved significantly** (**JD**: Wilcoxon p*; **CB:** Wilcoxon p***)*	*not assessed*	***1/3 participants improved significantly** (**JD**: Wilcoxon p**)*	*not assessed*	*no significant changes*
Faroqi-Shah & Graham (2011) (n=2)	*To explore the effect of verb class relatedness on the effectiveness of (**combination**) treatment (i.e. within level generalization).*	*naming from video, semantic feature generation, semantic feature judgement task, sentence construction*	***1/2 participants improved significantly (P1** McNemar p*)*	***both participants improved significantly** (McNemar p*)*	*not assessed*	*not assessed*	*not assessed*	*not assessed*

Fink, Schwartz, Sobel & Myers (1997) (n=5)	*To investigate the effectiveness of multimodal (**combination**) treatment for action naming.*	*Sentence prompt, comprehension task, gesture cue*	***significant improvement** for the group *****	*none*	*not assessed*	*not assessed*	*not assessed*	*not assessed*
Knoph, Lind & Simonsen (2015) (n=1)	To explore the effectiveness of **SFA treatment** with a quadrilingual speaker.	SFA	**Participant improved significantly** (McNemar p***) **ES d =10.01**	None	not assessed	not assessed	not assessed	**significant improvements reported in** verb production in a narrative (v**erb types: ES d = 1.98); verb tokens (d = 1.54**)
Kristensson, Behrns & Saldert (2015) (n=3)	To explore the impact of **SFA treatment** on conversation & functional communication.	SFA	not assessed	no significant improvements	not assessed	not assessed	no significant improvements reported by participants or carers	no significant improvements

Kurland, Wilkins & Stokes (2014) *(n=5)*	*To assess the efficacy of a multi component (**combination**) treatment delivered by **computer***	*Video stimulus, video of articulation of target word, semantic, phonological, orthographic cues, comprehension (semantic) tasks (6)*	***significant improvement for all 5 Ps:** NWS & PBS: **small ES (d=4.0**); MCR &SSM: **medium ES (d =7.0**); ACL **large ES (d=10.10**)*	*no significant improvements*	*not assessed*	*not assessed*	*not assessed (anecdotal reports of improvement)*	*not assessed*
Marangolo, Cipollari, Fiori, Razzano & Caltagirone (2012) (n=7)	To **compare** the relative effectiveness of 2 treatment conditions (all involving action observation).	observation of video clips of 1) human actions and 2) non human actions; no verbal cues given	**significant improvement for the group** for treatment 1 only ****	**significant improvement** for **1/7** participants for treatment 1 only**	not assessed	not assessed	not assessed	**significant improvement** in picture description for 6/7 participants***
Marangolo, Bonifazi, Tomaiuolo, Craighero, Coccia, Altoè et al., (2010) (n=6)	To **compare** the relative effectiveness of 3 treatment conditions (all involving **action observation**).	action observation plus 1) say verb, 2) gesture verb and say it or 3) produce meaningless gesture and say verb	**all 4 nonfluent (phonologically impaired) Ps improved significant** (logistical model p* - ***) after treatment 1 and 2; no significant improvement for 2 fluent (semantically impaired) Ps for any treatment	not assessed	not assessed	not assessed	not assessed	not assessed
Marshall, Pring, & Chiat (1998) *(n=1)*	*To investigate (**combination**) treatment for a selective deficit in retrieving phonological representations of verbs.*	*comprehension tasks (semantic judgements) paired with reading aloud followed by verb generation task*	***Participant improved significantly** (McNemar, p**)*	*no significant improvement.*	***significant improvement** (McNemar, p***)*	***significant improvement** (McNemar, p*)*	*not assessed*	*not assessed*

McCann & Doleman (2011) (n=3)	*To replicate the (**combination**) treatment carried out by Edwards & Tucker (2006) with nonfluent participants.*	*see Edwards & Tucker (2006)*	***all 3 participants improved significantly** (**P1:** McNemar, p*; **P2**: McNemar p*; **P3**: McNemar p***)*	*not assessed*	*not assessed*	***only 1/3 participants improved significantly** (P2: McNemar p*)*	*not assessed*	*not assessed*
McNeil, Doyle, Spencer, Jackson Goda, Flores & Small (1998) (n=1)	To **compare** the effectiveness of **(a single**) treatment on words of different grammatical classes.	L -SAIT (lexical semantic activation inhibition treatment) requiring production of an antonym or of a synonym for the target verb	**Significant improvement** (reported via visual inspection)	None	not assessed	not assessed	not assessed	not assessed

Mortley, Wade & Enderby (2004) *(n=6)*	*To assess the efficacy of a multi component (**combination**) treatment delivered remotely by **computer**.* (StepByStep).	*StepByStep: video of articulation of target words, semantic, phonological, orthographic cues, comprehension tasks (5)*	**Significant improvement for all 6 participants**: (z scores): JW***; BD***; BH***; TW***; MJW***; GL***	*not assessed*	*not assessed*	*not assessed*	*Not formally assessed but interviews with carers and participants report positive benefits (see Wade et al., 2003)*	*Not assessed*
Palmer et al., (2012) *(n=15)*	*To assess the **feasibility** of an RCT investigating a multi component (**combination**) treatment delivered remotely by **computer**. (StepByStep).*	*as above (5)*	*19.8% (95% CI, 4.4%–35.2%; p= 0.014) mean improvement in change in percentage of **all** words named correctly for the treated group.* *Specific effect on verbs unclear because nouns and verbs analysed together*	*not assessed*	*not assessed*	*not assessed*	*Not assessed but interviews with carers and participants report positive benefits (see Palmer et al., 2013)*	*Not assessed*
Raymer, Ciampitti , Holliway, Singletary, Blonder et al., (2007) (n=8)	*To assess the effectiveness of (**combination**) semantic-phonological treatment on noun and verb retrieval.*	*repetition and Y/N questions about semantic and phonological attributes of target words*	***5/8 participants improved significantly** (**P1, P2, P3, P4 & P5:** C statistic, p*-**) **ESs: P1: d=17.34; P2: d=11.12; P3: d=4.61; P4: d=10.64; P5: d=2.56***	*no significant changes*	*not assessed*	*not assessed*	*positive changes reported for group (no individual results reported).*	*no significant changes*

Raymer, Singletary, Rodriguez, Ciampitti, Heilman, & Gonzalez Rothi (2006) (n=9)	*To assess the effectiveness of (**combination**) gesture-verbal (phonological) treatment on noun and verb retrieval.*	*repetition and gesture*	***4/9 participants improved significantly ESs (P1: d=8.61, P2: d=3.42, P7: d=4.14, P9: d=2.61**)*	*none reported*	*not assessed*	*not assessed*	*not assessed*	*not assessed*
Raymer & Kohen (2006) (n=2)	*To **compare** the effect of (**combination**) treatment in the context of the sentence on noun vs verb retrieval.*	*written and spoken sentence cue, reading aloud*	***1/2 participants improved significantly** (**P2**: **d=25.5**)*	***1/2 participants improved significantly (P2: d=2.84**)*	*No significant improvement*	*No significant improvement*	*Not assessed (Improvement reported anecdotally for P1)*	*not assessed*
Raymer , Rodriguez, Ciampitti, Singletary, Fuller et al., (2005) (n=7)	To investigate the effectiveness of **pantomime training** on noun and verb retrieval.	pantomime training	**5/7 participants improved significantly** (C Statistic p**; **ES d>2.5**)	no significant improvement reported	not assessed	not assessed	not assessed	not assessed

Raymer & Ellsworth, (2002) (n=1)	To **compare** the relative effectiveness of 3 (**single**) treatment conditions.	1) semantic treatment 2) phonological treatment 3) repetition treatment	**Significant improvement for all 3 treatments** (1): McNemar, p***; 2): McNemar, p***: 3): McNemar, p*)	**Significant improvement** (at the end of all 3 treatments) (McNemar, p*)	**Significant improvement for all 3 treatments** (1): McNemar, p*; 2): McNemar, p**: 3): McNemar, p*)	not assessed	not assessed	not assessed
Rochon & Reichman (2003) (n=1)	*To **compare** the relative effectiveness of 2 (**combination**) treatment conditions.*	*Written then spoken production of verb, plus written cue if required 1) verb in isolation treatment 2) sentence level treatment (results for treatment 1) only reported here)*	***yes** but only percentage improvement reported*	*not significant*	*not significant*	*not significant*	*not assessed*	*not significant*
Rodriguez, Raymer & Gonzalez-Rothi (2006) (n=4)	*To **compare** the relative effectiveness of 2 (**combination**) treatment conditions.*	*1) semantic-phonological treatment = repetition + 2 Y/N questions re the semantics or phonology of the V respectively, 2) gesture-verbal treatment = repetition plus gesture of V*	***improvement for 1/4 participants (P1 ESs: treatment 1): d =9.92; treatment 2): d = 3.34)***	*not significant*	*not assessed*	*not assessed*	*not assessed*	*not assessed*

Rose & Sussmilch (2008) (n=3)	To **compare** the relative effectiveness of 4 treatment conditions .	1) SFA; 2) *SFA + gesture*; 3) gesture only 4) repetition only	**2/3 participants Improved significantly KC**: treatment 1) **d=4.2**; treatment 2) **d=8.2**; treatment 3) **d= 8.1** **MW:** treatment 1) **d=5.91**; treatment 2) **d=5.84**; treatment 4) **d= 3.43**	**2/3 participants Improved significantly** after all 3 treatments (**KC**: McNemar, <p**; **MW**: McNemar, p<**)	not assessed	not significant	**all 3 participants Improved significantly** on the LCQ (KC & MW = spouse report; MW & MT = self report) (McNemar Test, p**)	**2/3 participants** showed increase increase in the number of verbs used (no statistical analysis)
Routhier, Bier & Macoir (2016) *(n=2)*	*To assess the efficacy of a multi component (**combination**) treatment delivered by **computer**.*	*video verb cue, sentence prime, phoneme/syllable cue, whole word written cue, whole word spoken cue (5)*	***Significant improvement for both Ps**: P1***, **large ES** (d=30.0);, P2*, **small ES** (d=2.36)*	***Partial generalisation** for P1; none for P2*	*not assessed*	*not assessed*	*not assessed*	*not assessed*
Routhier Bier & Macoir (2015) (n=2)	To **compare** the relative effectiveness of 2 treatment conditions (both involving **action observation**).	1) *Naming from an action video (action observation) + semantic & phonological cues* 2) action observation alone	**yes for both participants for treatment 1 only** (P1: Wilcoxon, p***; P2: p***) **ESs: P1: d= 9.9; P2: d=3.3**	not significant	not assessed	not assessed	not assessed	not assessed

Schneider & Thompson (2003) (n=7)	To **compare** the relative effectiveness of 2 treatment conditions.	1) semantic verb retrieval = semantic cue (definition) plus repetition if required treatment and 2) verb argument structure retrieval treatment (treatment 1 only reported here)	**significant improvement** for the group***	None	**significant improvement** for the group***	Significant improvement for the group*	not assessed	Narrative showed some improvements in grammatical sentences & production of verb arguments: for group not statistically significant
Takizawa, Nishida, Ikemoto, & Kurauchi (2015) (n=6)	*To **compare** the relative effectiveness of 2 (**combination**) treatment conditions.*	*1) verb in isolation treatment (semantic and phonological cueing with repetition if required) 2) sentence treatment (only treatment 1 reported here)*	***Improved verb retrieval for all participants with treatment 1)*** *(significance levels not reported)*	***only 1 participant improved significantly*** *(**P6**: χ2, p*)*	*not assessed*	*not assessed*	*not assessed*	***2/6 participants showed significant improvement in MLU*** *[**P1**: t(73) = −2.953, p = .004, **P6**: t(39) = −2.141,*
Wambaugh Mauszycki &Wright (2014) (n=4)	To explore the effectiveness of **SFA** on action naming.	SFA	**Significant improvement for 3/4 participants. ES**s: **P1: d=7.0 & 10.78; P2: d=10.69-17.0; P4: d=7.7 & 4.43)**	None	not assessed	not assessed	not assessed	**Improvement in CIUs for P1** (no stats)
Wambaugh & Ferguson, (2007) (n=1)	To explore the effectiveness of **SFA** on action naming.	SFA (plus repetition if required)	**Significant improvemen**t (Set 1: **d=1.5;** Set 2: **d=1.72** representing medium effect sizes (ESs calculated using Auerbach et al., (1995) d-index)	None	not assessed	not assessed	not assessed	**Improvements in CIUs and WPM**

Wambaugh, Cameron, Kalinyak-Fliszar, Nessler & Wright (2004) (n=5)	To **compare** the relative effectiveness of 2 (**single**) treatment conditions.	1) semantic cueing Treatment (SCT) 2) phonological cueing treatment (PCT), both including sentence closure and repetition if required	**improvement reported for P1 and P2 only** (reached 90% naming accuracy plus visual inspection of graphs)	None	not assessed	not assessed	not assessed	not assessed
Wambaugh, Doyle, Martinez, & Kalinyak-Fliszar (2002) (n=3)	To **compare** the relative effectiveness of 2 (**single**) treatment conditions.	as above	**improvements reported for treatment 1** for **both participants** (**P1 & P3**) and for **treatment 2** for **1/2 (P3)** (90% correct criterion)	**partial for 1/3 (P1 after treatment 1)**	not assessed	not assessed	not assessed	not assessed
Percentage (raw number) of participants: significant change reported			**80% (104)**	**!5% (18)**	**40% (2)**	**26% (6)**	**50% (4)**	**31% (15)**
Percentage (raw number) of participants: no significant change reported			**20% (26)**	**85% (101)**	**60% (3)**	**74% (17)**	**50% (4)**	**69% (33)**
Total number of participants individual results reported			**100% (130)**	**100% (119)**	**100% (5)**	**100% (23)**	**100% (8)**	**100% (48)**

Note. LCQ = LaTrobe Communication Questionnaire; SFA = semantic feature analysis; CI = confidence interval; Y/N = yes/no.
Italics indicate study investigated treatment that involved a combination of cues (21 of the 37 studies, 57%).
[1] Participant MLM omitted because improvement was from 1/25 to 2/25.
*p<.05; **p<.01; ***p<.001; ****p<.0001

Encouragingly all but one of the 37 reviewed studies reported significant improvement for *at least one individual* or for the group of participants (Table 2, column 4). (The remaining study Kristenssen et al., (2015) did not assess treated verbs). This represents improvement in trained verb retrieval for 104 of the 130 (80%) participants for whom individual results are reported. Thirty one studies (84%) reported improvement in treated verbs via inferential statistics (reporting either probability levels ($p<.05$) and/or effect sizes), and four (11%) reported improvements according to pre-established performance criteria, with three also using visual inspection of graphs of performance before, during and after treatment. In terms of effect sizes for trained verbs, when reported these were predominantly small in size (reported on 17 occasions across all studies – see Table 2 column 4), with medium effect sizes reported on 11 occasions and large effect sizes on 9. Whilst there is debate about the best way to analyse improvement in single case studies and case series (see e.g. Howard, Best & Nickels, 2015), the finding that 84% of studies used statistical analysis to assess the effectiveness of treatment adds to the rigour of the evidence for verb treatments.

In terms of (ii) untrained verbs (Table 2, column 5), 31 studies investigated this, with 11 (30%) of these studies reporting significant improvement in untrained verbs following verb treatment. This represented significant improvement for 18 out of 119 participants (15%) for whom individual results are reported.

Regarding the impact of verb treatment on *sentence production* involving either (iii) trained verbs or (iv) untrained verbs, interpretation of the evidence is hampered by inconsistent assessment (Table 2, columns 6 and 7). For trained verbs, only five studies (13%) assessed sentence production. Four of these studies (with a total of five particiants) report *individual* results, finding significant improvement for two of the five participants (Marshall et al., 1998; Raymer & Ellsworth, 2002), and no significant improvement for the remaining three participants (Raymer & Kohen, 2006; Rochon & Reichman, 2003). Schneider & Thompson (2003) report significant improvement in sentence production using trained verbs for the *group* (n=7). For untrained verbs, nine studies (24%) assessed sentence production: five studies reported significant improvement for six of the 23 participants (26%) for whom individual results are reported; Schneider & Thompson (2003) again report significant improvement for their group study (n=7); and three studies found no significant improvement.

4. What is the evidence of positive gains for verb production treatments in terms of a) improved functional communication and b) improved production of discourse?

The evidence concerning the impact of verb treatment on functional communication and discourse is also limited by lack of assessment (see Table 2, columns 8 and 9). Only four studies (11%) investigated the impact of verb treatment on functional communication. Three of these reported effectiveness as measured by the LaTrobe Communication Questionnaire (LCQ: Douglas, O'Flaherty & Snow, 2000) or the Communicative Effectiveness Index (CETI: Lomas et al., 1989). These studies were Boo and Rose (2011), Raymer et al., (2007), and Rose and Sussmilch (2008). This represented self-reported improvement for four of the five participants for whom individual results are reported, with spouses also reporting improvements for two of these participants. Raymer et al. report improvements in CETI ratings for their group as a whole (n=8) as rated by participants' carers. Kristenssen et al. (2015) found no effect of SFA treatment on functional communication for their three participants, as measured by the Communication Outcome after Stroke scale (COAST: Long, Hesketh, Paszek, Booth, & Bowen, 2008) given to both participants and carers.

Twelve studies (32%) investigated the impact of treatment on discourse. The outcome measures used to assess discourse were varied ranging from complex picture description to analysis of narrative production and conversation (see Table 2, column 9). Seven studies reported a significant effect of treatment. This represented improvement for 15 of the 48 (31%) participants for whom individual results were reported.

5. What are the potential active ingredients of verb treatments?

The treatment techniques (i.e. potential active ingredients) reported in the reviewed studies are summarized in column 6 of Table 1. The most commonly reported technique was the use of *phonological cues* (including repetition, initial phoneme and rhyme cues) which was reported in 25 (67.5%) of studies.[2] The next most common technique reported was *semantic cueing* reported in 20 studies (54%) including those using SFA. The use of *orthographic cues* was reported in 13 (35%) of studies as were *comprehension tasks*. *Gesture cues* were reported in 10 (31%) of studies and *sentence closure* in 8 (22%). The use of *video (verb) stimuli* was reported in 5 (13.5%) of studies. The least commonly reported techniques (reported in 3 (8%) of studies respectively) were 1) *construction of a sentence* using a treated verb, 2) *action observation* and 3) video cues (of the articulation) of a target verb.

The frequency with which a treatment technique is reported does not *per se* indicate its potency as an active ingredient, that is, we cannot assume that the most frequently reported techniques are necessarily the most effective. Indeed, approximately half of the reviewed studies (17

[2] When a study used more than one type of cue (i.e. a combination treatment) it is counted as a study for each type of cue used (e.g. Raymer et al. (2007) would be counted as a study using semantic cues and as a study using phonological cues).

or 46%) attempted to establish the active status of treatment techniques by using them in isolation and *comparing* their effectiveness (*comparison* studies) whilst 20 studies (54%) assessed the effectiveness of one treatment *(treatment evaluation* studies).

Given the lack of certainty regarding the active ingredients of verb treatment, it is perhaps unsurprising that most of the studies in the review investigated treatment which used a *combination of cues*: 21 of the 37 studies (57%) investigated treatment which involved a combination of cues (these are highlighted by italics in Table 2). For example, Edwards and Tucker (2006) and McCann & Doleman (2011) used sentence completion, naming to definition, semantic and progressive phonemic cues in their clinical studies. Carragher et al., (2013) used repetition, SFA and gesture cues, whilst Marshall et al., (1998) used comprehension tasks paired with reading aloud followed by a verb generation task in their single case study. All of these studies reported significant benefits to the participants. It is interesting to note that of the eighteen participants for whom generalization of treatment to untrained verb production is reported, sixteen of these received a combination treatment. (Wambaugh et al., (2002) report partial generalization for one of their 3 participants after (single) semantic cueing treatment and Marangolo et al. (2012) report generalization for one of their seven participants following action observation treatment).

All five of the studies evaluating verb treatments *delivered by computer* also used a combination of cues with the finding that 23 of the 28 participants (82%) (for whom individual results were reported) demonstrating improved retrieval of treated verbs (Adrian et al., 2011; Kurland et al., 2014; Mortley et al., 2004; Routhier et al., 2016). Additionally, significant improvement was reported for the group as a whole (n=15) in the pilot RCT conducted by Palmer and colleagues (Palmer et al., 2012). These findings are particularly impressive given that in all but one of these five studies (Adrian et al., 2011), treatment was self-administered.

Comparison studies of treatments utilising a single type of cue compared the use of: semantic cues with phonological cues (e.g. Raymer & Ellsworth, 2002; Wambaugh et al., 2004; Wambaugh et al., 2002;), semantic, phonological and gestural cues (e.g. Boo & Rose, 2011; Rodriguez et al., 2006); a word versus a sentence cue (Conroy et al., 2009a; Raymer & Kohen, 2006) and increasing versus decreasing cues (Conroy et al., 2009b and c). An emerging area of research is the comparison of action observation alone versus observation accompanied by execution of the action as a treatment technique, with initial findings being promising (Bonifazi et al., 2013; Marangolo et al., 2012; Marangolo et al., 2010).

None of these studies was able to reach a clear conclusion as to whether any type of cue is more effective than another. However, there are indications that participants whose verb retrieval deficit is semantic (rather than phonological) in nature may be less responsive to treatment (Bonifiazi et al., 2013; Marangolo et al., 2010; Rodriguez et al., 2006; Wambaugh et al., 2004), and participants with a more severe deficit may also be less responsive (Conroy et al., 2009c; Palmer et al., 2012).

In summary, the review found that the evidence for verb treatments is currently at a low level. In terms of the fidelity of verb treatment studies, whilst the vast majority of studies reported treatment in sufficient detail to enable replication, the fidelity of studies was poor. In particular, the dose of treatment was not accurately reported and treatment receipt was almost entirely neglected in terms of the views of participants about their treatment. Regarding the effectiveness of verb treatments, they resulted in improvements in the production of treated verbs in isolation for 80% of participants. However, generalization of treatment effects to untrained verbs occurred for only 15% of participants. The impact of verb treatments on sentence production, functional communication and discourse could not be evaluated because these were not consistently assessed in the reviewed studies. Regarding the active ingredients of verb treatments, it was not possible to identify these, although the review highlighted potentially active ingredients which merit further investigation such as action observation.

Discussion.

The existing body of evidence for verbs in isolation treatments is almost entirely of a low level, with the exception of one pilot RCT and one group study. This represents a challenge to researchers to conduct studies using designs which constitute higher levels of evidence such as well-designed, larger scale RCTs (see, for example, Leff & Howard, 2012). Before progressing with larger scale studies however, researchers need to be cognizant of the current phase of verb treatment research as defined by Robey and Schultz (1998). In terms of Robey and Schultz's model, verb treatment research is very predominantly at Phase I and II in that it is still seeking to establish therapeutic effects, refine treatment protocols and establish optimal dosages for example. Researchers must therefore be careful not to "put the cart before the horse" and ensure that they are trialing treatments where the effective ingredients have been clearly established (see discussion below in relation to research question 5).

Regarding the *fidelity of verb treatment* research, there are also challenges for researchers particularly in terms of the reporting of treatment dose, the gathering of data on treatment receipt and the monitoring of treatment enactment. Taking *treatment delivery* first, there were encouraging findings in that treatment protocols were almost always described in sufficient detail to enable replication. However, because this detail was most commonly contained within the body of the paper (e.g. within the Method section) rather than in an appendix or supplement, it was not easy to extract. This may be an issue for busy clinicians who might wish to implement a treatment in their practice and for researchers conducting replication studies. It is therefore recommended that treatment protocols are given in the appendices of papers (e.g. Boo & Rose, 20111; Wambaugh et al.,

2002) and that this is clearly flagged to facilitate easy extraction of treatment protocols and thus potentially their implementation into clinical practice. This would also represent a step towards the development of treatment manuals which state explicitly how treatment should be delivered and which can therefore enhance the fidelity of treatment delivery.

Reporting the *dose* of treatment is vital to ensure that treatment is delivered to the recommended amount and yet only eight studies reported the exact treatment schedule. To address this, it is recommended that the minimum detail studies should report is dose, dose form, dose frequency, session duration, and total intervention duration (as recommended by Baker, 2012a & 2012b). Treatments delivered by computer have the potential to report these data in detail but only two of the five studies which reported treatment delivered by computer had the technological capacity to do this (Mortley et al., 2004; Palmer et al., 2012). Reporting of this detail would facilitate accurate replication of dose and also enable progress towards identifying the optimal dose of treatment. Dosages reported in the reviewed studies varied hugely (from a total of 7 to nearly 93 hours of treatment, delivered over an estimated duration of a minimum of 2 weeks to a maximum of 8 months). Studies delivering a relatively small amount of treatment non intensively (e.g. Carragher et al., 2013: 8 hours over 8 weeks) reported significant benefit to participants as well as those delivering large amounts of treatment (e.g. Mortley et al., 2004: c46 hours - c93 hours over 6 months). Mortley et al.'s treatment was self-delivered by computer and treatments delivered in this way clearly have the potential to increase the dosage delivered significantly. However the amount of treatment received in the four studies using this delivery mode still varied hugely between participants (from 20 hours delivered over five weeks in Routhier et al. (2016), 25 and 26 hours delivered over 5 and 6 months in Kurland et al. (2014) and Palmer et al. (2012) respectively, to c46 – c93 hours in Mortley et al. (2004) as described above). Thus it is currently not possible to evaluate the potential efficacy of intensively delivered verb treatments or indeed to identify the optimal dose of treatment. This has implications not only in terms of the efficacy and efficiency with which treatment is delivered but also for the compliance of participants. Brady et al., (2016) conducted a Cochrane review of aphasia treatment and found that whilst treatment delivered to a high dose at a high intensity could be beneficial, this was confounded by a higher drop-out rate of participants in comparison to treatment of a lower dose and intensity. Determining the optimal dose for verb treatments is vital at many levels.

Turning now to *treatment receipt*, only two studies of verb treatment investigated the views of the participants on their treatment (Wade et al., 2003; Palmer et al., 2013) and this largely focused on the participants' views of the *mode* of treatment delivery (self-delivered via computer) rather than the treatment itself. This is a very significant gap in the evidence base but the pioneering studies by Wade et al. and Palmer et al. represent a potential way forward here in their use of structured interviews and thematic framework analysis to investigate the views of both people with aphasia and their carers on treatment.

Finally in terms of the fidelity of treatment*, treatment enactment* was monitored in only 16% of studies, and the way in which this was done varied considerably. For example, Carragher et al., (2103) state that they "regularly" reviewed an unspecified number of videoed taped sessions (p. 858), whilst adherence to the treatment protocol was rated for approximately 50% of sessions in Faroqi-Shah and Graham's (2011) study versus 17% for Wambaugh and Ferguson (2007) albeit with approximately 100% adherence found in both studies. As well as routinely monitoring treatment enactment, it is therefore recommended that a standard protocol to do this be developed. Apropos of this, Kaderavek & Justice (2010) regard direct observation as the gold standard for assessing the fidelity of treatment enactment and give an example of a Fidelity Coding Catalog for use during direct observation (Kaderavek & Justice, 2013, Appendix A, p.377), as well as making a number of recommendations to assess and enhance treatment fidelity in clinical practice (ibid, Table 1, p.375).

Turning now to treatment efficacy, this review found that the production of treated verbs in isolation improved for 80% of participants (for whom individual results were reported). This encouraging finding is slightly in excess of that of the meta-analysis carried out by de Aguiar et al. (2016) who found that treated verbs improved for 76.1% of participants. This may partly be accounted for by the inclusion of studies of self-delivered treatment via computer (n=4). These studies reported improvement in trained verb production for 100% of participants which bodes well for the development of more treatments delivered in this way. Optimism must however be tempered by the as yet small number of participants for whom (individual) data are available (n=13).

The efficacy of verb treatments is also tempered by lack of generalization of treatment to untrained verbs which only occurred for 15% of participants (similar to 14.5% in de Aguiar et al. (ibid)). The lack of generalization of verbs treatments is perplexing in that verbs are proposed to have looser semantic networks than nouns. Verbs tend not to be tied to particular actors or objects but instead constrain their semantic network through their argument structure, which is expressed in looser semantic terms (e.g. the looser notions of 'agent' and 'theme' rather than specific concepts such as 'teacher' and 'pen'). This should facilitate generalization at least to semantically related verbs, such as all verbs that have an 'agent' and a 'theme' (see e.g. Boyle, 2017). Given that 54% of studies incorporated semantic cues in treatment the lack of generalization is again hard to explain. There are several possibilities. Boyle (ibid) speculates that SFA treatment may have resulted in disappointing generalization because "the current lack of agreement about the semantic representation of verbs means that we have not yet identified the features that might be most potent in promoting generalization of improved verb retrieval" (p.58). So, for example, generating the instrument of the action denoted by a verb, during SFA, might be more potent than generating the agent, if the former is stored as a feature of the verb but the latter is not. Another possibility is that generalisation of treatment to untrained verbs may require verbs to be treated in the context of a sentence. This explanation is supported by the suggestion that sentence treatments result in generalization more frequently than verb treatments (see the review by Webster & Whitworth, 2012). The reportedly more frequent generalization seen in sentence treatments may occur because sentence treatments require the production not just of the targeted verb during treatment but also its arguments. As noted above, most verbs are unconcerned about the particular phrase that appears in an argument slot, so long as the general semantic requirement of 'agent' or 'theme' is met. As a consequence, the production of a verb with its arguments may enable activation to spread more easily through the network. Thus sentence treatments may exploit verbs' looser semantic networks better than verb treatments. Alternatively, sentence treatments may result in greater generalisation because they are activating the syntactic information that verbs encode, resulting in a "syntactic bootstrapping" effect which verb treatments fail to capitalize upon.

Another possible explanation is drawn from anomia treatment research. The lack of generalisation of verb treatments mirrors that for anomia treatments in that they also predominantly improve only treated nouns (see e.g. Best et al., 2013). This is unsurprising in that verb treatments unashamedly use the same techniques as anomia treatments. It may indeed be the case that, as Howard (2000) has argued in relation to anomia treatments, all treatments whether they claim to be semantic, phonological (or, in the case of verbs, gestural) actually function as "mapping" treatments, here in the sense of strengthening mapping between semantic and phonological stages of lexical retrieval. If this is the case, then treatment effects will be item specific (i.e verb or noun specific) because the mechanism of treatment is lexically based. Finally it could be the case that verb treatments do not result in generalisation simply because they are not powerful enough and/or not enough treament was given to enable generalization.

Whatever the reason, the lack of generalization of verb treatments strongly suggests that treatments should target verbs which are *functionally useful* to participants. However, only two reviewed studies did this: Carragher et al. (2013) included five functionally relevant verbs in a treated set of forty, and Palmer et al. included 48 personally relevant nouns/verbs in a treated set of 96, although the results for these verbs are not reported on specifically in either study. Recommendations that functionally relevant verbs should be the target of treatment have been made before (e.g. Webster & Whitworth, 2012) but as yet this recommendation has not been followed in verb treatment research. As the evidence to date indicates that treated verbs only improve in the vast majority of cases, then we owe it to our participants to work on targets that are maximally meaningful to their lives.

Turning to the impact of verb treatments on sentence production using treated and untreated verbs, interpretation of the evidence here is seriously restricted by lack of assessment as only 13% and 24% of studies assessed sentence production using trained verbs and untrained verbs respectively. Whilst these studies reported improvements in sentence production for some of their participants, no firm conclusions can be drawn due to the small numbers involved, and it is thus recommended that verb treatments routinely include an assessment of sentence production as an outcome measure.

Regarding broader impact, the impact of verb treatments on functional communication and discourse was also difficult to evaluate because it was infrequently measured. Only four studies assessed functional communication, and whilst they found some evidence of improvement, the very small number of participants (16) means again that no firm conclusions can be drawn. Similarly, it is recommended that functional communication is routinely assessed in verb treatment research, and that ideally, researchers should reach agreement on the measure used to do this.

Discourse was more frequently assessed (in twelve or 32% of studies) and improvements were reported for 31% of participants (for whom individual results were reported). Given the limited generalization of verb treatments, it is perhaps not surprising that the impact of treatment on discourse is limited. It seems plausible that treating functionally relevant verbs might produce a greater improvement in discourse as these verbs may be used more frequently in discourse in real life. This however leads to the problem of the best way to assess discourse as the reviewed studies used a variety of methods, with few analyzing samples of discourse from daily life. Indeed, the assessment of discourse is currently a topic of debate for many reasons, including the plethora of measures available (Bryant, Ferguson & Spencer, 2016), the varied psychometric quality of those measures (Pritchard, Hilari, Cocks, & Dipper, 2017), and the potential for a core set of discourse outcome measures (e.g. Dietz & Boyle, 2017). Whilst there continues to be disagreement about the best way to assess discourse (Wallace, Worrall, Rose & Le Dorze, 2017), it is difficult to reach agreement regarding a Discourse Core Outcomes Set to be used in verb treatment studies and in aphasia treatment research in general. However, this is high priority for future research.

Finally, regarding what induces change in verb treatment, the most commonly reported treatment technique was *phonologica*l cueing (reported in 67.5% of studies) followed by *semantic* cueing (54% of studies). Verb treatments are similar to anomia treatments here as the latter also commonly use semantic and phonemic cues (e.g. Nickels, 2002) and this is unsurprising given that most verb in isolation treatments are derived from anomia treatments. What is perhaps surprising is that only 22 of the reviewed studies (59%) incorporated treatment techniques which were designed to exploit the unique, action related properties of verbs (such as the use of gesture cues, action observation, video stimuli and the adaptation of SFA for verbs). All of these treatments were designed to capitalize upon the unique features of verbs, and their impact is promising. Indeed it seems plausible that treatments which deliberately target verbs' unique properties have the potential to be more powerful than those which transpose techniques directly from anomia treatments because these unique properties may be the (most) active ingredients of verb treatments. Thus, for example, the use of *video* rather than static picture stimuli in treatment may be more effective in eliciting verb production (Blankestijn-Wilmsen et al., 2017) because seeing an action performed primes production of the related verb's lexical form (either by mirror neurons or some other mechanism). This is also supported by the emerging area of research which indicates that *action observation* (either face-to-face or via video stimuli) has the potential to be an active ingredient of treatment (Bonifazi et al., 2013; Marangolo et al., 2012 and 2010). The use of both video stimuli and action observation warrants further investigation.

Also warranting further investigation is the *self-delivery* of verb treatments via *computer* as studies investigating this reported improvements in treated verbs for 100% of participants. Whilst the significance of this finding is greatly tempered by the small number of participants (n=13), that verb treatment remained as (or possibly even more) effective when self-delivered is an alluring prospect, especially given that treating verbs has been considered more complex than treating nouns and might therefore be deemed unsuitable for self-delivery. There are many potential reasons why computer delivered verb treatments appear to work which may include the delivery of higher dosages of treatment (see Mortley et al., 2004). However there are potential disadvantages of computer-delivered verb treatments which include the need to train participants to use the treatment program, as well as the lack of supervision and feedback from a clinician (which was indeed seen as a disadvantage of self-delivered computer treatment by some participants and carers when interviewed by Palmer et al. (2013) and Wade et al. (2003)). There is thus an urgent need to explore in more detail how computer delivery affects (all) aphasia treatments whose efficacy has been established face-to-face, and indeed this is an emerging area of research. For example, Ball et al. (2018) found that whilst participants complied with the treatment protocol in only 45% – 61% of sessions of self-delivered anomia treatment, there were actually *more* successful naming attempts when participants did *not* comply with the treatment protocol than when they did. Ball et al. speculate that this may be because, through repeated interaction with the treatment program, participants identified the level of cueing which was most successful for them and adopted this. This points to the need for greater research into how participants *qualitatively* interact with computers during self-delivered computer treatment as well as how much, *quantitatively*, they receive.

The active status of verb treatment ingredients whether delivered face-to-face or by computer may depend on other factors, two of which will now be briefly discussed. The number of verbs treated in studies may influence the outcome. This review found that the minimum number of verbs treated in the reviewed studies was five (Fink et al., 1997) and the maximum 120 (Bonifazi et al., 2013). However, studies which treated larger number of verbs tended to treat them in smaller sets, usually to compare treatment techniques. For example, Boo & Rose (2011) and Rose and Sussmilch (2008) each treated 80 verbs, in sets of 20, with either semantic cues, repetition, gesture or combined cues to compare the effectiveness of each of these treatment techniques. Indeed, treating verbs in sets of 20 was the commonest design found in the reviewed studies, with four studies treating a total set of 20 verbs and seven studies treating 80 verbs in sets of 20. However, it is unclear whether this is the optimal number for treatment. Edwards & Tucker (2006) treated a set of 50 verbs. This was replicated by McCann & Doleman (2011) who treated 100 verbs but in two sets of 50, in a crossover design. These would appear to be the largest sets of verbs treated in the studies of face-to-face treatments, with significant improvement in treated verbs for five of the six participants. The largest set of verbs treated in the studies of computer delivered treatments however was 100 (alongside a set of 162 nouns) in Mortley et al. (2004). The use of computers to deliver treatment clearly has the potential to greatly increase the number of verbs treated. However, given the self-administered nature of treatment in this study it is not guaranteed that every verb will be treated in every treatment session (see Ball et al. (2018) who found that participants adhered to the treatment protocol in only 45-61% of self-administered sessions).

Whilst it may seem logical to treat a larger number of verbs (given that treatment is likely to improve retrieval of treated verbs only), treating a large number of verbs might effectively “dilute” the dose of treatment. The number of times each verb might be treated (i.e. go through the prescribed treatment protocol) during a one-hour treatment session will be considerably less for a verb which is part of a 50-verb set compared to one from a set of 20. In this scenario, treatment of a large set of verbs might conceivably be less effective than treatment of a small set. Indeed it is possible that an active ingredient of verb treatments which aim to improve the *production* of verbs is the *number of attempts at producing* a verb which are required during a treatment session and this may be less for a large set of treated verbs. However, of the reviewed studies only Conroy et al., (2009a; 2009b; 2009c) specified that verb naming was attempted at least 100 times for each of the treated verbs during their treatment programs. Reporting the number of times a verb was attempted has the additional benefit of allowing the optimal dose of treatment to be more accurately determined than if just the number and length of sessions is reported (although even this is inconsistently reported as discussed earlier) and it is therefore recommended. Two recent studies underline the importance of detailed reporting not just of the amount of treatment delivered but also how much of each potential active ingredient is delivered. Quique, Evans and Dickey (2018) conducted a meta-analysis of SFA (for anomia) and found that response to treatment was positively correlated with the amount of treatment. However, Gravier et al. (2018) found that response to treatment was more accurately predicted by the *number of features generated* by a participant during SFA than either total treatment time or the average number of treatment trials in their treatment study also using SFA for anomia with 17 participants.

Thus it seems that the number of features generated by a participant in SFA may be an active ingredient of treatment. Indeed Gravier et al. (ibid) go on to hypothesize that it may be the generation of features *specific* to a noun which predict improvement of *treated* items (e.g. *fuzzy* in relation to the treated noun peach), whereas generation of features *shared* within a category (such as *fruit, round* and *stone*) may be predict generalisation to *untreated* items (such as plum and nectarine) and this is the subject of future research for the authors. Whilst it is unclear how these findings might translate to SFA for verbs, It is such fine-grained research as this which has the potential to tease apart the active ingredients of aphasia treatments (see for example the series of studies by Marangolo and colleagues which are teasing apart the active ingredients of action observation treatment.)

A second factor which may influence the effectiveness of verb treatment ingredients is the *nature* and *severity* of the *deficit* underlying a verb retrieval difficulty. So for example, we might speculate that verb retrieval difficulties caused by a semantic difficulty would respond better to semantic treatment and likewise that phonologically based retrieval difficulties would respond better to phonological treatments. That is, semantic ingredients would be active during treatment of semantic deficits and phonological ingredients inactive, and vice versa for phonologically based deficits. We might also speculate that more severe verb retrieval deficits may be less responsive to treatment. The findings of this review only allow very tentative conclusion to be drawn here, with the suggestion that a variety of treatments (including semantic) were less effective in treating semantically based verb retrieval difficulties (Bonifiazi et al., 2013, Marangolo et al., 2010 Rodriguez et al., 2006; Wambaugh et al., 2004), and that participants with a more severe verb deficit may also be less responsive (Conroy et al., 2009c; Palmer et al., 2012). However, far more research is needed to establish what are the most effective (or active) ingredients of treatment for different types of verb retrieval deficits as this review also found no clear relationship between treatment given to participants and the nature of their underlying verb retrieval difficulties in line with Webster and Whitworth (2012).

Finally, whilst nearly half of the studies included in this review (16) compared treatment techniques to try to establish their status as active ingredients, no clear winners emerged from these studies. Indeed, what was clear from this review is that verb treatments generally constitute complex interventions (as defined by the Medical Research Council) and it is thus likely to be difficult to tease apart what are the individual, active ingredients of treatment. It is indeed entirely possible that it is the *combination* of ingredients in verb treatments that is actually the "active" ingredient. The finding that generalisation to untreated verbs was almost entirely restricted to treatment that combined techniques adds weight to this suggestion.

Before final conclusions are drawn limitations of this review will be briefly discussed. The review is not a systematic review. Thus studies were not blind reviewed by two or more reviewers but solely by the first author. This is because the review was completed as part of a doctoral research study. Studies were also not evaluated with a published, standard tool as this is lacking for case series which was the design used for 29 (78%) of the studies reviewed. Whilst the systematic reviews of SFA (Efstratiadou et al., 2018; Maddy et al., 2014) and of gestural treatment for verbs (Rose et al., 2013) used the Single-Case Experimental Design (SCED) Scale (Tate et al., 2008) to review both single cases and case series, this does not address those aspects of case series which are beyond the scope of single case design, such as whether it is appropriate to report group results with the small n of case series (as was the case for six studies in this review).

The review also does not cover treatments which target verbs *together with their arguments* (as opposed to *in isolation.)* Although SFA adapted for verbs does include arguments in the array of features (e.g. via the question 'Who usually does this?') the excluded treatments target a verb's arguments *within a sentence structure* (such as Verb Network Strengthening Treatment (VNeST: e.g. Edmonds, 2016, Edmonds, Obermeyer & Kiernan, 2015) and mapping treatments (e.g. Byng, Nickels & Black, 1994; Marshall, Chiat & Pring, 1997). This is an important limitation because the current evidence suggests that these sentence level treatments are more likely to results in improvements in *sentence production* (using treated verbs) than verb in isolation treatments (e.g. Webster & Whitworth, 2012). Sentence level treatments are therefore reviewed in a companion paper (Hickin, Cruice & Dipper, 2022)

In conclusion, whilst this review identified many challenges for verb treatment research, researchers should not be despondent. The predominant finding that treatment improves the retrieval of trained verbs for 80% of participants is very encouraging. The challenge is to demonstrate generalisation of treatment to untrained verbs, and this remains a challenge for anomia treatment too. Functional communication and discourse need to be routinely measured to establish if verb treatments affect these. As Carragher, Sage and Conroy (2015) vividly express it, demonstrating that treatment improves communication in daily life remains the "holy grail" for all aphasia treatments. We encourage verb treatment researchers to don their hats, crack their whips and, Indiana Jones like, pursue their quest.

Acknowledgements.

This research was supported by a doctoral studentship to the first author jointly funded by the Worshipful Company of Saddlers and the School of Health Sciences, City, University of London. The authors would also like to thank the two anonymous reviewers for their helpful comments on earlier drafts of this paper.

References.

Adrian J.A., Gonzalez. M., Buiza, J.J. & Sage, K. (2011). Extending the use of spanish computer-assisted anomia rehabilitation program (CARP-2) in people with aphasia. *Journal of Communication Disorders, 44*, 666–677.

Auerbach, C., Schnall, D., Heft LaPorte, H., & SingWin. (1995). In Bloom, M., Fischer, J. & Orme J.G, (editors). *Evaluating practice: guidelines for the accountable professional*. (2nd ed.) Boston (MA): Allyn & Bacon

Baker, E. (2012a). Optimal intervention intensity. *International Journal of Speech-Language Pathology, 14*, 401-409. https://doi.org/10.3109/17549507.2012.700323

Baker, E. (2012b). Optimal intervention intensity in speech-language pathology: Discoveries, challenges, and unchartered territories. *International Journal of Speech-Language Pathology,* 14, 478-485. https://doi.org/10.3109/17549507.2012.717967

Ball, A.L., de Riesthal, M. & Steele, R.D. (2018). Exploring treatment fidelity in persons with aphasia autonomously practicing with computerized therapy materials. *American Journal of Speech-Language Pathology, 27*, 454–463.

Beeson, P., & Robey, R. (2006). Evaluating single-subject treatment research: Lessons learned from the aphasia literature. *Neuropsychological Review*, 16, 161–169.

Berndt, R.S.,Haediges, A.N., Mitchum,C.C. and Sandson, J. (1997). Verb Retrieval in aphasia. 1. Characterizing single word impairments. *Brain and Language,* 56, 68–106.

Best, W., Greenwood, A., Grassly, J., Herbert, R., Hickin, J. & Howard, D. (2013). Aphasia rehabilitation: Does generalisation from anomia therapy occur and is it predictable? A case series study. *Cortex, 49*, 2345-2357. https://doi.org/10.1016/j.cortex.2013.01.005

Black, M & Chiat, S. (2003). Noun - verb dissociations: a multi-faceted phenomenon. *Journal of*

Neurolinguistics, 16, 231-250.

Blankestijn-Wilmsen, J., Damen, I., Voorbraak-Timmerman, V., Hurkmans, J., Koning, B. D., Pross, A., & Jonkers, R. (2017). The effect of static versus dynamic depictions of actions in verb and sentence production in aphasia. *Aphasiology,* 31, 1166-1182. https://doi.org/10.1080/02687038.2016.1258537

Bonifazi, S., Tomaiuolo, F., Altoè, G., Ceravolo, M. G., Provinciali, L. & Marangolo, P. (2013). *European Journal of Physical and Rehabilitation Medicine*, 49, 473-481.

Boo, M., & Rose, M. L. (2011). The efficacy of repetition, semantic, and gesture treatments for verb retrieval and use in broca's aphasia. *Aphasiology, 25*, 154-175. https://doi.org/10.1080/02687031003743789

Boyle, M. (2017). Semantic treatments for word and sentence production deficits in aphasia. *Seminars in Speech & Language, 38*, 52-61. https://doi.org/10.1055/s-0036-1597256

Brady, M.C., Kelly, H., Godwin, J., Enderby, P. & Campbell, P. (2016). 'Speech and language therapy for aphasia following stroke', *Cochrane Database of Systematic Reviews,* 6(Art. No.: CD000425). https://doi.org/10.1002/14651858.CD000425.pub4.

Bryant, L., Ferguson, A., & Spencer, E. (2016). Linguistic analysis of discourse in aphasia: A review of the literature. *Clinical Linguistics & Phonetics*, *30*, 489-518.

Byng, S., Nickels, L. & Black, M. (1994). Replicating therapy for mapping deficits in agrammatism: Remapping the deficit?, *Aphasiology, 8*, 315-341, https://doi.org/10.1080/02687039408248663

Carragher, M., Sage, K. & Conroy, P. (2015). Outcomes of treatment targeting syntax production in people with Broca's-type aphasia: evidence from psycholinguistic assessment tasks and everyday conversation. *International Journal of Language & Communication Disorders*, *50*, 322-336.

Carragher, M., Sage K, & Conroy, P. (2013). The effects of verb retrieval therapy for people with non-fluent aphasia: Evidence from assessment tasks and conversation. *Neuropsychological Rehabilitation*, *23*, 846-887. https://doi.org/10.1080/09602011.2013.832335

Cochrane Consumer Network (undated) Available at: http://consumers.cochrane.org/levels-evidence. [Accessed: 22 November 2017].

Conroy, P. & Scowcroft, J. (2012). Decreasing cues for a dynamic list of noun and verb naming targets: A case-series aphasia therapy study. *Neuropsychological Rehabilitation*, *22*, 295-318.

Conroy, P., Sage, K., & Lambon-Ralph, M. (2009a). A comparison of word versus sentence cues as therapy for verb naming in aphasia. *Aphasiology, 23*, 462 – 482.

Conroy, P., Sage, K., & Lambon-Ralph, M. (2009b). The effects of decreasing and increasing cue therapy on improving naming speed and accuracy for verbs and nouns in aphasia, *Aphasiology*, 23, 707–730.

Conroy, P., Sage, K. & Ralph, M.A.L. (2009c). Errorless and errorful therapy for verb and noun naming in aphasia. *Aphasiology, 23*, 1311-1337.

Conroy, P., Sage, K. & Ralph, M.A.L. (2006). Towards theory-driven therapies for aphasic verb impairments: A review of current theory and practice. *Aphasiology,* 20, 1159-1185. https://doi.org/10.1080/02687030600792009

de Aguiar, V., Bastiaanse, R. & Micel, G. (2016). Improving production of treated and untreated verbs in aphasia: a meta-analysis. *Frontiers in Human Neuroscience, 10,* 468. https://doi.org/10.3389/fnhum.2016.00468

Dietz, & Boyle, M. (2017). Discourse measurement in aphasia research: have we reached the tipping point? *Aphasiology*, https://doi.org/10.1080/02687038.2017.1398803

Dijkers, M. (2015). What is a scoping review. *KT Update, 4. An e-newsletter from the Center on Knowledge Translation for Disability and Rehabilitation Research*.

[http://ktdrr.org/products/update/v4n1]

Douglas, J. M., O'Flaherty, C. A., & Snow, P. C. (2000). Measuring perception of communicative ability: The development and evaluation of the La Trobe communication questionnaire. *Aphasiology, 14*, 251–268.

Edmonds, L. A. (2016). A review of verb network strengthening treatment. *Topics in Language Disorders, 36*(2), 123-135. https://doi.org/10.1097/TLD.0000000000000088.

Edmonds, L. A., Obermeyer, J., & Kernan, B. (2015). Investigation of pretreatment sentence production impairments in individuals with aphasia: Towards understanding the linguistic variables that impact generalisation in verb network strengthening treatment. *Aphasiology, 29*, 1312-1344. https://doi.org/10.1080/02687038.2014.975180

Edwards, S. & Tucker, K. (2006). Verb retrieval in fluent aphasia: A clinical study. *Aphasiology*, *20*, 644-675. https://doi.org/10.1080/02687030600631827

Efstratiadou, E.A., Papathanasiou, I., Holland, R., Archonti, A. & Hilari, K. (2018). A Systematic Review of Semantic Feature Analysis Therapy Studies for Aphasia. *Journal of Speech, Language, and Hearing Research*, *61*, 1261–1278.

Embry, D. D., & Biglan, A. (2008). Evidence-based kernels: Fundamental units of behavioral influence. *Clinical Child and Family Psychology Review, 11*, 1573 – 2827.

Faroqi-Shah. Y & Graham, L.E. (2011). Treatment of semantic verb classes in aphasia: acquisition and generalization effects. *Clinical Linguistics & Phonetics, 25*, 399–418.

Fillingham, J. K., Sage, K., & Lambon Ralph, M. A. (2005a). Further explorations and an overview of errorless and errorful therapy for aphasic word finding difficulties: The number of naming attempts during therapy affects outcome. *Aphasiology, 19*(7), 597–614.

Fink, R.B., Schwartz, M.F., Sobel, P.R. & Myers, J.L. (1997). Effects of multilevel training on verb retrieval: is more always better? *Brain & Language, 60*, 1, 41-45.

Garrett, M. (1988). Processing in language production. In F. Newmeyer (Ed.), *Linguistics: The Cambridge Survey, Vol 3.* Cambridge: Cambridge University Press.

Gravier, M.L., Dickey, M.W., Hula, W.D, Evans, W.S., Owens, R.L., Winans-Mitrik, R.L. & Doyle, P.J. (2018) What Matters in Semantic Feature Analysis: Practice-Related Predictors of Treatment Response in Aphasia. American Journal of Speech-Language Pathology, 27, 438–453. https://doi.org/10.1044/2017_AJSLP-16-0196

Hickin, J., Cruice, M., & Dipper, L. (2022). A systematically conducted scoping review of the evidence and fidelity of treatments for verb and sentence deficits in aphasia: sentence treatments. American Journal of Speech Language Pathology, 31, 1: 431-462. https://doi.org/10.1044/2021_AJSLP-21-00120.

Hinckley, J. J. & Douglas, N. F. (2013). Treatment fidelity: Its importance and reported frequency in aphasia treatment studies. *American Journal of Speech-Language Pathology*, 22, 279–284.

Howard, D., Best, W. & Nickels, L. (2015) Optimising the design of intervention studies: critiques and ways forward, *Aphasiology*, *29*, 526-562, https://doi.org/10.1080/02687038.2014.985884

Howard D. (2000). Cognitive neuropsychology and aphasia therapy: the case of word retrieval. In: Papathanasiou, I, ed. *Acquired Neurogenic Communication Disorders: A Clinical Perspective*. London: Whurr.

Jones, E. V. (1986). Building the foundations for sentence production in a non-fluent aphasic. *International Journal of Language & Communication Disorders, 21*, 63-82. https://doi.org/10.3109/13682828609018544

Kaderavek, J.N. and Justice L.M. (2010). Fidelity: An Essential Component of Evidence-Based Practice in Speech-Language Pathology. *American Journal of Speech-Language Pathology, 19*, 369–379.

Kim, M., & Thompson, C. K. (2000). Patterns of comprehension and production of nouns and verbs in agrammatism: implications for lexical organization. *Brain and Language, 74,* 1-25.

Knoph, M. I. N., Lind, M., & Simonsen, H. G. (2015). Semantic feature analysis targeting verbs in a quadrilingual speaker with aphasia. *Aphasiology, 29*, 1473-1496. https://doi.org/10.1080/02687038.2015.1049583

Kristensson, J. Behrns, I. & Saldert, C. (2015). Effects on communication from intensive treatment with semantic feature analysis in aphasia, *Aphasiology, 29*, 466-487, https://doi.org/10.1080/02687038.2014.973359

Kurland, J., Wilkins, A. R., & Stokes, P. (2014). iPractice: Piloting the effectiveness of a tablet-based home practice program in aphasia treatment. *Seminars in Speech & Language, 35*, 51-63. https://doi.org/10.1055/s-0033-1362991.

Leff, A.P. & Howard, D. (2012). Stroke: Has speech and language therapy been shown not to work? Nature Reviews Neurology. https://doi.org/10.1038/nrneurol.2012.211

Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge MA: MIT Press.

Lomas, J., Pickard, L., Bester, S., Elbard, H., Finlayson, A. & Zoghaib, C. (1989). The Communicative Effectiveness Index: development and psychometric evaluation of a functional communication measure for adult aphasia. *Journal of Speech and Hearing Disorders*, *54*, 113-124.

Long, A. F., Hesketh, A., Paszek, G., Booth, M., & Bowen, A. (2008). Development of a reliable self-report outcome measure for pragmatic trials of communication therapy following stroke: The Communication Outcome after Stroke (COAST) scale. *Clinical Rehabilitation*, *22*, 1083–1094. https://doi.org/10.1177/0269215508090091

Luzzatti, C., Raggi, R., Zonca, G., Pistarini, C., Contardi, A., & Pinna, G. (2001). On the nature of the selective impairment of verb and noun retrieval. *Cortex, 37*, 724–726.

Maddy, K.M., Capilouto, J.G. & McComas, ,K.L. (2014). The effectiveness of semantic feature analysis: An evidence-based systematic review. *Annals of Physical and Rehabilitation Medicine, 57*, 254–267. http://dx.doi.org/10.1016/j.rehab.2014.03.002

Maher, C. G., Sherrington, C., Herbert, R. D., Moseley, A. M., & Elkins, M. (2003). Reliability of the PEDro scale for rating quality of randomized controlled trials. Physical Therapy, 8, 713–721.

Marangolo P, Cipollari S, Fiori V, Razzano C. & Caltagirone C. (2012). Walking but Not Barking Improves Verb Recovery: Implications for Action Observation Treatment in Aphasia Rehabilitation. *PLoS ONE 7(*6): e38610. https://doi.org/10.1371/journal.pone.0038610

Marangolo P, Bonifazi S, Tomaiuolo F, Craighero L, Coccia M, et al., (2010). Improving Language without words: first evidence from aphasia. *Neuropsychologia, 48*, 3824–33.

Marshall, J., Pring, T. & Chiat, S. (1998) Verb Retrieval and Sentence Production in Aphasia. *Brain and Language, 63*, 159–183.

Marshall, J., Chiat, S. & Pring, T. (1997). An impairment in processing verbs' thematic roles: A therapy study. *Aphasiology, 11*, 855 — 876. http://dx.doi.org/10.1080/02687039708250461

McCann C. & Doleman, J. (2011). Verb retrieval in nonfluent aphasia: A replication of Edwards & Tucker, 2006. *Journal of Neurolinguistics, 24*, 237–248.

McNeil , M.R., Doyle, P.J. , Spencer, K. Jackson Goda, V.,Flores, D. & Small, S.L. (1998) Effects of training multiple form classes on acquisition,generalization and maintenance of word retrieval in a single subject, *Aphasiology, 12*, 575-585. https://doi.org/10.1080/02687039808249559

Medical Research Council (2006) Developing and evaluating complex interventions: new guidance. Available at: www.mrc.ac.uk/complexinterventionsguidance. [Accessed 7th February 2018].

Moher D., Liberati A., Tetzlaff J. & Altman DG. The PRISMA Group. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med* 6(6):

e1000097. https://doi.org/10.1371/journal.pmed1000097

Mortley, J., Wade, J. & Enderby, P. (2004). Superhighway to promoting a client-therapist partnership? Using the Internet to deliver word-retrieval computer therapy, monitored remotely with minimal speech and language therapy input. *Aphasiology*, *18*, 193-211.

Nickels, L. (2002). Therapy for naming disorders: Revisiting, revising, and reviewing. *Aphasiology, 16*, 935–979.

Palmer, R., Enderby, P. & Paterson, G. (2013). Using computers to enable self-management of aphasia therapy exercises for word finding: the patient and carer perspective. *International Journal of Language & Communication Disorders, 48*, 508-521.

Palmer, R., Enderby, P., Cooper, C., Latimer, N., Julious, S., Paterson, G., . . . Dixon, S. (2012). Computer therapy compared with usual care for people with long-standing aphasia poststroke: A pilot randomized controlled trial. *Stroke (00392499), 43*, 1904-1911. https://doi.org/10.1161/STROKEAHA.112.650671

Peters, M.D.J., Godfrey, C.M., Khalil, H., McInerney, P., Parker, D. & Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *International Journal of Evidence-Based Healthcare,* 13 3: 141-146. https://doi.org/10.1097/XEB.0000000000000050

Pritchard, M., Hilari, K., Cocks, N., & Dipper, L. (2017). Reviewing the quality of discourse information measures in aphasia. International Journal of Language & Communication Disorders, 52, 689-732.

Quique, Y.M., Evans, W.S. & Dickey, M.W. (2018). Acquisition and Generalization Responses in Aphasia Naming Treatment: A Meta-Analysis of Semantic Feature Analysis Outcomes. *American Journal of Speech-Language Pathology, 28*, 1-17. https://doi.org/10.1044/2018_AJSLP-17-0155

Raymer, A.M. , Ciampitti , M., Holliway, B., Singletary, F., Blonder, L.X., Ketterson, T., Anderson , S., Lehnen, J., Heilman, K.M. & Gonzalez Rothi, L.J. (2007). Semantic-phonologic treatment for noun and verb retrieval impairments in aphasia, *Neuropsychological Rehabilitation, 17*, 244-270, https://doi.org/10.1080/09602010600814661

Raymer, A.M., Singletary, F., Rodriguez, A., Ciampitti, M., Heilman, K.M. & Gonzalez Rothi, L.J. (2006). Effects of gesture+verbal treatment for noun and verb retrieval in aphasia. *Journal of the International Neuropsychological Society*, *12*, 867–882. https://doi.org/10.10170S1355617706061042

Raymer, A., & Kohen, F. (2006). Word-retrieval treatment in aphasia: Effects of sentence context. *Journal of Rehabilitation Research & Development, 43*, 367-377.

Raymer, A., Rodriguez, A., Ciampitti, M., Singletary, F. Fuller, R., Heilman, K. & Gonzalez Rothi, L. (2005). Effects of Pantomime Training for Noun and Verb Retrieval in Aphasia. *Neurorehabilitation and Neural Repair 19*, 378.

Raymer, A.M. & Ellsworth, T.A. (2002). Response to contrasting verb retrieval treatments: A case study. *Aphasiology*, *16*, 1031-1045. https://doi.org/10.1080/026870401430000609

Robey, R. R., & Schultz, M. C. (1998). A model for conducting clinical outcome research: An adaptation of the standard protocol for use in aphasiology. *Aphasiology, 12*, 787–810.

Rochon, E., Laird, L., Bose, A. & Scofield, J. (2005). Mapping therapy for sentence production impairments in nonfluent aphasia. *Neuropsychological Rehabilitation, 15*, 1-36. https://doi.org/10.1080/09602010343000327

Rochon, E. & Reichman, S. (2003). A Modular Treatment for Sentence Processing Impairments in Aphasia: Sentence Production. *Journal of Speech-Language Pathology and Audiology,* 27, 202 -210.

Rodriguez, A. D., Raymer, A. M., & Gonzalez Rothi, L. J. (2006). Effects of gesture+verbal and semantic‐phonologic treatments for verb retrieval in aphasia. *Aphasiology*, *20*, 286-297.

https://doi.org/10.1080/02687030500474898

Rose, M.L., Raymer, A.M., Lanyon, L.E. & Attard, M.C. (2013). A systematic review of gesture treatments for post-stroke aphasia, *Aphasiology, 27*, 1090-1127. https://doi.org/10.1080/02687038.2013.805726

Rose, M., & Sussmilch, G. (2008). The effects of semantic and gesture treatments on verb retrieval and verb use in aphasia. *Aphasiology, 22*, 691-706. https://doi.org/10.1080/02687030701800800.

Routhier, S., Bier, N. & Macoir, J. (2016) Smart tablet for smart self-administered treatment of verb anomia: two single-case studies in aphasia. *Aphasiology, 30*, 269-289, https://doi.org/10.1080/02687038.2014.973361

Routhier, S., Bier, N. Macoir, J. (2015). The contrast between cueing and/or observation in therapy for verb retrieval in post-stroke aphasia. *Journal of Communication Disorders 54*, 43–55. http://dx.doi.org/10.1016/j.jcomdis.2015.01.003

Schneider, S. & Thompson, C. (2003). Verb production in agrammatic aphasia: The influence of semantic class and argument structure properties on generalization. *Aphasiology, 17*, 213-241. https://doi.org/10.1080/729255456

Takizawa, T., Nishida, N., Ikemoto, A., & Kurauchi, N. (2015). Comparison of single-word therapy versus sentence therapy for verb retrieval and sentence production in Broca’s aphasia. *Aphasiology, 29*, 169-194. https://doi.org/10.1080/02687038.2014.961121

Tate, R.L., McDonald, S., Perdices, M., Togher, L., Schultz, R. & Savage, S. (2008). Rating the methodological quality of single-subject designs and n-of-1 trials: Introducing the Single-Case Experimental Design (SCED) Scale. *Neuropsychological Rehabilitation, 18*, 385–401. https://doi.org/10.1080/09602010802009201

Thompson, C.K., Lange, K.L, Schneider, S.L. & Shapiro, L. P. (1997). Agrammatic and non-brain-damaged subjects' verb and verb argument structure production. Aphasiology, 11, 473-490, https://doi.org/10.1080/02687039708248485

Verhagen, A., de Vet, H., de Bie, R., Kessels, A., Boers, M., Bouter, L., & Knipschild, P. (1998). The Delphi List: A criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi Consensus. *Journal of Clinical Epidemiology*, *51*, 1235–1241.

Wade, J., Mortley, J. & Enderby, P. (2003). Talk about IT: Views of people with aphasia and their partners on receiving remotely monitored computer-based word finding therapy. *Aphasiology, 17*, 1031-1056. https://doi.org/10.1080/02687030344000373

Wallace, S.J., Worrall, L.E. Rose, T. & Le Dorze, G.(2017). Discourse measurement in aphasia research: have we reached the tipping point? A core outcome set … or greater standardisation of discourse measures? *Aphasiology,* https://doi.org/10.1080/02687038.2017.1398811

Wambaugh, J. L., Mauszycki, S., & Wright, S. (2014). Semantic feature analysis: Application to confrontation naming of actions in aphasia. *Aphasiology, 28*, 1-24. doi:10.1080/02687038.2013.845739

Wambaugh, J.l. & Ferguson, M. (2007). Application of semantic feature analysis to retrieval of action names in Aphasia. *Journal of Rehabilitation Research and Development*, *44*, 381-394.

Wambaugh, J., Cameron, R., Kalinyak-Fliszar, M., Nessler, C., & Wright, S. (2004). Retrieval of action names in aphasia: Effects of two cueing treatments. *Aphasiology, 18*, 979-1004. https://doi.org/10.1080/02687030444000471

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*, 455-466.

Webster, J. and Whitworth, A. (2012). Treating verbs in aphasia: exploring the impact of therapy at the single word and sentence levels. *International Journal of Language & Communication Disorders, 47*, 619–636. doi:10.1111/j.1460-6984.2012.001

Chapter 2. Hickin, J., Cruice, M., & Dipper, L. (2022). A systematically conducted scoping review of the evidence and fidelity of treatments for verb and sentence deficits in aphasia: sentence treatments. *American Journal of Speech Language Pathology, 31,* 1: 431-462. https://doi.org/10.1044/2021_AJSLP-21-00120*.*

A systematically-conducted scoping review of the evidence and fidelity of treatments for verb and sentence deficits in aphasia: sentence treatments.

Authors: Julie Hickin,[1] Madeline Cruice[1] and Lucy Dipper[1]

[1] Division of Language and Communication Science, City, University of London, UK

Corresponding Author: Julie Hickin, Division of Language and Communication Science, City, University of London, Northampton Square, London EC1V 0HB. Email: julie.hickin@city.ac.uk.

Conflict of Interest: There are no relevant conflicts of interest.

Funding Statement: This research was supported by a doctoral studentship to the first author jointly funded by the Worshipful Company of Saddlers and the School of Health Sciences, City, University of London.

Purpose. This paper synthesizes and evaluates the evidence for sentence production treatments in aphasia, systematically charting impairment based and functional communication outcomes. It reports i) level of evidence and fidelity of sentence treatments, ii) impact of treatment on production of trained and untrained verbs and sentences, functional communication and discourse, and iii) discusses potential active ingredients of treatment.

Method. The search included studies January 1980 to June 2019. The level of evidence of each study was documented, as was fidelity in terms of treatment delivery, enactment and receipt. Studies were also categorised according to treatment methods used.

Results. Thirty-three studies were accepted into the review and predominantly constituted Level 4 evidence (e.g. case control studies and case series). Thirty studies (90%) described treatment in sufficient detail to allow replication, but dosage was poorly reported, and fidelity of treatment was rarely assessed. The most commonly reported treatment techniques were mapping (10 studies: 30%), Predicate Argument Structure treatment (6 studies: 18%), and Verb Network Strengthening Treatment (5 studies: 15%). Production of trained sentences improved for 83% of participants, and improvements generalised to untrained sentences for 59% of participants. Functional communication was rarely assessed but discourse production improved for 70% of participants.

Conclusions. The evidence for sentence treatments is predominantly generated from Level 4 studies. Treatments were effective for the majority of participants

regarding trained sentence and discourse production. However, there is inconsistent use of statistical analysis to verify improvements, and diverse outcome measures are used which makes interpretation of the evidence difficult. The quality of sentence treatment research would be improved by agreeing a core set of outcome measures and extended by ascertaining the views of participants on sentence treatments.

Introduction.

There has been increasing interest in the treatment of verb and sentence production deficits (e.g. Carragher, Sage & Conroy, 2015; Edmonds, Obermeyer & Kernan, 2015; Kurland, Liu & Stokes, 2018; Newton, Kirby & Bruce, 2017). However, the most recent review of studies investigating sentence production treatment for individuals with aphasia included studies to March 2011 only (Webster & Whitworth, 2012). At this time eight studies of sentence production treatment were available for review. An up-date was needed to capture more recent literature. Additionally, no existing review has systematically charted the impact of treatment on production of verbs in isolation, on sentence and discourse production and on functional communication: the review reported here charts all of these. Finally, existing reviews of sentence production treatments do not report on the *fidelity* of sentence treatment studies – a more recent focus of aphasia research.

The review is a companion to that carried out by Hickin, Dipper and Cruice (2020) who reviewed treatments that aimed to improve verb and sentence production deficits in aphasia by treating verbs *in isolation* (i.e. treatments which focused on the *lexical properties* of verbs in isolation). In this companion review, treatments that aimed to improve sentence production by treating verbs *in sentences* (that is treatments which focused on the *relationship* between a verb and its *arguments* in the context of a *sentence*) are reviewed.

The following introductory section sets the stage for the ensuing review by first summarizing how far theories of sentence processing have informed the development of sentence production treatments for individuals with aphasia. Second, the existing reviews of sentence production treatments are briefly discussed to elucidate the additional contributions made by the study reported here. Third, the importance of assessing the fidelity of sentence production treatment studies is highlighted.

Sentence production deficits in aphasia: theory and therapy.

Whilst theory is increasingly informing sentence production treatments, there is much left to learn about which theories underpin the most effective sentence production treatments, and for whom they are most effective. The review reported in this paper set out to begin to address this gap in knowledge with an up-dated review of sentence treatment studies. One issue addressed by the review was to establish whether there is any evidence to guide *clinical decision making* as to what might be the most effective, theoretically motivated treatment for a particular sentence production deficit. Jones (1986) proposed a treatment based on the interaction (or mapping) between linguistic levels. She argued that classical treatments which targeted the syntactic complexity of sentences at surface level only overlooked the need for mapping between syntax and semantics. Jones's successful treatment of a patient with a hypothesized mapping deficit gave rise to a raft of further *mapping treatment* studies (e.g. Marshall, 1997; Rochon, Laird, Bose & Scofield, 2005; Schwartz, Saffran, Fink, Myers & Martin, 1994). The review also identified treatment research underpinned by other linguistic theories. These include treatment targetting a theorized deficit in *predicate argument structure* (PAS treatment: e.g. Bazzini et al., 2012; Biran & Fisher, 2015; Webster, Morris & Franklin, 2005) and treatment predicated on the spreading activation theory of semantic processing proposed by Collins and Loftus (1975) called Verb Network Strengthening Treatment (VNeST: e.g. Edmonds, Mammino & Ojeda, 2014; Edmonds, Nadeau & Kiran, 2009).

Sentence production treatments for people with aphasia have also been influenced by theories of *cognitive* processing. For example, Kolk (1995) viewed agrammatic sentence production as an adaptive response to *limited* cognitive processing resources (such as deficits in attention) resulting in the production of elliptical or telegrammatic utterances. A compensatory treatment based on this theory is REduced Syntax Therapy (REST: e.g. Springer, Huber, Schlenck & Schlenck, 2000)[3] and one reviewed study incorporated elements of REST (together with mapping treatment) with the aim of restoring sentence production (Carragher, Sage & Conroy, 2015). VNeST on the other hand aims to recruit *intact* cognitive abilities in treatment (i.e. episodic and autobiographical memory) by requiring participants to produce *personally relevant* agent-verb-theme exemplars during treatment (e.g. Edmonds et al., 2014). Finally, learning theory - in terms of *constraint induced (CI) learning* - has influenced the development of Constraint Induced Language Treatment (CILT: e.g. Goral & Kempler, 2009). Constrained induced (CI) learning was first used to treat recovery of movement following stroke and was implemented in aphasia treatment by Pulvermuller et al. in 2001. CILT embraces the CI principles of *massed* practice, *shaping* of responses, *constraint* of the less impaired/more easily accessible communication modality/ies and implementation of treatment that is *functionally relevant*.

In summary, theories of sentence processing appear to be increasingly informing the development of treatments for sentence production deficits in aphasia, with VNeST being the most recent example of this. However, there is a need for further research to identify which of the theoretically motivated treatments are most effective, and for whom.

Existing reviews of sentence treatments.

To date, two reviews have been published that examine the nature of sentence production treatment for individuals with aphasia (Conroy, Sage & Lambon-Ralph, 2006; Webster & Whitworth, 2012). Conroy et al. (2006) reviewed 10 studies, four of which investigated sentence treatments, whilst Webster and Whitworth (2012) reviewed 26 studies including seven sentence treatment

[3] Papers investigating compensatory approaches such as REST were not included in this systematic review since the aim of these approaches is not to restore normal sentence production.

studies, three of which were included by Conroy et al[4]. Outcomes of treatment studies were compared in terms of the impact on verb retrieval and sentence production using both treated and untreated verbs, and in terms of changes in connected speech. In terms of the efficacy of sentence treatments, Webster and Whitworth's review indicated that sentence treatments appeared to be more effective in improving *sentence* production than verb-in-isolation treatments: verb-in-isolation treatments resulted in improved sentence production for seven out of 15 participants compared to sentence treatments which resulted in improved sentence production for seven out of eight participants (magnitude of gain not reported). That sentence treatments may be more effective at improving sentence production than verb treatments is important both clinically and theoretically. However, the review of a larger number of studies (with more participants) is required to investigate the robustness of this finding: establishing the impact of sentence treatments on (trained and untrained) sentence production was a key aim of this review.

The influence of both linguistic and cognitive processing theories on sentence production treatments highlights that they are complex. This makes identifying the potential active ingredients of sentence treatments (another key aim of this review) a difficult process. In this regard, one important issue is whether treating *verb retrieval* is an active ingredient of sentence treatments, in addition to the active ingredient of treating deficits in *syntactic* processing. This issue is important to explore since treatment which works on a lexical basis may be less likely to generalise than one which improves syntactic processing. The review reported here thus systematically charts the impact of treatment on both treated and untreated verbs *and* treated and untreated sentence production with the aim of elucidating this issue. As well as verb and sentence production, the review charts the impact of treatment on discourse production and on functional communication

[4] The other studies reviewed by Conroy et al. and Webster and Whitworth investigated verb-in-isolation treatments

(where this has been assessed). This is because in their review, Webster and Whitworth (2012) note that the impact of verb-in-isolation and sentence treatments on *connected speech* was difficult to determine since it was infrequently assessed. They conclude that a more systematic approach to evaluation of the outcomes of treatment for spoken sentence production deficits in aphasia is required, and this is the comprehensive approach taken in the review reported here. Finally Brady and colleagues' Cochrane review (2016) reported on 57 randomised controlled trials (RCTs) of aphasia treatment. Of these RCTs, only one specifically targeted sentence production (Rochon, Laird, Bose & Scofield, 2005), and this study also had the smallest n of all the those reviewed (n=5)[5]. There is thus a clear need for further research into sentence production treatments.

Fidelity of sentence treatment studies.

The fidelity of sentence production treatment studies has not been reviewed to date and so the fidelity of the studies included in this review was evaluated. The neglect of fidelity in aphasia treatment research was highlighted by Hinckley and Douglas (2013) who reviewed aphasia treatment studies published in the previous ten years and found that only 14% of 149 studies reviewed assessed the fidelity of treatment. They recommended that three levels of treatment fidelity needed to be addressed to improve the quality of aphasia therapy research namely, *treatment delivery* (e.g. by the use of treatment manuals and training), *treatment receipt* (e.g. by the use of homework record sheets and establishing the views of recipients regarding their treatment), and *treatment enactment* (e.g. by observation of treatment delivery). The fidelity of the studies accepted into this review is reported according to these three levels.

[5] Brady et al. (2016) report (p.55) that their review used both published and unpublished data from Rochon et al's study (2005) (n=3 plus 2 control participants) that presumably enabled its classification as an RCT.

Kaderavek and Justice (2010) argued that poor reporting of treatment fidelity had the potential to impede the implementation of research into clinical practice. This is because if treatment is not reported in terms of not only the *planned* but also the *actual* implementation, it is not possible for clinicians to accurately replicate treatment and validly transfer research into practice. More recently Brogan, Ciccone and Godecke (2019) reviewed the implementation and reporting of treatment fidelity in 42 RCTs of aphasia treatment published between 2012 and 2017. They found that whilst 88% of studies addressed treatment fidelity in terms of reporting dosage, still only 21% of studies addressed the fidelity of treatment procedures per se (e.g. by monitoring adherence to the planned treatment protocol). Brogan et al. reiterate the importance of reporting (and indeed implementing) treatment fidelity procedures in order to both strengthen the evidence base of aphasia treatment and to facilitate knowledge transfer. In particular they note that fidelity measures should attend to reporting the *rationale* underlying treatment to ensure, for example, that theoretically motivated (and therefore valid) outcomes measures are selected in treatment studies, and also to shed light on what might be the potential active ingredients of treatment - contained in what has sometimes been referred to as the treatment "black box." The underlying rationales for treatments are reported in this review, as are the potential active ingredients. Finally, evaluation of the fidelity of reporting in reviewed studies was informed by the template for intervention description and replication (TIDieR) checklist (Hoffmann et al., (2014) available at https://www.equator-network.org/wp-content/uploads/2014/03/TIDieR-Checklist-PDF.pdf). The TIDieR checklist includes noting the Why? of treatment (e.g. rationale), the What? (e.g. materials and procedures), How much? (including duration and intensity) and How well? (i.e. how far treatment was delivered *as planned* and how this was monitored).

Review Methodology.

The purpose of this review was to synthesise and evaluate the evidence for treating sentence production in aphasia, whilst systematically charting the impact of treatment on impairment based and functional communication outcomes.

The research questions for this review were:

6. What are the levels of evidence for sentence production treatments for people with aphasia?
7. What is the fidelity of aphasia research investigating sentence production treatments in terms of treatment delivery, receipt, and enactment (as defined by Hinckley & Douglas, 2013)?
8. What is the evidence of positive gains for sentence production treatments for people with aphasia in terms of improved production of a) trained and untrained verbs in isolation and b) trained and untrained verbs in sentences (within and across level generalization)?
9. What is the evidence of positive gains for sentence production treatments for people with aphasia in terms of a) improved functional communication and b) improved production of discourse?
10. What are the potential active ingredients of treatments for sentence production deficits in aphasia?

A scoping review paradigm was selected to evaluate the sentence production treatment evidence base because of the suitability of this method for answering broad evaluations questions, such as research questions 1-4, above (Tricco et al., 2018). This research design is also appropriate for synthesising this complex, heterogeneous, evidence base which is not amenable to a more precise systematic review (Dijkers, 2015; Peters, Godfrey, Khalil, McInerney, Parker, & Soares, 2015). The diversity of outcome measures used to evaluate the efficacy of sentence treatment also supports the use of this methodology, in that measures relate to verb and/or sentence and/or discourse production, and to functional communication – research questions 3 and 4. Finally, the scoping review was conducted using systematic procedures to ensure that these were rigorous, explicit and replicable and these are reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher, Liberati, Tetzlaff & Altman, 2009) adapted for scoping reviews (PRISMA – ScR checklist: Tricco et al., 2018 - see Appendix 1).

Method.

Cinahl Complete and Medline Complete databases were searched using the terms *sentence* OR *sentence production* AND *aphasia* AND *treatment* OR *therapy.* Studies which were original research and which were published in peer-reviewed journals, in English from 1980 up to June 2019 were considered for inclusion. Studies were excluded if they:

a) investigated other types of aphasia treatment (e.g. anomia treatment, dysgraphia treatment, conversation training), addressed other aspects of aphasia management (such as assessment) or investigated sentence treatment but:

b) had participants with another form of aphasia (e.g. progressive aphasia)

c) whose primary aim was to improve sentence *comprehension*

d) whose aim was not to restore normal sentence production (i.e. compensatory treatments such as REST e.g. Springer et al., 2000)

e) studies which were reviews or meta-analyses rather than original research

f) studies which investigated Treatment of Underlying Forms (TUF). TUF is predicated on the hypothesis that generalization of treatment to untrained sentences occurs because *complex* structures generalize to (related) *simpler* structures but not in the opposite direction (Complexity Account of Treatment Effectiveness: CATE e.g. Thompson et al., 2003)). It is acknowledged by the authors of TUF that it “requires considerable linguistic knowledge as well as a substantial amount of training to administer” (Thompson, Choy, Holland & Cole, 2010 p.1244). The scoping review of sentence treatments reported here was carried out to inform the content of a novel Sentence Production Treatment program to be *self-delivered* by people with aphasia (PwA) via computer. Studies investigating TUF were therefore excluded as it was anticipated that the treatment would be too hard for PwA to self-administer, independently, at home[6].

The titles of all papers identified in the database search were read. If the purpose of the study was not clear from the title, the abstract was read to determine whether the study should be included. All papers accepted into the review were read in full.

In order to answer research question 1 (regarding the level of evidence for sentence treatments), the aims and design of the study, and the number of participants were charted. Research question 2 addresses the fidelity of sentence treatment studies. As described in Hickin et al. (2020), there is no pre-established protocol for evaluating fidelity and therefore a tailored approach was developed to evaluate the studies in this review. The fidelity of the reviewed studies was charted in line with the recommendations of Hinckley and Douglas (2013) and Brogan et al. (2019) in terms of reporting treatment fidelity. Two papers by Baker (2012a; 2012b) were influential in determining how to evaluate the fidelity of reporting of treatment *dose* in reviewed studies. Baker discusses the level of reporting required to determine the *optimum* dose of treatment. She points out that the concept of an optimum treatment dose originates in pharmacology, with the aim being to identify exactly how much of a drug is needed to cure a disease in terms of dose strength, form, intensity and duration. To determine the optimum dose of a *therapeutic* rather than drug

[6] Thompson et al. (2010) report the successful implementation of TUF via computer (Sentactics). However, a clinician was present in the room for all treatment sessions to initiate treatment and ensure there were no technical issues (p.1249) i.e. it was not self-delivered, independently.

treatment, Baker states that it is necessary to describe treatment session *duration*, session *frequency* and the *duration of treatment* as a whole. All of these details were charted for the reviewed studies where reported.

In order to answer the next two research questions, the impact of sentence production treatments on production of trained and untrained verbs in isolation and in sentences was charted (research question 3), as was the impact on functional communication and discourse (research question 4). Research question 5 addressed the potential active ingredients of sentence treatments. In order to try to identify these, the two papers published by Baker in 2012 were again useful. Baker points out that the concept of an active ingredient of treatment is again drawn from pharmacology. In terms of *behavioural* treatments, Baker also points out that it is treatment *activities* which are the potential active ingredients of treatments. She makes a useful distinction between treatment activities which comprise *therapeutic inputs* (i.e. acts of the clinician) - such as giving a cue, and those which are *client acts* or responses expected of the client - such as repeating a word. Thus, in order to inform research question 5, these were recorded. Also influential was a paper by Byng, Nickels and Black (1994). They discuss the need for a "theory of therapy" to inform the process of determining what type of aphasia treatment is appropriate for a client. They discuss that this includes not only describing treatment activities but also the nature of the *interaction* between therapist and client since this is also likely to influence the outcome of treatment (e.g. how/whether the rationale for treatment is discussed and whether corrective feedback is given during treatment). Thus these aspects of treatment were also charted (when reported) to inform the potential active ingredients of treatment.

The data extracted from the reviewed studies were entered into a Microsoft Excel spreadsheet and are reported in Tables 2 and 3 below. All 33 studies were reviewed by the first author only (JH), with 10% of studies blind reviewed by the second and third author respectively

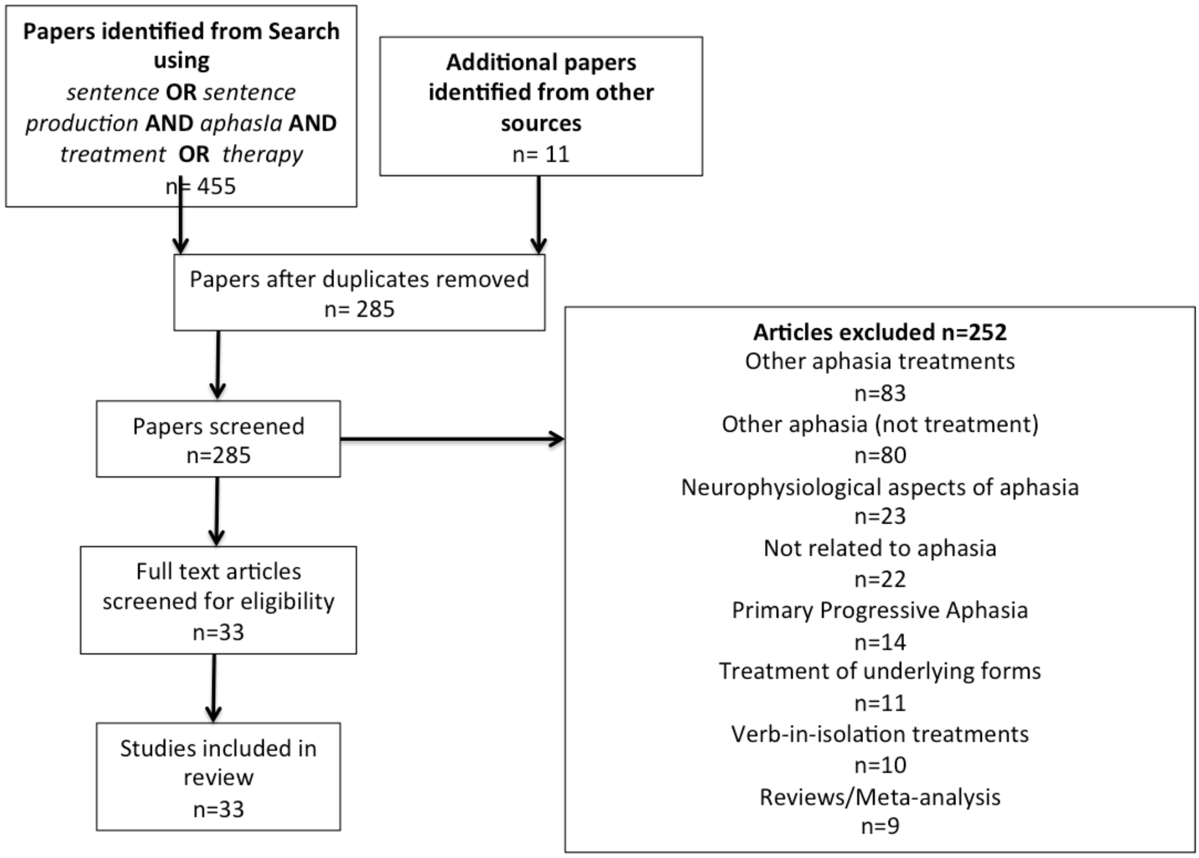
(MC and LD: three studies each). There was 95% agreement between reviewers with any disagreements resolved via discussion.

The results of treatment for *individual* participants are described whenever these are reported. When a study reported results for the *group* only, this is reported separately i.e. the participants in these studies were not added to the totals for individual participants because it was not clear how many participants in the group had (or had not) benefitted from treatment.

Results.

The results of the scoping review are summarized in Figure 1, reported according to PRISMA guidelines (Moher et al., 2009; Tricco et al., 2018). The combined searches resulted in 455 papers, with another 11 papers identified from additional sources. Once duplicates were removed, the titles and abstracts of 285 papers were screened. As a result of this screening, 252 papers were excluded (see Figure 1), the full text of the remaining 33 papers was screened, and all papers were included in the subsequent review.

Figure 1. Results of the systematically conducted scoping review for sentence treatments.



The details of the included studies are summarized in Table 1 and the impact of treatment on language and functional communication is reported in Table 2. Thirty-two papers reported sentence treatment delivered face-to-face, and one reported treatment delivered via computer (Furnas & Edmonds, 2014).

Table 1. The design, aims, type and amount of treatment given in studies accepted into the scoping review. **(Abbreviations**: C = complement; min = minimum; max = maximum; S = sentence; V = verb; O = object; TMA = trans-motor aphasia Tx= treatment).

1.Paper	2.Study Design	3.Study Aim	4.Type of treatment	5.Amount of Treatment
Bastiaanse, Hurkmans, & Links, 2006	case series (n=2)	to test the efficacy of the Verb Production at the Word and Sentence Level (VWS) treatment program.	Verb Production at the Word and Sentence Level (VWS) treatment program	12 weeks of treatment, (3 phases of 4 weeks) 3 times per week for 30 minutes (total =18 hours), plus homework
Bazzini et al., (2012)	case series + group results (n=8)	to explore the effectiveness of a treatment program for verb	Verb Argument Structure (VAS) treatment	6/8 Ps had 30 hours treatment, 1P had

		argument on the speed and accuracy of production.		32 hours and the final P 50 hours.
Biran & Fisher (2015)	case series (n=2)	to assess the effectiveness of a predicate argument structure Predicate Argument Structure (PAS) treatment.	PAS treatment	7 hours for AB & 10 hours for HY; 1 or 2 sessions per week
Byng, Nickels & Black (1994)	case series n=3	to replicate mapping treatment carried out by Byng (1988)	mapping treatment	3 phases of 6 weeks (=18 weeks) x2 weekly for 1 hour (total = 36 hours)
Carragher, Sage & Conroy (2015)	case series + group results (n=9)	to investigate the effects of a "hybrid" theoretically motivated therapy for sentence production.	Reduced Syntax Treatment (REST) plus elements of mapping therapy	8 x c1 hour sessions over 8 weeks, plus homework
Davis & Tan (1987)	single case study	to assess the effectiveness of sentence stimulation on sentence production	sentence stimulation	3 hours a week for 6 weeks (total = 18 hours)
Edmonds, Mammino & Ojeda (2014)	case series + group results (n=11: 5 anomic, 2 conduction, 2 TMA, 1 Wernicke's, 1 mixed)	to extend and replicate previous findings regarding Verb Network Strengthening Treatment (VNeST)	VNeST	35 hours, twice per week, over 10 weeks; each verb trained once a week
Edmonds & Babb (2011)	case series (n=2)	to establish if VNesT is effective with people with more severe aphasia	VNeST	P1 had 45 hours over 15 weeks and P2 37.5 hours over 12 weeks
Edmonds, Nadeau & Kiran (2009)	case series (n=4) (2 TMA, 2 conduction)	to establish if VNeST is effective..	VNeST	unclear but appears to vary from 4 - 6 weeks, twice a week for total of 3 hours. (min = 12 hours – max 18 hours)
Furnas & Edmonds (2014)	case series n=2 (aphasia type not specified)	to investigate the effectiveness of VNeST when delivered by computer.	VNeST C	2 hour sessions x3 a week for 8 weeks (total = 48 hours)
Fink, Schwartz, & Myers (1998)	case series (n=9) (study 1)	to explore the effectiveness of a simplified version of mapping treatment (Schwartz et al., 1994): Sentence Query Approach.	mapping treatment (Schwartz et al., 1994): Sentence Query Approach	not reported
Goral & Kempler (2009)	single case study	to explore the effectiveness of Constraint Induced Language Treatment (CILT).	modified CILT	4 x 75 minute sessions per week for 4 weeks (total = 40)
Helm Estabrooks & Ramsberger (1986)	group study n=6	to establish the effectiveness of the Helm Elicited Language Program for Syntax Stimulation (HELPSS) programme.	HELPSS	24-113, 30 minute sessions (mean 80) (total 12 hours - 56.5 hours)
Hoover, Caplan, Waters &	group study n=12	to compare the effectiveness of VNeST treatment when delivered 1) individually, 2) in a group context or 3) combined	VNeST treatment delivered individually, in a	6.75 hours, over two days, for 6 weeks (2.25 hours a week for each treatment

Budson, (2015)			group context or combined	condition: total = 13.5 hours)
Jones (1986)	single case study	clinical case study of mapping treatment.	mapping treatment	not specified (clinical case study) but c9 months of treatment, x3 sessions a week (c108 sessions)
Kempler & Goral (2011)	case series n=2	to compare the effectiveness of drill vs communication based CILT.	drill vs communication based CILT.	x2 phases of 30 hours over a 4 week period c7.5 hours per week (total 60 hours over 8 weeks)
Le Dorze, Jacob, & Coderre (1991)	single case study	replication of Jones (1986) clinical case study.	mapping treatment	45-60 minutes, 4-5 times per week, for 1 month (min 12hours – max 20hours)
Links, Hurkmans & Bastiaanse (2010)	case series + group results n=11	to further explore the effectiveness of the VWS treatment used in Baastianse et al., (2006)	Verb Production at the Word and Sentence Level (VWS) treatment program	30 minutes x3 weekly for 12 weeks (total 18 hours)
Marshall, Chiat & Pring (1997)	single case study (Wernicke's aphasia)	to report a treatment program for a selective verb deficit involving difficulty mapping thematic roles.	mapping treatment	2 x 1 hour sessions per week for 6 weeks (total = 12 hours) plus homework
Maul, Conner, Kempler, Radvanski & Goral (2014)	case series n=4	to assess the effectiveness of CILT.	modified CILT	7.5 hours per week, over 3-4 days, for 1 month (total = 30 hours)
Mitchum, Greenwald & Berndt (1997)	single case study (fluent, mild anomia)	to describe a treatment programme for a mapping deficit specific to production	mapping treatment	18 2 hour sessions (total = 36 hrs)
Newton, Kirby & Bruce (2017)	case series n=2	to assess the effectiveness of shape coding treatment	shape coding treatment	1 hour, x2 weekly for 4 weeks (total = 8 hours)
Nickels, Byng & Black (1991)	single case study	to replicate the mapping therapy carried out by Byng (1988).	mapping treatment	2 phases, each of 6 weeks. x2 weekly sessions, c1.5 hours per session (total treatment = c36 hours)
Park, Goral, Jerkuilen & Kempler, (2013)	case series n=3	to investigate the effect of phonological and conceptual relatedness of verbs to nouns on response to treatment.	CILT (See Goral & Kempler, 2009)	2.5-3 hours, 3 times a week for 4 weeks (min 30 hours – max 36 hours)

Rochon, Laird, Bose, & Scofield (2005)	case series n=3	to elucidate if mapping therapy is effective when the emphasis is on production (as opposed to comprehension)	mapping treatment	average of 19 hours, c1 hour, twice weekly, c2.5 months
Rochon & Reichman (2003)	single case study (mixed aphasia: fluent/non-fluent)	to investigate the effectiveness of i) verb retrieval treatment and ii) sentence treatment (sentence treatment only reported)	grammatical frame & mapping treatment	14 x 1 hour sessions (total =14 hours)
Schneider & Thompson (2003)	case series n=7	to compare the effectiveness of i) semantic verb retrieval treatment and ii) verb argument structure treatment (treatment ii only reported here)	verb argument structure treatment.	12 sessions
Schwartz, Saffran, Fink, Myers & Martin (1994)	case series n=8	to explore the effectiveness of mapping treatment	mapping treatment	48-72 hours, in sessions of 60-90 minutes, x3 weekly
Silagi, Hirata & De Mendonca (2014)	single case study	to assess the effectiveness of HELPSS (Helm Estabrooks & Ramsberger, 1986).	HELPSS (see Helm Estabrooks & Ramsberger, 1986)	30 x weekly 30 minute sessions (total = 15 hours)
Takizawa, Nishida, Ikemoto, & Kurauchi (2015)	case series + group results n=6	to assess the relative effectiveness of i) single word (SW) and ii) sentence (S) treatment (S treatment only reported here)	sentence treatment = PAS	1-5, 40-minute sessions per week over, 2–8 months, plus homework (min = 5.3 hours - max = 106.6 hours
Webster, Morris, & Franklin (2005)	single case study	to explore the effectiveness of treatment for verb retrieval and PAS.	PAS treatment	x5 weekly session of 45 minutes, for 10 weeks (total = 37.5 hours) (+ break of 4 weeks when home practice took place)
Webster & Gordon (2009)	single case study	to explore why PAS treatment was not effective whilst verb-noun association treatment was in a single case study.	verb-noun association treatment	2 phases of 4 weeks, x2 weekly, 45 minute sessions (total =12 hours)
Whitworth, Webster & Howard (2015)	single case study	to assess the success of treatment for a PAS deficit.	PAS treatment	2 phases of 5 weeks, x2 weekly, c1 hour sessions (total= 20 hours) plus homework

1. What are the levels of evidence for sentence treatments for people with aphasia?

There are a number of published guidelines which assist researchers and clinicians to determine the level of the evidence in support of a particular treatment. This review used the Oxford

Centre for Evidence-Based Medicine (OCEBM) revised Levels of Evidence (2011). Systematic reviews of RCTs and of n of 1 trials are considered the highest level of evidence because they are designed to be unbiased and have less risk of systematic errors. However, cohort studies and case series (Levels 2-4) also provide evidence to motivate treatment choice if they are well controlled. Consequently, all studies graded level 1-4 were considered to be suitable to be included in this review.

In terms of the level of evidence for sentence treatments, it was predominantly Level 4 comprising 14 case series, 12 single case studies, five case series with group results reported, and two group studies. The total number of participants in the reviewed studies was 126 (see Table 1, column 2) **with the largest number of participants in a (group) study being 12 (Hoover et al., 2015).**

2. What is the fidelity of aphasia research investigating sentence production treatments in terms of treatment delivery, receipt, and enactment?

In terms of *treatment delivery,* this is enhanced by the existence of a treatment manual or tutorial (Hinckley and Douglas, 2013). VNeST, investigated in five studies (Edmonds, Mammino & Ojeda, 2014; Edmonds & Babb, 2011; Edmonds, Nadeau & Kiran, 2009; Furnas & Edmonds, 2014; Hoover et al., 2015), has a published tutorial containing a very detailed description of the treatment protocol (Edmonds, 2014). The Helm Elicited Language Program for Syntax Stimulation in aphasia (HELPSS) investigated in two studies (Helm-Estabrooks & Ramsberger, 1986; Silagi, Hirata, Iracema & de Mendonça, 2014) has a published manual[7]. No other treatments reported the existence of a treatment tutorial or manual. However, 30 (90%) of the reviewed papers were judged to describe treatment in sufficient detail to enable replication of the treatment protocol. The exceptions were

[7] An updated version of the HELPSS is available via the following link: https://www.proedinc.com/Products/9085/sentence-production-program-for-aphasia-formerly-the-helpss-program.aspx

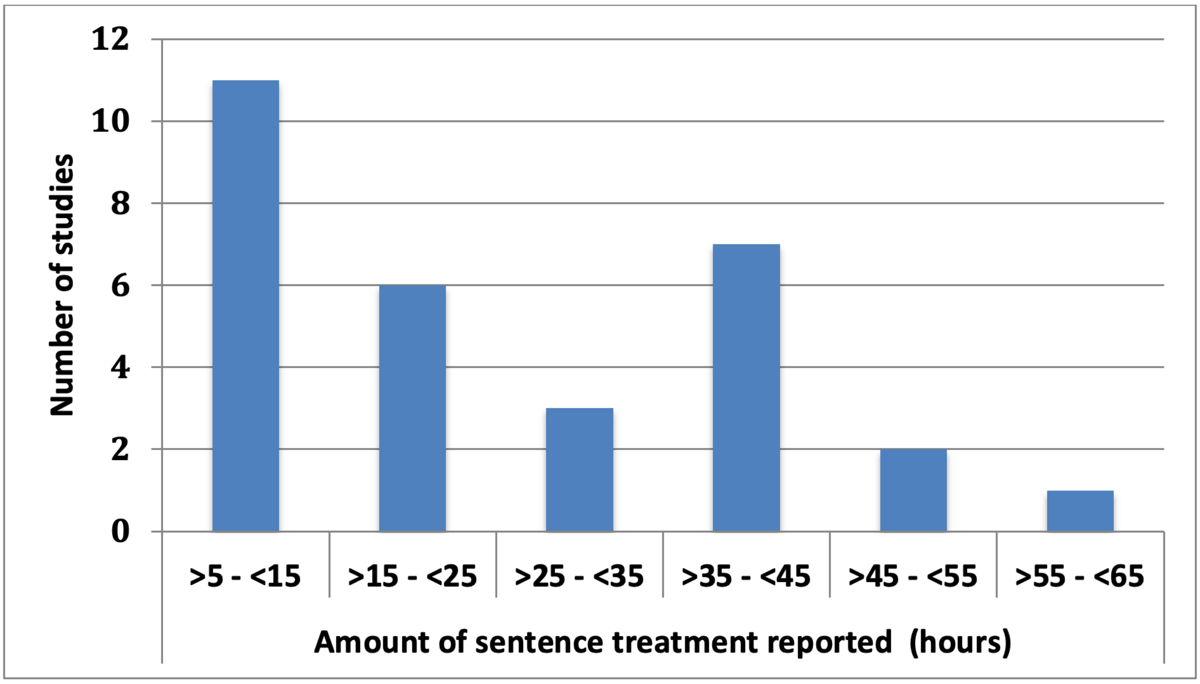
Bazzini et al. (2012), Davis and Tan (1987) and Fink, Schwartz and Myers (1998). (See Table 1, column 4: Type of Treatment; column 5: Amount of Treatment; Table 2, column 2: the Ingredients of Treatment; column 3: Number of Verbs/Sentences Treated and Type).

Regarding *dose* of sentence treatments, 23 studies (70%) reported the exact amount of treatment given (see Table 1 Column 4). Six studies (18%) reported the minimum and maximum amount of treatment given (or gave sufficient detail to allow this to be calculated). These studies were: Edmonds et al. (2009: 12-18 hours), Helm-Estabrooks and Ramsberger (1986: 12-56.5 hours); Le Dorze, Jacob and Corderre (1991: 12–20 hours); Park, Goral, Verkuilen, and Kempler (2013: 30–36 hours); Schwartz, Saffran, Fink, Myers and Martin (1994: 48-72 hours) and Takizawa, Nishida, Ikemoto and Kurauchi (2015: 5.3-106.6 hours). Rochon, Laird, Bose and Scofield (2005) reported the average amount of treatment given (19 hours) and Schneider and Thompson (2003) reported the number of treatment sessions (12) but not their length. Neither Fink et al. (1998) nor Jones (1986) reported on amount of treatment given. In summary, dose varied greatly, with the minimum being 5.3 hours and the maximum 106.6 hours (both reported in Takizawa et al.'s clinical study). Five to 15 hours of treatment was the most commonly reported (in 11 (33%) of studies), with 15 to 25 hours reported in six studies, 25 to 35 hours in three studies, 35 to 45 hours reported in seven studies, and more than 45 hours in three studies (see Figure 2).

In terms of *treatment receipt,* no studies of sentence treatment reported the *views* of participants about their treatment. In terms of *home practice*, six studies reported participants carried this out (Bastiaanse, Hurkmans & Links, 2006; Carragher, Sage & Conroy, 2015; Marshall, 1997; Takizawa et al., 2015; Webster, Morris & Fanklin, 2005; Whitworth, Webster & Howard, 2015) but none reported how *much* home practice was completed.

Figure 2. The number of hours of sentence production treatment reported in the reviewed studies. (Schneider and Thompson (2003), Fink et al. (1998), and Jones (1986) are not included, as the

amount of treatment is not reported. For studies that reported a minimum and maximum amount of treatment, the minimum amount of treatment is reported in the graph).



In terms of *treatment enactment*, six studies (18%) reported this was assessed. Kempler and Goral (2011) and Maul, Conner, Kempler, Radvanski and Goral (2014) discussed treatment with clinicians and then assessed treatment enactment by direct observation of sessions. Schneider and Thompson (2003) videoed four random sessions and assessed these, finding 100% adherence to the treatment protocol. Edmonds and colleagues assessed adherence to the VNeST protocol in three studies by observing 25% of sessions either live or via video. They rated adherence to the protocol at over 95% in each study (Edmonds et al., 2014; Edmonds & Babb, 2011; Edmonds et al., 2009).

3. What is the evidence of positive gains for sentence production treatments for people with aphasia in terms of improved production of a) trained and untrained verbs in isolation and b) trained and untrained verbs in sentences?

The evidence for the **efficacy** of sentence treatment on the production of a) trained and untrained verbs in isolation, and b) on the production of sentences using trained and untrained verbs is summarized in Table 2 Columns 4 - 7. Significance levels and effect sizes are given for each individual participant when available.

Table 2. Sentence treatments: ingredients and impact of treatment at each level of communication (Note: ANELT = Amsterdam–Nijmegen Everyday Language Test; MLU = mean length of utterance; P = patient; Tx = treatment; ANOVA = analysis of variance; HELPSS = Helm Elicited Language Program for Syntax Stimulation; CETI = Communicative Effectiveness Index; ASHA FACS = American Speech-Language-Hearing Association Functional Assessment of Communication Skills for Adults).

1.Study	2.Ingredients of Treatment	3.Trained Verbs in Isolation	4.UnTrained Verbs in Isolation	5.Trained Verbs in Sentences	6.Untrained Verbs in Sentences	7.Functional Communication	8.Discourse
Bastiaanse, Hurkmans, & Links (2006) (n=2)	VWS treatment: **picture prompts; sentence completion** using **written sentence frame; reading aloud** of sentence; **production of verb morphology;** moving verb within sentence; **production of verb;** with **external argument**; **internal argument/adjunct**; **sentence anagram cards, semantic, written, phonemic & repetition** cues given if needed (**13**)	not assessed	not assessed	not assessed	**Mr M sig improvement for finite verbs only: Chi Square p***, sig improvement in sentence construction on the Aachen Aphasia Test; Mrs F sig improvement for finite verbs only: Chi Square p*; sig improvement in sentence construction on the Aachen Aphasia Test;**	**Mr M: sig improvement on the ANELT: P*; Ms F: sig improvement on the ANELT: p*;**	semi-structured interview: **Mr M sig improvement in MLU and proportion of finite verbs: t**-***; Mrs F improvement in production of lexical verbs & sig improvement in MLU: t*****
Bazzini et al., (2012) (n=8)	VAS treatment: **hierarchy of sentence structures** treated: SV -< SVO ->SVO + complement or adjunct; drilled sentence completion involving **stepped production** of **verb;external argument; internal argument/s; adjunct,**; **emphasis on speed of production**; gender and number agreement within the NP; production of prepositions within the PP; **verbal explanations of thematic role** of sentence constituents given; (**10**)	not assessed	**significant improvement for group on naming test of 50 verbs: p** ;**	**significant improvement for group: p**; significant improvement for 7/8 Ps**	**significant improvement for group: p**; significant improvement for 5/8 Ps**	not assessed	picture description and personal narrative: **significant improvement for group: p*; significant improvement for 6/8 Ps**
Biran and Fisher (2015) (n=2)	PAS treatment: with **instruction phase** that focussed on the action represented by a verb, the role and number of arguments taken by the target verbs; **ID arguments in sentence**; **sentence hierarchy; stepped production** of **external argument**; **verb**; **internal argument**; **shape cues**; **wh-question cues for thematic role**; **written cues**; **written production** followed by **spoken production** (**8**)	not assessed	not assessed	**significant improvement for HY p**; significant improvement for AB p****;**	**significant improvement for HY p****; significant improvement for AB p****;**	not assessed	**significant improvement for HY in story telling P*;** AB no significant change

Byng, Nickels & Black (1994) (n=3)	mapping treatment: **picture prompts, colour coding; sentence frame; written cue cards; identify thematic roles; identify the part of the sentence changed; sentence hierarchy** & **cueing hierarchy used; construction of written & spoken sentence; production of verb, internal** & **external arguments** & **adjuncts; questions used** to **encourage self monitoring** rather than corrective feedback**; generalisation phase** using **personally relevant/functional material;** (**18**)	not assessed	**significant improvement for all 3 Ps: AER: p**; EM: p**; L.C. p***	not assessed	not assessed	not assessed	**significant improvement for all 3 Ps** (Cinderella narrative)**: AER: p****; EM: p**; LR: p*****
Carragher, Sage & Conroy (2015) (n=9)	combination Tx**: picture prompt** of black and white line drawings**; hierarchy of sentence structures; comprehension task; written cues; colour coding, sentence frame**; **production of light** as well as **heavy verb** encouraged; **production of external;** & **internal arguments**; & **adjuncts**; **production of PR agent encouraged; repetition; gesture encouraged; phonemic cueing;** (**15**)	not assessed	not assessed	**significant improvement for group: ANOVA p***; significant improvement for 8/9 participants** (not GL)	**significant improvement for group: ANOVA p***; significant improvement for 8/9 participants** (not GL)	not assessed	**story telling**: **numerical improvements in proportion of VPs,** not statistically significant for group; **numerical improvements in VPs reported for 6/9 participants,** no statistical analysis; **conversation:** no significant changes for group; **numerical changes reported for 5/9 participants,** no statistical analysis
Davis & Tan (1987) (n=1)	sentence stimulation treatment: **picture prompt; repetition**; use of **wh questions** to stimulate sentence constituents; **spoken production of verb; external**; & **internal arguments**; **backward chaining** to support sentence production (**7**)	not assessed	not assessed	**visual inspection of graphs indicates improvement in treated sets, plus %age improvement reported**	visual inspection of graphs indicates no improvement i in untreated sets (crossover design)	not assessed	not assessed
Edmonds, Mammino & Ojeda (2014) (n=11)	VNeST**: written cues**; ;**written;** & **spoken Wh- question** prompts to elicit thematic roles; **spoken production of verb; external; & internal argument; spoken production of personally relevant agent/theme encouraged**; **reading aloud** of **target agent-verb-theme**; **wh questions to stimulate production of adjunct; sentence judgement task**; **spoken production of verb-in-isolation; spoken production of target sentence with no cues**; (**12**)	not assessed;	**significant improvement for group: p**; significant improvement for P4, P5, P8, P10, P11: > 2 SDs**	**significant improvement for group: p**; significant improvement for P1, P4, P6,: large ES; P3, P5, P7, P10: medium ES; P2, P11: small ES**	**significant improvement for group: p**; significant improvement for P1, P2, P4, P5: large ES; P6, P9: medium ES; P7, P8, P10, P11: small ES**	**significant improvement for group: p**; significant improvement for 9/9 Ps** (carer rating of CETI)	Nicholas & Brookshire's (1993) discourse tasks: **significant improvement for group: % complete utterances: p*; % CIUs: p*; significant improvement for P1, P3, P5, P8, P9, P10: % complete utterances - large /medium ES; improvement in %CIUs: P1, P5, P8, P9, P11: >2 SEM**

Edmonds & Babb (2011) (n=2)	VNeST as for Edmonds, Mammino & Ojeda (2014) without production of verb-in-isolation but plus **written sentence production** accepted for P2 (**11/12**)	not assessed	**significant improvement for P2: p******	**significant improvement for P1: d= 5.73 and for P2 d=10**	**significant improvement for P1: d= 3.86 and for P2 d= 5.66**	**significant improvement on the CETI for P1 and for P2: p***	Nicholas & Brookshire's (1993) discourse tasks: **numerical improvements reported for P1**
Edmonds, Nadeau & Kiran (2009) (n=4)	VNeST as for Edmonds, Mammino & Ojeda (2014) without production of verb-in-isolation (**11**)	not assessed	not assessed	**numerical improvements reported for all 4 Ps**	**numerical improvements reported for all 4 Ps**	not assessed	Nicholas & Brookshire's (1993) discourse tasks: **numerical improvements in production of complete utterances for P1, P2 & P3**
Furnas & Edmonds (2014) (n=2)	VNeST adapted for computer delivery: As for Edmonds et al (2014) but **written cues** presented on computer screen; only **written Wh- question** prompts to elicit thematic roles; addition of spoken **semantic cues** ; **orthographic; typed sentence production also allowed** (**13**)	not assessed	P1 and P2 : No significant improvement	**Significant improvement for P1: d=8.29 and P2: d=2.91**	P1 and P2: No significant improvement	not assessed	Nicholas & Brookshire's (1993) discourse tasks: **P1 & P2: numerical improvements in complete utterances**
Fink, Schwartz, & Myers (1998) (n=9)	mapping treatment - sentence query approach**: sentence hierarchy; wh-questions** to identify **thematic roles** first **agent** then **theme**, then **adjunct**; **icon cues paired with wh-questions; corrective feedback given** (**4**)	not assessed	not assessed	**Numerical improvements >20% for 5 Ps for at least 1 sentence type**	none	not assessed	not assessed
Goral & Kempler (2009) (n=1)	CILT**:** which emphasised the production of verbs within **informative exchanges; functionally relevant material; output restricted to verbal only** - **barrier tasks** used; **picture prompt**; **spoken production of verb; external** & **internal argument; sentence repetition; reading aloud; picture sequences; story generation; scripted phone calls; video retell: conversation task** (**14**)	not assessed	not assessed	not assessed	not assessed	**significant improvement** p** (bespoke questionnaire)	personal narratives: **significant improvement in total number of words: d=8** and **%age of verbs: d=8.2**
Helm Estabrooks & Ramsberger (1986) (n=6)	HELPSS: **picture prompt; story completion; +/- a spoken sentence prime;** followed by a **spoken prompt question; spoken production of verb; external** & **internal argument** (**7**)	not assessed	not assessed	not assessed	**significant improvement for group: p****	not assessed	**significant improvement for group on Cookie Theft picture description: content units: p**; grammatical morphemes: p***

Hoover, Caplan, Waters & Budson, (2015) (n=12)	Adapted version of VNeST treatment: individual Tx: **written cues; spoken production of verb; external; internal argument;** & **adjunct; wh- question prompt for adjunct**; **repetition;** group treatment involved **conversation/discussion; language games; functional scripts** (**10**)	**significant improvement for group: p*** **	**significant improvement for group: p** **	not assessed.	**significant improvement for group: p***	**significant improvement for group: on the ASHA-FACS: p* (& the ALA: p*)**	Nicholas & Brookshire's (1993) discourse tasks: **significant improvement for group in number of complete sentences: p***
Jones (1986) (n=1)	mapping treatment: **output discouraged in early treatment**; **written sentence cue; comprehension tasks focused on verb identification; wh-questions** to identify **thematic roles first agent** then **theme**, then **adjunct**; **chart of Wh question words & relationship to verb; syntactic hierarchy; sentence judgement task with written;** & **spoken sentences; picture description; story telling; conversation based tasks** (**11**)	not assessed	not assessed	not assessed	not assessed	not assessed	**improvement in picture description, & personal narrative reported by pre and post treatment language samples; improvement in spontaneous output reported anecdotally**
Kempler & Goral (2011) (n=2)	CILT: drill Tx: **picture prompts, spoken** & **written cues; hierarchy of sentences; hierarchy of cues; production of verb; external; internal argument;** & **adjunct; repetition; choral reading; picture description; map task; memory task** (**14**) communication Tx: predicated on **exchange of novel information: picture prompts; hierarchy of sentences; hierarchy of cues; production of verb; external; internal argument;** & **adjunct; barrier tasks; picture description; Go Fish; memory task; map task; picture sequence description; story construction** (**15**)	**numerical improvement for P2 after drill Tx**	none	not assessed	not assessed	not assessed	personal narrative production: **significant improvement after drill Tx only**: **P1: ES=5.95 p***; P2: ES=11.16 p****
Le Dorze, Jacob, & Coderre (1991) (n=1)	mapping treatment: modification of Jones (1986) (because P had impaired reading): **output discouraged**; **picture prompt for verb; comprehension tasks focused on verb identification; picture prompts for sentence constituents** to identify **thematic roles first agent** then **theme**, then **adjunct; syntactic hierarchy; "sentence" judgement task** with picture sequence to identify missing constituent**;** (6)	not assessed	not assessed	**significant improvement: p *****	**yes** (no statistical analysis)	not assessed	**picture description**: **significant improvement: p ***

Links, Hurkmans & Bastiaanse (2010) (n=11)	VWS treatment: see Bastiaanse et al., (2006) (**13**)	not assessed	not assessed	not assessed	**significant improvement** for the **group on finite verbs p** and infinitives p**; improvement in infinitive verbs** for **1 individual; improvement in finite verbs** for **5 individuals** (no statistical analysis)	ANELT: **significant improvement for the group: P****; improvement for 10/11 individuals** (1 individual had a likely ceiling effect)	semi-standardised interview: **significant improvement for the group: P**; improvement for 9/11 individuals** (no statistical analysis)
Marshall, Chiat & Pring (1997) (n=1)	mapping treatment: **picture prompts for thematic roles**; **colour coding; written** & **spoken sentence cue**; **identify thematic roles; movement of theme card; feedback emphasizing relationship between thematic roles & syntactic structure; syntactic hierarchy; hierarchy of cueing; spoken production of verb;** with **external;** & **internal argument;** using **sentence frames;** & **sentence completion; conversational opportunities used**; (**15**)	not assessed	not assessed	**numerical improvement reported for treated verbs**	**significant improvement for untreated exemplars of same verb class: p***; no significant improvement for untreated thematically similar verbs or dissimilar verbs	bespoke tasks: **significant improvement for treated sentences: p***	Story retell: **significant improvement: for treated verbs: p*;** untreated not assessed
Maul, Conner, Kempler, Radvanski & Goral (2014) (n=4)	CILT: Tx predicated on exchange of novel information: use of **barrier tasks; massed practice; picture prompts; spoken sentence modelled;** & **spoken cues for fuller sentence; shaping of target sentence; sentence repetition; spoken production of verb;** with **internal** & **external argument**; language games**: Go Fish; memory; picture sequence description; story construction; map task; feedback given** (**16**)	not assessed	not assessed	**significant improvement for 2/4 Ps: P3 & P4; p***	**significant improvement for 1/4 Ps: P1: p*;** no change in untrained tasks (picture sequence description and responses to wh- questions) for any P	not assessed	not assessed
Mitchum, Greenwald & Berndt (1997) (n=1)	mapping treatment: **active & passive versions of target sentence contrasted**; **written;** & **spoken sentence cues**; **sentence anagram cards**; **sentence hierarchy; spoken production of verb;** with **internal;** & **external argument;** (**8**)	not assessed	not assessed	**significant improvement p*****	not assessed	not assessed	not assessed
Newton, Kirby & Bruce (2017) (n=2)	shape coding treatment: **spoken production of verb** with **external argument; internal argument**; & **adjunct**; **simple picture prompts; colour** & **shape cues; written cues; sentence frames; syntactic hierarchy; cueing hierarchy; contrastive drills; composite picture cues; picture sequence cues; video cues; conversation tasks; sentence judgement task; personally relevant material;** homework: **written sentence production tasks** (**19**)	not applicable	**significant improvement for 1 participant: TW: p*****	not applicable	**significant improvement for 1 participant: AS: p***	no significant improvement on the ANELT	no significant improvement

Nickels, Byng & Black (1991) (n=1)	mapping treatment: stage 1: comprehension: **picture prompt; colour coding; written cue cards; sentence frame; cueing hierarchy; sentence hierarchy; reading aloud of target sentence; wh-questions** & **feedback highlighting the thematic roles; self-monitoring encouraged;** stage 2: production**: sentence ordering task; spoken production of verb; with external; & internal argument; phonological cue for verb; production of proper (functionally relevant) nouns** for subject/object encouraged**; (16)**	not assessed	**significant improvement : p****	not assessed	**significant improvement: p******	not assessed	**significant improvement: p****** (Cinderella narrative)
Park, Goral, Jerkuilen & Kempler, (2013) (n=3)	CILT (see Kempler & Goral, 2011: communicative treatment) (**15**)	**significant improvement for 1 participant: P2: p****	no significant improvement	not assessed	not assessed	not assessed	not assessed
Rochon, Laird, Bose, & Scofield (2005) (n=3)	mapping treatment: **picture prompt; icons** used to identify thematic roles in sentence; plus **verbal explanation**; **written cues; spoken production of verb;** with **external** & **internal argument; cueing hierarchy; correct sentence modelled; corrective feedback** given (**10**)	not assessed	not assessed	**significant improvement for all 3 Ps: SM: p***; QO: p**; NS: p***** (novel exemplars of treated structures)	no significant improvement reported	not assessed	**yes numerical improvements for all 3 Ps** (no statistical analysis)
Rochon & Reichman (2003) (n=1)	verb & mapping treatment: verb treatment: **picture prompt; written;** followed by **spoken verb production; written cues; (4);** mapping treatment: **picture prompt; sentence frame; wh question prompts; spoken** followed by **written production of verb; external;** & **internal argument;; reading aloud; corrective feedback given; passive** as well as active **sentence production; production of verb morphology; sentence judgement of own (spoken) sentence production;** (**12**)	not assessed	not assessed	**significant improvement: p***	no (significant deterioration (of active sentences: p*)	not assessed	video retell: **yes numerical improvement on 1 measure (lexical: nonlexical verbs)**

Schneider & Thompson (2003) (n=7)	verb argument structure treatment: **picture prompt; spoken definition** of **argument structur**e of verb; **spoken sentence production of verb**; with **external;** & **internal argument**; **correct spoken sentence modelled** for **repetition** (**7**)	**significant improvement for the group: p: *****	no significant improvement for untrained verbs within either semantic or syntactic category (partial generailsation for Ps 3 & 4)	**significant improvement s for the group: p: ****	for the group untrained sentences improved significantly less than trained * HOWEVER **statistically significant improvement on NWVPB**	not assessed	**Cinderella narrative showed some improvements in production of grammatical sentences & verb arguments**: not statistically significant
Schwartz, Saffran, Fink, Myers & Martin (1994) (n=8)	mapping treatment: **reading aloud** of target sentence with **spoken model** if required; **wh questions** to **identify thematic roles; colour coding**; **syntactic hierarchy; picture prompts; spoken & written cues; sentence judgement; video retell** (**11**)	not assessed	not assessed	numerical improvement for 5/7 (type A sentences following type A Tx)	numerical improvement for 3/7 (type B sentences following type A Tx)	not assessed	Cinderella narrative or picture description: **improvement for 4/6 participants** (not IC or JH) in (no statistical analysis)
Silagi, Hirata & De Mendonca (2014) (n=1)	HELPSS: see Helm Estabrooks & Ramsberger (1986) above (**7**)	not assessed	not assessed	**improvement reported (no analysis)**	not assessed	not assessed	**significant improvement on (Cookie Theft) picture description: p***
Takizawa, Nishida, Ikemoto, & Kurauchi (2015) (n=6)	combination sentence treatment: **picture prompt**; **spoken model of verb; identify thematic roles; corrective feedback given; sentence frame** to cue **spoken production of verb;** with **external;** & **internal argument; oral** or **written model of sentence** if required; **written homework** (**11**)	**significant improvement for 6/7 participants** (not P1) significance levels not reported	**significant improvement for 2/7 participants (P5 & P6: p*)**	not assessed	not assessed	not assessed	personal narrative & picture description, no significant improvements in production of grammatical sentences, verb retrieval or MLU for the group; **1/6 Ps sig improvement in MLU***
Webster, Morris, & Franklin (2005) (n=1)	PAS treatment: **picture prompt; written;** & **spoken semantic (comprehension) tasks**; **spoken verb production;** with **repetition** if required; **written noun-verb association task for agent; & theme; sentence frame; spoken production of external**; & **internal arguments** & **adjunct; spoken sentence production; discussion with therapist re completeness of sentence produced** & **role of arguments;** (**14**)	**significant improvement : p*****	none	not assessed	**significant improvement: p ****	not assessed	Cinderella narrative **numerical increase in 2 argument structures** (no statistical analysis)

Webster & Gordon (2009) (n=1)	verb-noun association treatment: **picture prompt; written cues; reading aloud of verb**; **comprehension task** (select associated noun**);** with **feedback; spoken sentence production of verb; with external; & internal argument**; correct **spoken sentence model** if required (**9**)	**significant improvement : p****	none	**significant improvement: p********	none	not assessed	not assessed
Whitworth Webster & Howard (2015) (n=1)	PAS treatment; **picture prompt; written cues; reading aloud of verb;** & **sentence; wh- questions** to **identify thematic roles;** & their **position;** in a **sentence frame; generate 3 agents;** & **3 themes; feedback given; multiple; sentence generation** per **verb**; with **internal** & **external argument**; (**14**)	not assessed	not assessed	**significant improvement: p***	**significant improvement: p***	not assessed	not assessed
Percentage (raw number) of participants: significant change reported		71% (10)	36% (15)	83% (57)	59% (48)	89% (25)	72% (58)
Percentage (raw number) of participants: no significant change reported		**29% (4)**	**64% (27)**	**17% (12)**	**41% (34)**	**11% (3)**	**28% (23)**
Total number of participants individual results reported		**100% (14)**	**100% (42)**	**100% (69)**	**100% (82)**	**100% (28)**	**100% (81)**

The impact of treatment on the production of a) *trained* and *untrained verbs in isolation* was assessed in only 7 studies (21%) for trained verbs, and 14 (42%) of studies for untrained verbs (see Table 2, columns 4 and 5). The seven studies investigating impact of treatment on production of *trained* verbs in isolation had a total of 33 participants. Individual results are reported for 14 of these participants with 10 individuals showing significant improvement (71%) (confirmed by statistical analysis for three of these participants: $p<.05$). Two studies reported group results with both finding significant improvement: (Hoover et al., 2015 (n=12); Schneider & Thompson, 2003 (n=7)). With regard to the production of *untrained* verbs in isolation, this was investigated in 14 studies with 62 participants. Individual results are reported for 42 of these participants, 15 of whom (36%) showed significant improvement (confirmed by statistical analysis for eight of these participants: $p<.05$). Significant improvement was also reported in two group studies: Bazzini et al. (2012) (n=8) and Hoover et al. (2015) (n=12).

The impact of sentence treatments on b) *sentence* production involving either trained or untrained verbs is summarised in Table 2 columns 6 and 7. Sentence production involving *trained* verbs was assessed in 20 (67%) studies which involved 75 participants, with individual results reported for 69 of these. Significant improvement in sentence production was reported for 57 participants (83%) (confirmed by statistical analysis for 25 of these participants: $p<.05$ for 12 participants; effect sizes reported for 13 participants).

Sentence production using *untrained* verbs was assessed in 25 (75%) studies with a total of 110 participants. Individual results were reported for 82 participants with significant improvement reported for 48 of these (59%) (confirmed by statistical analysis for 21 of these: $p<.05$ for 10 participants; effect sizes reported for 11 participants). Evidence for generalisation of treatment effects to untrained sentences currently appears strongest for VNeST (reported for 16 of the 19 participants (84%)) confirmed by statistical analysis for 12 participants). Group results were reported in seven studies (total participants n=64) with all of these studies reporting significant improvement for the group. Four studies reported individual participant results alongside the group results (Bazzini et al., 2012; Carragher et al., 2015; Edmonds et al., 2014; Links et al., 2010). Twenty-nine of the 39 participants in these studies (74%) showed significant improvement in the production of sentences using untrained verbs (see Table 2, Column 6).

4. What is the evidence of positive gains for sentence production treatments for people with aphasia in terms of a) improved functional communication and b) improved production of discourse?

The impact of treatment on a) functional communication was assessed in only eight studies (24%) with a total of 42 participants (see Table 2 Column 8). A variety of measures were used to assess functional communication: the Amsterdam—Nijmegen everyday language test (ANELT: Blomert, Kean, Koster & Schokker, 1994) was used in three studies, the Communicative Effectiveness Index (CETI: Lomas et al., 1989) in two studies, the American Speech-Language Hearing Association Functional Assessment of Communication Skills for Adults (ASHA FACS: Frattali et al., 1995) in one study, and bespoke tasks in the remaining two studies. Individual results are reported for 28 participants, 25 (89%) of whom showed significant improvement (confirmed by statistical analysis for six of these participants). Three studies reported significant improvement for the group of participants (Edmonds et al., 2014 (n=11); Hoover et al., 2012 (n=12); Links et al., 2010 (n=11)). Of the studies that found a significant impact of treatment on functional communication, three used VNesT (Edmonds et al., 2014; Edmonds & Babb, 2011; Hoover et al., 2015), two used the Verb Production at the Word and Sentence Level (VWS) treatment program (Bastiaanse et al., 2006; Links et al., 2010), with CILT and mapping treatment respectively used in the two remaining studies (Goral

& Kempler, 2009; Marshall, 1997). The study of shape coding[8] treatment reported by Newton, Kirby and Bruce (2017) was the only study to report no significant impact.

The impact of sentence treatment on b) discourse was explored in 26 studies (79%) with a total of 106 participants (see Table 2 Column 9). The stimuli used to elicit discourse for assessment were diverse, ranging from picture description to conversation. The most popular elicitation protocol was story retell (used in eight studies). Outcome measures were also diverse, the most common being the number of complete utterances, the number/proportion of verb phrases, content information units and mean length of utterance. Individual results were reported for 81 participants with 58 participants (72%) demonstrating improvement, for 27 of whom this was confirmed by statistical analysis (p<.05 for 18 participants whilst effect sizes were reported for nine participants). When improvement was reported in a study but was not confirmed by statistical analysis this was because numerical improvements were reported in terms of the number (or percentage) of verbs or clauses produced for example, but these were not subsequently subjected to statistical analysis.

Eight studies reported group results (total n=70) with six of these studies reporting significant improvement (confirmed by statistical analysis in five of these studies). Five of the studies reporting group results also reported results for the participants as individuals (Bazzini et al., 2012 (n=8); Carragher et al., 2015 (n=9); Edmonds et al., 2014 (n=11); Links et al., 2010 (n=11); Takizawa et al., 2015 (n=6)). Improvement was reported for 29 of these 45 participants (64%).

Of the 22 studies that both investigated the impact of treatment on discourse *and* reported *individual* results, (see column 7 Table 2) eight studies used mapping treatment reporting improvements for 15 of the 17 participants. VNeST was used in four studies with improvements

[8] Shape coding is a treatment approach used with children with developmental language disorder (e.g. Ebbls, van der Lely & Dockrell, 2007) which was adapted for use with PwA by Newton et al. Shape coding uses both shapes and colours to represent the different syntactic elements of a sentence (including morphology), and it provides a visual frame for a sentence.

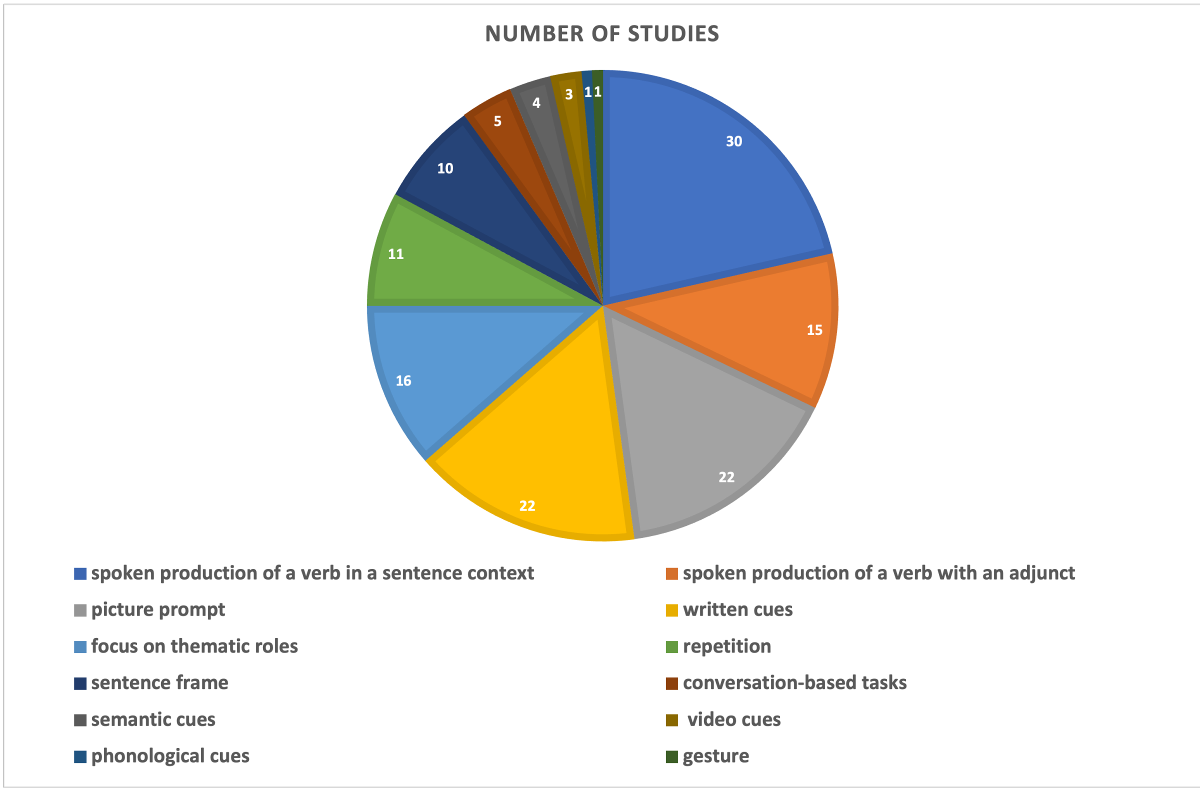
reported for 13 of the 19 participants. Verb/predicate argument structure was used in three studies with improvement reported for eight of the 11 participants. The less frequently employed treatments were: VWS treatment (with improvement reported for 11 of the 13 participants), a combination treatment (reporting improvement for seven out of 15 participants) and CILT (reporting improvement for all three participants). The remaining study used the HELPSS program and reported improvement for the single participant.

5. What are the potential active ingredients of treatments for sentence production deficits in aphasia?

In terms of the potential *active ingredients* of sentence treatments the *type* of treatment explored in the reviewed studies is given in Table 1 column 4. Ten studies (30%) explored mapping treatment, 6 (18%) studies investigated PAS treatment, 5 studies (15%) investigated VNeST, 4 studies investigated CILT, 2 studies investigated the VWS program, and 2 the HELPSS program. One study each investigated a particular approach: sentence stimulation (Davis & Tan, 10987), shape coding (Newton et al., 2017) and verb-noun association (Webster & Gordon, 2009). One study investigated a combination of ReST and mapping treatment (Carragher et al., 2015). Three studies also compared verb-in-isolation and sentence level treatments with the aim of elucidating which type of treatment was more effective (Rochon & Reichman, 2003; Schneider & Thompson, 2003; Takizawa et al., 2015). Rochon and Reichman found that only sentence level (mapping) treatment resulted in improvements in sentence and narrative production, whereas Takizawa et al. (2015) found the opposite pattern i.e. that verb-in-isolation treatment resulted in significant improvement in narrative production for the group as a whole whilst sentence treatment did not, although responses at the individual level were variable. Schneider and Thompson (2003) found no significant difference in the efficacy of verb-in-isolation (semantic) and sentence (PAS) treatments.

The treatment used in each study is described in more detail in Table 2 column 2 in order to try to identify the *individual* active ingredients within each *type* of sentence treatment. Potential active ingredients are emboldened in the text, and the number of ingredients given by the figure in brackets. The ingredients used in sentence treatments are also represented in a pie chart (Figure 3) which shows the number of studies that used a particular ingredient.

Figure 3. The number of studies that used an ingredient of treatment.



Sentence treatments are complex in that the number of ingredients in most treatments is large. The largest number of ingredients (19) was reported by Newton et al. (2017) in their shape-coding study (19), and the smallest number was four (Schwartz et al., 1994; Rochon & Reichmann, 2003: verb treatment component). The majority of studies (18 or 67%) used 11 or more ingredients. Figure 3 shows that the most common ingredient of sentence treatments was the *spoken production of a verb in a sentence context*, usually with an external argument (agent) and an internal argument (theme) (in 30 studies), with 15 studies also requiring *spoken production of a verb with an adjunct* (i.e. non-argument phrases providing additional information that is not necessary to complete the meaning of the verb). The next most common ingredients were the use of a *picture prompt* to stimulate sentence production and the use of *written* cues (in 22 studies respectively). The least used ingredients were *conversation-based tasks* (to facilitate generalisation of sentence production skill to real life communication) (5 studies), *semantic cue*s (4 studies), *video cues* (3) and the use of *phonological cues* and *gesture* (1 study each).

The spoken production of a verb in a sentence context was the most common ingredient of sentence treatments. However, the number of times this was required during a treatment session was only rarely reported. Kempler and Goral (2011) state that in the drill-based treatment phase of their sentence treatment study each of 32 verbs was practiced approximately 40 times during 30 hours of treatment. Rochon et al. (2005) state that six exemplars of each sentence structure were treated per session, and Edmonds and colleagues are specific that three to four agent-verb-theme exemplars have to be produced by a participant undergoing VNeST every time a verb is treated, although how many times a verb is treated during a treatment session is unclear (e.g. Edmonds & Babb, 2011).

In terms of the number of verbs/sentences treated in reviewed studies, the minimum number of verbs treated was six (Rochon et al., 2005) but in 144 exemplar sentences. The maximum number of verbs treated was 100 (Whitworth et al., 2015) in two sets of 50. The most common number of verbs to be treated was 10 in six studies (four of which investigated VNeST), with another three studies treating a larger number of verbs but in sets of 10 (30 in sets of 10: Davis & Tan, 18997; Schwartz et al., 1994; 40 in sets of 10: Schneider & Thompson, 2003).

Discussion.

The systematically conducted scoping review reported here synthesised and evaluated 33 studies of sentence production treatments for people with aphasia with a total of 126 participants and included 24 studies published since previous reviews were carried out (Conroy et al., 2006; Webster & Whitworth, 2011).

In summary, the systematic review reported in this paper found that whilst the reviewed studies predominantly represented Level 4 evidence, 83% of participants showing improved production of trained sentences and 70% of participants showing improved discourse production. Improvements in functional communication were also reported but this was not assessed frequently enough to make this a robust finding. Certain aspects of sentence treatments are under researched including the use of video and gesture cues in treatment, and delivery of sentence treatments by computer. In terms of the latter, given that people with aphasia increasingly rely on computer-based treatments (e.g. Kurland, 2014), and the emerging evidence that verb treatments can be effectively delivered in this way (e.g. Kurland, Liu & Stokes, 2018; Routhier, Bier & Macoir, 2016), it is imperative that computer-based sentence treatments are also explored. Lastly, whilst verb and sentence treatment studies generally reported treatment protocols to a good level of detail, researchers must improve the fidelity of reporting particularly with regard to the dose of treatment given, and with regard to the views of participants on sentence treatments which has not been investigated to date.

Levels of evidence for sentence production treatments.

Research question 1 addressed the levels of evidence for sentence production treatments in aphasia. The evidence was predominantly Level 4 as it was dominated by case series and single case studies. Whilst it is now acknowleged that well-controlled case series can be used to support clinical decision making, this should generally only be the case if high quality systematic reviews are not available (https://www.cebm.ox.ac.uk/resources/levels-of-evidence/explanation-of-the-2011-ocebm-levels-of-evidence). There is, therefore, a need for sentence production treatment studies using designs which constitute higher levels of evidence such as well-designed, larger scale RCTs (e.g. Palmer et al., 2019). However, it should also be noted that single case reports and case series *are* well suited to the current phase of sentence treatment research, which is predominantly Phase I and II (as defined by Robey and Schultz, 1998) in that sentence treatment protocols are still being refined and optimal dosages established, for example. The detailed description of treatment afforded by case reports and series (which is often omitted in RCTs) is also likely to be important in informing the sentence treatments used in such larger scale studies. Indeed, the significant role that case reports have played in advancing medical science in this way is now acknowledged (e.g. Murad et al., 2018).

Fidelity of sentence production treatments.

Fidelity of aphasia treatment is currently an important focus of research both in terms of implementation and reporting (e.g. Brogan et al., 2019; Conlon, Braun, Babbitt & Cherney, 2020; Dipper, Franklin, de Aguiar, Baumgaertner, Brady, Best et al., 2021). It is acknowledged that, historically, fidelity has been poorly addressed in aphasia research (e.g. Hinckley & Douglas, 2010) and that improving fidelity has the potential to improve the quality of aphasia treatment research in multiple ways. These include increasing the power of studies to detect treatment effects which may otherwise have been obscured due to variance (e.g. Spell, Richardson, Basilakos, Stark, Teklehaimanot, Hillis et al., 2020) and facilitating the implementation of research into practice because interventions will have been accurately described in terms of what treatment was *actually* given (e.g. Brogan et al, 2019; Kaderavek & Justice, 2013).

In terms of the fidelity of sentence treatments regarding *treatment delivery,* the majority of studies (90%) were sufficiently detailed to enable replication, however, use of manualized treatments was minimal. In terms of the reporting of treatment *dose* specifically, the exact amount of treatment was reported in 70% of studies. Increased accuracy of reporting treatment dose is vital if the optimal dose of treatment is to be determined (e.g. Baker 2012a & 2012b). It was also of interest that the dose of *sentence* treatment given tended to be larger than that for *verb-in-isolation* treatment as reviewed by Hickin et al. (2020). Thus, whereas 13 sentence treatment studies reported doses of more than 25 hours of treatment, this was the case for only four verb treatment studies, and in each of the latter treatment was self-delivered via computer. Any future research that aims to establish the relative efficacy of the two types of treatment must ensure that there is a level playing field in terms of the dose of treatment given. Thus, it is recommended that the minimum detail studies should report is dose, dose form, dose frequency, session duration, and total intervention duration as recommended by Baker (2012a & 2012b). Delivery of treatment by computer has the potential to record this level of detail accurately, as well as having the capacity to increase the dose of treatment given. Delivery of sentence treatment via computer is currently underexplored (in only one study: Furnas & Edmonds, 2014), and is an avenue of research that warrants more attention.

With regard to the fidelity of sentence treatment *receipt*, there was little indication of this being monitored in relation to sentence treatments, with the views of participants on their treatment not reported in any study. To elaborate, fidelity of treatment receipt includes demonstrating that a participant comprehends their treatment and can utilize the required skills (e.g. cognitive and communicative) during treatment sessions (e.g. Brogan et al., 2019; Conlon et al., 2020). Whilst there are currently few validated tools to monitor the fidelity of treatment receipt, the response recording sheet developed for VNeST (Edmonds, 2014, Appendix B) represents a way forwards here, as does the co-design (with people with aphasia) of a feedback questionnaire for computer-based aphasia treatment reported by Kearns, Kelly, and Pitt (2020). Finally, fidelity of treatment *enactment* was rarely reported, being evaluated in only six studies (18%). Studies of sentence treatments are not alone in this regard. Dipper, Franklin et al. (2021) conducted an umbrella review of the description of aphasia intervention within studies included in systematic reviews. Using the TIDieR checklist (Hoffmann et al., 2014) they reviewed 93 studies and found that reporting of fidelity measures was rare. Improving the fidelity of aphasia treatment research is important at many levels. As Brogan et al. (2019) put it, as a profession we cannot afford to conduct studies which are "under specified, under researched and under reported" (p.761) and exhort that greater attention be paid to both implementing and reporting fidelity procedures in aphasia treatment research. However, there are an increasing number of studies which demonstrate that it is indeed feasible to implement treatment fidelity measures, noting also that these must be planned and accounted for from initial study design (e.g. Carragher, Brooke, Worrall, Thomas, Rose, Simmons-Mackie et al., 2019; Conlon, Braun, Babbitt & Cherney, 2020; Hilari, Behn, Marshall, Simpson, Thomas & Northcott et al., 2019).

Evidence of treatment effects and generalization for sentence production treatments.

Conclusions regarding the impact of sentence production treatments on production of *trained verbs in isolation* must be treated with caution because this was assessed in only a small proportion of sentence treatment studies, with a similar pattern in relation to the assessment of *untrained* verb production in isolation. Having said this, sentence treatments were effective in improving the production of *trained* verbs in isolation for the majority (71%) of participants. This compares to improvement for 80% of participants in verb-in-isolation treatments (Hickin et al.,2020). Regarding production of *untrained* verbs in isolation, this improved for 36% of participants following sentence treatments compared to 15% of participants following verb-in-isolation treatments (ibid). This provides indicative evidence that sentence treatments may generalise to untrained verb-in-isolation production more successfully than verb treatments (within level generalisation). However, as stated earlier, because the number of sentence treatment studies that evaluated production of untrained verbs in isolation was small, this finding must be treated with caution. Given that sentence treatments are *ipso facto* investigating treatment involving the production of treated verbs simultaneously with other sentence constituents (i.e. not in isolation) it is understandable that verb production in isolation is not routinely assessed. However, if a research aim is to investigate the importance of verb retrieval *per se* to sentence production then assessing verb retrieval in isolation as well as in a sentence would seem important. In particular, routine assessment of verb retrieval in isolation (as is the case after verb in isolation treatment) would make comparison of the two types of treatment easier.

There are a number of challenges in interpreting the effect of sentence treatments on *sentence* production. These include the differing ways in which generalisation is assessed and how the *significance* of improvement is determined. Thus, with regard to within level generalization, there were two almost equally frequent ways of assessing this: i) assessing sentence production in an *untrained task* (such as a standardised sentence production test), and ii) assessing *untrained exemplars* of sentences (i.e. using matched but untrained *verbs*), with a minority of studies assessing generalisation via production of untreated sentence (syntactic) *structures* (6 studies). Interpretation of the evidence would be assisted by reaching a consensus on how to assess generalisation. The evidence would also be strengthened by the use of inferential statistics to evaluate the significance of treatment effects. Statistics were used to confirm treatment effects just under 50% of the time in relation to both trained and untrained sentence production. With these caveats in mind, sentence treatments appeared effective in improving *trained* sentence production (for 83% of participants) and *untrained* sentence production (for 59% of participants).

Broader generalization of sentence production treatments for people with aphasia.

Conclusions regarding the impact of sentence treatments on *functional communication* must be tentative because this was not routinely assessed. Nonetheless, 89% of the participants for whom individual results were reported improved (25/28 participants), and it is recommended that future research routinely assesses the impact of treatment on functional communication. VNeST currently provides the strongest evidence of the functional impact of treatment: for all 11 participants for whom individual results are reported. One possible reason for this may be that VNeST requires the production of functionally relevant verb-agent-theme exemplars during treatment with the explicit aim of both increasing the salience of treatment stimuli and of facilitating a functional impact of treatment. However, the amount of VNeST treatment and the intensity with which it was delivered was relatively high and so this may also have contributed to it having a functional impact.

Discourse production improved for 70% participants, with this finding strengthened by the frequent assessment of discourse production in sentence treatment studies (79%). The finding must, though, be tempered by the proviso that improvement was confirmed statistically for less than 50% of participants (as was the case for generalisation of treatment effects to sentence production). Interpretation of the evidence is also impeded by the use of diverse outcome measures. The most common means of sampling discourse was story retell, with only one study using real life conversation as the context for measuring discourse outcomes (Carragher et al., 2015). However, the best way to assess discourse remains subject to debate (e.g. Bryant, Ferguson & Spencer, 2016).

The potential active ingredients of treatments for sentence production deficits in aphasia.

Key to establishing what contributes to the efficacy of sentence treatments is identifying the active ingredients of treatment. The review found that the most common ingredient of sentence treatments was the spoken production of a verb in a sentence context (in 91% of studies) making it a likely active ingredient. However, this task varied considerably in the way it was implemented. For example, some studies allowed and even encouraged the use of pronouns (as well as full lexical forms) in target sentences with the aim of increasing the variety of utterances attempted and making these potentially more functional (e.g. Carragher et al., 2015; Nickels et al., 1991). Other studies specifically discouraged the production of pronouns, notably those investigating VNeST wherein the rationale is that the treatment is predicated upon strengthening (priming) the semantic network of a verb by requiring production of a verb alongside its arguments, and the production of pronouns circumvents this process (see e.g. Edmonds 2016).

Other common ingredients of sentence production treatments were the use of a *picture prompt* to stimulate sentence production, the use of *written* cues and a *focus on the thematic roles* of a verb's arguments during treatment. The use of a picture prompt to stimulate sentence production varied across studies. Edmonds and colleagues specifically state that pictures are *not* used in VNeST because they can constrain a verb's meaning to what is imaged and this may limit potential responses. Edmonds et al. regard this as particularly problematic for verbs whose semantic networks are "loose" in comparison to nouns (Edmonds, 2016, p.126). The use of pictures is therefore theorized to potentially constrain the amount of activation (i.e. strengthening) of a treated verb's network, and to limit engagement of autobiographical and episodic memory which are also regarded as active ingredients of VNeST. The use of pictures to prompt sentence production thus warrants further investigation (as does the use of video stimuli – see discussion below).

A focus on the thematic roles of a target verb and how they map onto the syntactic structures was a key component of studies investigating mapping and PAS treatments and this raises an important question regarding how key to treatment is raising the *metalinguistic awareness* of participants. Metalinguistic awareness was raised using a variety of strategies including colour cues, wh- questions, icons and written labels of thematic roles accompanied, in all cases, by discussion. In most studies, corrective feedback was also given, but in a minority of studies it was not (e.g. Byng et al., 1994; Nickels et al., 1991) with the specific aim of encouraging participants to self-monitor their production. A study carried out by Webster and Gordon (2009) gives an insight into how important latent metalinguistic awareness may be to the success of treatment. They report two different treatments given to their participant, only one of which was successful in improving verb and sentence production. The first treatment given by Webster and Gordon was a mapping treatment. However, their participant was confused by the use of linguistic terminology during treatment: this caused her to become frustrated and to disengage with the treatment which was ultimately ineffective. The second treatment specifically did *not* aim to improve metalinguistic awareness and used a noun-verb association task, with no linguistic terminology used and no discussion about errors in sentence production. This second treatment was accepted by the participant and resulted in statistically significant improvement in trained verb and sentence production. It may, therefore, be useful to establish the level of metalinguistic awareness of participants prior to planning treatment (e.g. by discussing how much "grammar" they know), and/or to trial metalinguistic treatments to establish their acceptability to a person with aphasia.

The least used ingredients of sentence treatments included the use of *conversation-based tasks* (to facilitate generalisation of sentence production skill to real life communication), *video* and *gesture cues*. Despite the lack of explicit treatment for conversation and discourse skills, there is evidence of generalization to these contexts because when discourse was assessed, it improved for the majority of participants (70%). However, it should be noted that this finding indicates that treated sentence skills carried-over into discourse and conversational *contexts* rather than constituting an improvement in discourse or conversation per se. This is because the discourse outcome measues used (e.g. proportion of complete utterances) represent a measure of sentence skill. For a fuller discussion of this distinction see Dipper, Marshall, Boyle, Hersh, Botting, and Cruice (2021). The limited use of video and gesture cues may be more of an issue. This is because these cues may exploit features which are *unique* to action verbs (i.e. that they encode movement) and thus they may be particularly effective (or active) ingredients of treatment (e.g. Blankestijn-Wilmsen etal., 2017). Indeed, there is emerging evidence that they are effective in relation to verb-in-isolation treatments (e.g. Bonifazi, Tomaiuolo, Altoè, Ceravolo, Provinciali, & Marangolo, 2013; Boo & Rose, 2011). It is therefore recommended that future studies of sentence treatment explore the use of video and gesture cues to establish their efficacy in this type of treatment.

Finally, in discussing the potential active ingredients of treatment, the relationship between the deficit underlying sentence production difficulties and how this may interact with response to treatment needs to be considered. Schwartz et al. (1994) found that participants with relatively pure agrammatism (n=3) responded better to mapping treatments than those with additional deficits (i.e. severe apraxia of speech and/or word retrieval deficits) (n=5). Edmonds et al. (2015) performed additional analysis of the background assessment results of 11 participants in a previous VNeST study (Edmonds et al., 2014). They found no relationship between overall severity of impairment and response to treatment. They therefore categorized participants in terms of their relative impairment of PAS, mapping, noun and verb retrieval. They found that participants with relatively better sentence construction and word retrieval responded best to treatment and thus that VNeST appears to be best suited to non-fluent participants with reasonably intact syntax and lexical retrieval. However, much more research is required to elucidate the relationship between type and severity of sentence production deficit and the type of treatment/ingredient.

In summary, all conclusions from this review come with the caveat that the evidence base for sentence treatments predominantly constitutes Level 4 case series and single cases. Based on Level 4 evidence, the sentence treatments described within improved people's ability to produce sentences using trained and untrained verbs, in discourse contexts and in functional communication for 59-89% of participants.

Limitations of this study.

The review is not a systematic review. Thus, all studies were not blind reviewed by two or more reviewers but solely by the first author. However, a subset of studies (20%) was blind reviewed by two of the authors of this paper with a high level of agreement. Studies were not evaluated with a published, standard tool as this is lacking for case series which was the design used for 19 (58%) of the studies reviewed. The review also does not cover studies of TUF which has been shown to be effective (e.g. Ballard & Thompson, 1999; Thomson et al., 2010; Thompson, Shapiro, Kiran & Sobecks, 2003) and studies of TUF should be included in future reviews.

Acknowledgements.

This research was supported by a doctoral studentship to the first author jointly funded by the Worshipful Company of Saddlers and the School of Health Sciences, City, University of London.

References.

Baker, E. (2012a). Optimal intervention intensity. *International Journal of Speech-Language Pathology, 14*, 401-409. https://doi.org/10.3109/17549507.2012.700323

Baker, E. (2012b). Optimal intervention intensity in speech-language pathology: Discoveries, challenges, and unchartered territories. *International Journal of Speech-Language Pathology,* 14, 478-485. https://doi.org/10.3109/17549507.2012.717967

Ballard K.J. and Thompson C.K. (1999) Treatment and generalization of complex sentence production in agrammatism. *Journal of Speech, Language & Hearing Research, 42*, 690-707.

Bastiaanse, R., Hurkmans, J. & Links, P.(2006) The training of verb production in Broca's aphasia: A multiple‐baseline across‐behaviours study, *Aphasiology, 20*(2-4), 298-311, DOI: 10.1080/02687030500474922

Bazzini, A., Zonca, G., Craca, A., Cafforio, E., Cellamare, F., Guarnaschelli, G., Felicetti, G. & Luzzatti, C. (2012). Rehabilitation of argument structure deficits in aphasia. *Aphasiology, 26*, 1440-1460, DOI:10.1080/02687038.2012.690023

Biran, M., & Fisher, S. (2015). Structured treatment can improve predicate argument structure impairment. *Aphasiology, 29(*1), 29-56. doi:10.1080/02687038.2014.961122

Blankestijn-Wilmsen, J., Damen, I., Voorbraak-Timmerman, V., Hurkmans, J., Koning, B. D., Pross, A., & Jonkers, R. (2017). The effect of static versus dynamic depictions of actions in verb and sentence production in aphasia. *Aphasiology,* 31, 1166-1182. https://doi.org/10.1080/02687038.2016.1258537

Blomert, L., Kean, M. L., Koster, C., & Schokker, J. (1994). Amsterdam—Nijmegen everyday language test: Construction, reliability and validity. Aphasiology, 8(4), 381-407. 10.1080/02687039408248666

Bonifazi, S., Tomaiuolo, F., Altoè, G., Ceravolo, M. G., Provinciali, L. & Marangolo, P. (2013). Action observation as a useful approach for enhancing recovery of verb production: new evidence from aphasia. *European Journal of Physical and Rehabilitation Medicine*, 49, 473-481.

Boo, M., & Rose, M. L. (2011). The efficacy of repetition, semantic, and gesture treatments for verb retrieval and use in broca's aphasia. *Aphasiology, 25*, 154-175. https://doi.org/10.1080/02687031003743789

Brady, M.C., Kelly, H., Godwin, J., Enderby, P. & Campbell, P. (2016). 'Speech and language therapy for aphasia following stroke', *Cochrane Database of Systematic Reviews,* 6(Art. No.: CD000425). https://doi.org/10.1002/14651858.CD000425.pub4.

Brogan, E., Ciccone, N., & Godecke, E. (2019). Treatment fidelity in aphasia randomised controlled trials. Aphasiology, 33(7), 759-779. doi:10.1080/02687038.2019.1576442

Bryant, L., Ferguson, A., & Spencer, E. (2016). Linguistic analysis of discourse in aphasia: A review of the literature. *Clinical Linguistics & Phonetics*, *30*, 489-518.

Byng, S., Nickels, L. & Black, M. (1994). Replicating therapy for mapping deficits in agrammatism: Remapping the deficit?, *Aphasiology, 8*, 315-341, https://doi.org/10.1080/02687039408248663

Byng, S. (1988) Sentence processing deficits: Theory and therapy, Cognitive Neuropsychology, 5, 629-676, DOI:http://dx.doi.org/10.1080/02643298808253277

Carragher, M., Brooke R., Worrall, L., Thomas, S., Rose, M., Simmons-Mackie, N., Asad, K., . . . Kneebone, I. (2019, May) Fidelity protocol for the Action Success Knowledge (ASK) trial: A psychosocial intervention administered by speech and language therapists to prevent depression in people with post-stroke aphasia. *BMJ Open 9*(5):e023560 DOI:10.1136/bmjopen-2018-023560

Carragher, M., Sage, K. & Conroy, P. (2015). Outcomes of treatment targeting syntax production in people with Broca's-type aphasia: evidence from psycholinguistic assessment tasks and everyday conversation. *International Journal of Language & Communication Disorders*, *50*, 322-336.

Collins, A.M. & Loftus, E.M. (1975). A Spreading-Activation Theory of Semantic Processing *Psychological Review, 82*(6), 407-428.

Conlon, E. L., Braun, E. J., Babbitt, E. M., & Cherney, L. R. (2020). Treatment fidelity procedures for an aphasia intervention within a randomized controlled trial: Design, feasibility, and results. *American Journal of Speech-Language Pathology, 29*, 412-424. doi:10.1044/2019_AJSLP-CAC48-18-0227

Conroy, P., Sage, K. & Ralph, M.A.L. (2006). Towards theory ‐ driven therapies for aphasic verb impairments: A review of current theory and practice. *Aphasiology, 20*(12), 1159-1185, DOI: 10.1080/02687030600792009

Davis, G.A. and Tan, L.L. (1987) Stimulation of sentence production in a case with agrammatism. Journal of Communication Disorders, 20, 447-457.

Dijkers, M. (2015). What is a scoping review. *KT Update, 4. An e-newsletter from the Center on Knowledge Translation for Disability and Rehabilitation Research*. [http://ktdrr.org/products/update/v4n1]

Dipper L., Marshall J., Boyle M., Hersh D., Botting N. & Cruice M. (2021). Creating a Theoretical Framework to Underpin Discourse Assessment and Intervention in Aphasia. *Brain Sciences*, 11(2), 183. doi: 10.3390/brainsci11020183. PMID: 33540723; PMCID: PMC7913065.

Dipper, L.T., Franklin, S., de Aguiar, V., Baumgaertner, A., Brady, M., Best, W. … Visch-Brink, E.G. (2021). An Umbrella Review of Aphasia Intervention descriPtion In Research: the AsPIRE project. *Aphasiology.* doi:10.1080/02687038.2020.1852001.

Ebbels, S.H. van der Lely, H. K.. & Dockrell, J.E. (2007) Intervention for Verb Argument Structure in Children With Persistent SLI: A Randomized Control Trial. *Journal of Speech, Language, and Hearing Research*, *50*, 1330 –1349.

Edmonds, L. A. (2016). A review of verb network strengthening treatment. *Topics in Language Disorders, 36*(2), 123-135. https://doi.org/10.1097/TLD.0000000000000088.

Edmonds, L. A., Obermeyer, J., & Kernan, B. (2015). Investigation of pretreatment sentence production impairments in individuals with aphasia: Towards understanding the linguistic variables that impact generalization in verb network strengthening treatment. *Aphasiology, 29*, 1312-1344. https://doi.org/10.1080/02687038.2014.975180

Edmonds, L. A. (2014). Tutorial for verb network strengthening treatment (VNeST): Detailed description of the treatment protocol with corresponding theoretical rationale. *Perspectives on Neurophysiology & Neurogenic Speech & Language Disorders, 24*(3), 78-88. from http://search.ebscohost.com/login.aspx?direct=true&db=ccm&AN=103898651&site=ehost-live

Edmonds, L. A., Mammino, K., & Ojeda, J. (2014). Effect of verb network strengthening treatment (VNeST) in persons with aphasia: Extension and replication of previous findings. *American Journal of Speech-Language Pathology, 23*(2), S312-29. doi:10.1044/2014_AJSLP-13-0098

Edmonds, L. A., & Babb, M. (2011). Effect of verb network strengthening treatment in moderate-to-severe aphasia. *American Journal of Speech-Language Pathology, 20*(2), 131-145. doi:10.1044/1058-0360(2011/10-0036)

Edmonds, L.A., Nadeau, S.E. & Kiran, S. (2009) Effect of Verb Network Strengthening Treatment (VNeST) on lexical retrieval of content words in sentences in persons with aphasia , *Aphasiology, 23*(3), 402-424, DOI: 10.1080/02687030802291339

Fink, R.B., Schwartz, M.F. & Myers, J.L. (1998). Investigations of the Sentence-Query Approach to

Mapping Therapy. *Brain and Language. 65,* 203-207.

Frattali, C.M., Thompson, C.K., Holland, A.L., Wohl, C.B. and Ferketic, M.M. (1995) *The American Speech-Language Hearing Association Functional Assessment of Communication Skills for Adults (ASHA FACS).* ASHA

Furnas, D. W., & Edmonds, L. A. (2014). The effect of computerised Verb Network Strengthening Treatment on lexical retrieval in aphasia. Aphasiology, 28(4), 401-420. doi: 10.1080/02687038.2013.869304

Goral, M. and Kempler, D. (2009) Training verb production in communicative context: Evidence from a person with chronic non-fluent aphasia. *Aphasiology,* 23 (12), 1383–1397.

Helm-Estabrooks, N. & Ramsberger, G. (1986) Treatment of agrammaticism in long term Broca's aphasia. *British Journal of Disorders of Communication, 21*, 39-45.

Hickin, J., Cruice, M. & Dipper, L. (2020). A systematically conducted scoping review of the evidence and fidelity of treatments for verb deficits in aphasia: verb in isolation treatments. *American Journal of Speech Language Pathology, 29*, 530-559. https://doi.org/10.1044/2019_AJSLP-CAC48-18-0234

Hilari, K., Behn, N., James, K., Northcott, S., Marshall, J., Thomas, S. … Goldsmith, K. (2021). Supporting wellbeing through peer-befriending (SUPERB) for people with aphasia: A feasibility randomised controlled trial. Clinical Rehabilitation. doi:10.1177/0269215521995671

Hinckley, J. J. & Douglas, N. F. (2013). Treatment fidelity: Its importance and reported frequency in aphasia treatment studies. *American Journal of Speech-Language Pathology*, 22, 279–284.

Hoffman, T.C., Paul P Glasziou, P.P., Boutron, I., Milne, R., Perera, R., Moher, D., Altman, D.G., . . . Michie, S. (2014). Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. British Medical Journal, 348:g1687 doi:

10.1136/bmj.g1687

Hoover, E.L., Caplan, D., Waters, G. & Budson, A. (2015) Effects of impairment-based individual and socially oriented group therapies on verb production in aphasia, Aphasiology, 29:7, 781-798, DOI: 10.1080/02687038.2014.989953

Jones, E. V. (1986). Building the foundations for sentence production in a non-fluent aphasic. *International Journal of Language & Communication Disorders, 21*, 63-82. https://doi.org/10.3109/13682828609018544

Kaderavek, J.N. and Justice L.M. (2010). Fidelity: An Essential Component of Evidence-Based Practice in Speech-Language Pathology. *American Journal of Speech-Language Pathology, 19*, 369–379.

Kagan, A., Simmons-Mackie, N., Rowland, A., Huijbregts, M., Shumway, E., McEwen, S., & Dickey, L. (2010). Assessment for living with aphasia. Toronto, ON: Aphasia *Institute.*

Kearns, A. Kelly. H & Pitt, I. (2020) Rating experience of ICT-delivered aphasia rehabilitation: co-design of a feedback questionnaire, Aphasiology, 34:3, 319-342, DOI:10.1080/02687038.2019.1649913

Kempler, D., & Goral, M. (2011). A comparison of drill- and communication-based treatment for aphasia. *Aphasiology*, 25(11), 1327-1346. 10.1080/02687038.2011.599364

Kolk, H. (1995). A time-based approach to agrammatic production. *Brain and Language, 50*, 282–303.

Kurland, J., Liu, A. and Stokes, P. (2018) Effects of a tablet-based home practice program with telepractice on treatment outcomes in Chronic aphasia. Journal of Speech, Language, and Hearing Research, 61, 1140–1156. https://doi.org/10.1044/2018_JSLHR-L-17-0277

Kurland, J. (2014). iRehab in aphasia treatment. *Seminars in Speech & Language*, 35(1), 3-4.

Le Dorze, G., Jacob, A. & Coderre, L. (1991). Aphasia rehabilitation with a case of agrammatism: A partial replication. *Aphasiology, 5,* 63-85. DOI:10.1080/02687039108248520

Links, P., Hurkmans, J., & Bastiaanse, R. (2010). Training verb and sentence production in agrammatic broca's aphasia. Aphasiology, 24(11), 1303-1325. doi:10.1080/02687030903437666

Lomas, J., Pickard, L., Bester, S., Elbard, H., Finlayson, A. & Zoghaib, C. (1989). The Communicative Effectiveness Index: development and psychometric evaluation of a functional communication measure for adult aphasia. *Journal of Speech and Hearing Disorders*, *54*, 113-124.

Marshall, J., Chiat, S. & Pring, T. (1997). An impairment in processing verbs' thematic roles: A therapy study. *Aphasiology, 11*, 855 — 876. DOI: 10.1080/02687039708250461

Maul, K. K., Conner, P. S., Kempler, D., Radvanski, C., & Goral, M. (2014). Using informative verbal exchanges to promote verb retrieval in nonfluent aphasia. *American Journal of Speech-Language Pathology, 23*(3), 407-420. doi:10.1044/2014_AJSLP-13-0004

Mitchum, C.C., Greenwald, M.L., & Berndt, R.S. (1997). Production-specific thematic mapping impairment: A treatment study. Brain and Language, 60, 121–123.

Moher D., Liberati A., Tetzlaff J. & Altman DG. The PRISMA Group. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med* 6(6): e1000097. https://doi.org/10.1371/journal.pmed1000097

Murad, M.H., Sultan, S., Haffar S. and Bazerbachi, F. (2018). Methodological quality and synthesis of case series and case reports *BMJ, Evidence-Based Medicine*, 23:60–63. doi:10.1136/bmjebm-2017-110853

Newton, C., Kirby, P., & Bruce, C. (2017). Getting into shape: The effect of shape coding on the spoken language production of two men with chronic aphasia. *Aphasiology,* 1-23. doi:10.1080/02687038.2017.1306639

Nicholas, L. E., & Brookshire, R. H. (1993). A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. Journal of Speech Language and Hearing Research, 36, 338–350. doi:10.1044/jshr.3602.338

Nickels, L., Byng, S. & Black, M. (1991). Sentence processing deficits: A replication of therapy. *International Journal of Language & Communication Disorders*, 26, (2), 175-199.

Oxford Centre for Evidence-Based Medicine 2011 Levels of Evidence. (2011). Available at: http://www.cebm.net/wp-content/uploads/2014/06/CEBM-Levels-of-Evidence-2.1.pdf [Accessed: 23 November 2017]

Palmer, R., Dimairo, M., Cooper, C., Enderby, P., Brady, M., Bowen, A. et al. (2019) Self-managed, computerised speech and language therapy for patients with chronic aphasia post-stroke compared with usual care or attention control (Big CACTUS): a multicentre, single-blinded, randomised controlled trial. *Lancet Neurology, 18*, 821–33.

Park, Y., Goral, M., Verkuilen, J. & Kempler, D. (2013). Effects of noun–verb conceptual/phonological relatedness on verb production changes in Broca's aphasia. *Aphasiology, 27*, 7, 811-827.

Peters, M.D.J., Godfrey, C.M., Khalil, H., McInerney, P., Parker, D. & Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *International Journal of Evidence-Based Healthcare,* 13 3: 141-146. https://doi.org/10.1097/XEB.0000000000000050

Pulvermuller, F., Neininger, B., Elbert, T., Mohr, B., Rockstroh, B., & Koebbel, P. et al. (2001). Constraint-induced therapy of chronic aphasia after stroke. Stroke, 32, 1621–1626.

Robey, R. R., & Schultz, M. C. (1998). A model for conducting clinical outcome research: An adaptation of the standard protocol for use in aphasiology. *Aphasiology, 12*, 787–810.

Rochon, E., Laird, L., Bose, A. & Scofield, J. (2005). Mapping therapy for sentence production impairments in nonfluent aphasia. *Neuropsychological Rehabilitation, 15*, 1-36. https://doi.org/10.1080/09602010343000327

Rochon, E. & Reichman, S. (2003). A Modular Treatment for Sentence Processing Impairments in Aphasia: Sentence Production. *Journal of Speech-Language Pathology and Audiology,* 27, 202 -210.

https://doi.org/10.1080/02687030701800800.

Routhier, S., Bier, N. & Macoir, J. (2016) Smart tablet for smart self-administered treatment of verb anomia: two single-case studies in aphasia, Aphasiology, 30, (2-3), 269-289, DOI: 10.1080/02687038.2014.973361

Schneider, S. & Thompson, C. (2003). Verb production in agrammatic aphasia: The influence of semantic class and argument structure properties on generalization. *Aphasiology, 17*, 213-241. https://doi.org/10.1080/729255456

Schwartz, M.F. Saffran, E.M., Fink, R.B. Myers, J.L. & Martin, N. (1994) Mapping therapy: A treatment programme for agrammatism. *Aphasiology, 8*, 19-54, DOI: 10.1080/02687039408248639

Silagi, M.L., Hirata, F.N., Iracema, L. & de Mendonça, Z. (2014) Sentence production and rehabilitation of agrammatism Sentence production in rehabilitation of agrammatism: A case study. *Dementia Neuropsychologia, 8*, 297-301.

Spell, L. A., Richardson, J. D., Basilakos, A., Stark, B. C., Teklehaimanot, A., Hillis, A. E., & Fridriksson, J. (2020). Developing, implementing, and improving assessment and treatment fidelity in clinical aphasia research. *American Journal of Speech-Language Pathology, 29*(1), 286-298. doi:10.1044/2019_AJSLP-19-00126

Springer, L., Huber, W., Schlenck, K.J. & Schlenck, C. (2000). Agrammatism: Deficit or compensation? Consequences for aphasia therapy. *Neuropsychological Rehabilitation*, 10 (3), 279–309.

Takizawa, T., Nishida, N., Ikemoto, A., & Kurauchi, N. (2015). Comparison of single-word therapy versus sentence therapy for verb retrieval and sentence production in Broca’s aphasia.

Aphasiology, 29, 169-194. https://doi.org/10.1080/02687038.2014.961121

Taub, E., Uswatte, G., King, D. K., Morris, D., Crago, J. E., & Chatterjee, A. (2006). A placebo-controlled trial of constraint induced movement therapy for upper extremity after stroke. Stroke, 37, 1045–1049. doi:10.1161/01.STR.0000206463.66461.97

Thompson, C.K., Choy, J.J., Holland, A. & Cole, R. (2010) SENTACTICS: computer-automated treatment of underlying forms. *Aphasiology, 24 (10),* 242-1266.

Thompson, C.K., Shapiro, S.P., 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*, 591–607. doi: 10.1044/1092-4388(2003/047)

Thompson, C.K., Lange, K. L., Schneider, S. L. & Shapiro, L. P. (1997) Agrammatic and non-brain-damaged subjects' verb and verb argument structure production, Aphasiology, 11:4-5, 473-490, DOI: 10.1080/02687039708248485

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., . . . Lewin, S. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, 169(7), 467-473. doi:10.7326/M18-0850

Webster, J. and Whitworth, A. (2012). Treating verbs in aphasia: exploring the impact of therapy at the single word and sentence levels. *International Journal of Language & Communication Disorders, 47*, 619–636. doi:10.1111/j.1460-6984.2012.001

Webster, J. & Gordon, B. (2009). Contrasting therapy effects for verb and sentence processing difficulties: a discussion of what worked and why. *Aphasiology,* 23 (10), 1231-1251.

Webster, J., Morris, J. & Franklin, S. (2005). Effects of therapy targeted at verb retrieval and the realisation of the predicate argument structure: a case study. Aphasiology, 19 (8), 748-764.

Whitworth, A., Webster, J., & Howard, D. (2015). Argument structure deficit in aphasia: It’s not all about verbs. *Aphasiology, 29*(12), 1426-1447. doi:10.1080/02687038.2015.1037823

Appendix.
Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) Checklist

SECTION	ITEM	PRISMA-ScR CHECKLIST ITEM	REPORTED ON PAGE #
TITLE			
Title	1	Identify the report as a scoping review.	1
ABSTRACT			
Structured summary	2	Provide a structured summary that includes (as applicable): background, objectives, eligibility criteria, sources of evidence, charting methods, results, and conclusions that relate to the review questions and objectives.	2 - 3
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known. Explain why the review questions/objectives lend themselves to a scoping review approach.	4 - 11
Objectives	4	Provide an explicit statement of the questions and objectives being addressed with reference to their key elements (e.g., population or participants, concepts, and context) or other relevant key elements used to conceptualize the review questions and/or objectives.	10 - 11
METHODS			
Protocol and registration	5	Indicate whether a review protocol exists; state if and where it can be accessed (e.g., a Web address); and if available, provide registration information, including the registration number.	11 - 14
Eligibility criteria	6	Specify characteristics of the sources of evidence used as eligibility criteria (e.g., years considered, language, and publication status), and provide a rationale.	12
Information sources*	7	Describe all information sources in the search (e.g., databases with dates of coverage and contact with authors to identify additional sources), as well as the date the most recent search was executed.	11 - 12
Search	8	Present the full electronic search strategy for at least 1 database, including any limits used, such that it could be repeated.	11 - 12
Selection of sources of evidence†	9	State the process for selecting sources of evidence (i.e., screening and eligibility) included in the scoping review.	12
Data charting process‡	10	Describe the methods of charting data from the included sources of evidence (e.g., calibrated forms or forms that have been tested by the team before their use, and whether data charting was done independently or in duplicate) and any processes for obtaining and confirming data from investigators.	13 - 14
Data items	11	List and define all variables for which data were sought and any assumptions and simplifications made.	Tables 1 & 2
Critical appraisal of individual sources of evidence§	12	If done, provide a rationale for conducting a critical appraisal of included sources of evidence; describe the methods used and how this information was used in any data synthesis (if appropriate).	13 - 14
Synthesis of results	13	Describe the methods of handling and summarizing the data that were charted.	17
RESULTS			
Selection of sources of evidence	14	Give numbers of sources of evidence screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally using a flow diagram.	15

SECTION	ITEM	PRISMA-ScR CHECKLIST ITEM	REPORTED ON PAGE #
Characteristics of sources of evidence	15	For each source of evidence, present characteristics for which data were charted and provide the citations.	Tables 1 & 2
Critical appraisal within sources of evidence	16	If done, present data on critical appraisal of included sources of evidence (see item 12).	13 - 14
Results of individual sources of evidence	17	For each included source of evidence, present the relevant data that were charted that relate to the review questions and objectives.	Tables 1 & 2
Synthesis of results	18	Summarize and/or present the charting results as they relate to the review questions and objectives.	15 - 25
DISCUSSION			
Summary of evidence	19	Summarize the main results (including an overview of concepts, themes, and types of evidence available), link to the review questions and objectives, and consider the relevance to key groups.	25 - 35
Limitations	20	Discuss the limitations of the scoping review process.	35 - 36
Conclusions	21	Provide a general interpretation of the results with respect to the review questions and objectives, as well as potential implications and/or next steps.	35
FUNDING			
Funding	22	Describe sources of funding for the included sources of evidence, as well as sources of funding for the scoping review. Describe the role of the funders of the scoping review.	36

JBI = Joanna Briggs Institute; PRISMA-ScR = Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews.

* Where *sources of evidence* (see second footnote) are compiled from, such as bibliographic databases, social media platforms, and Web sites.

† A more inclusive/heterogeneous term used to account for the different types of evidence or data sources (e.g., quantitative and/or qualitative research, expert opinion, and policy documents) that may be eligible in a scoping review as opposed to only studies. This is not to be confused with *information sources* (see first footnote).

‡ The frameworks by Arksey and O'Malley (6) and Levac and colleagues (7) and the JBI guidance (4, 5) refer to the process of data extraction in a scoping review as data charting.

§ The process of systematically examining research evidence to assess its validity, results, and relevance before using it to inform a decision. This term is used for items 12 and 19 instead of "risk of bias" (which is more applicable to systematic reviews of interventions) to include and acknowledge the various sources of evidence that may be used in a scoping review (e.g., quantitative and/or qualitative research, expert opinion, and policy document).

From: Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMAScR): Checklist and Explanation. Ann Intern Med. 2018;16

Chapter 3. The treatment of personally relevant verbs and different sentence structures in aphasia.

3.1 Introduction.

This PhD study investigated the feasibility, fidelity, acceptability and compliance with a low dose, clinician delivered sentence production treatment (SPT), supplemented by self-managed home practice via computer-based treatment exercises. The study also included preliminary efficacy testing.

The two preceding chapters report systematically conducted scoping reviews of the research relating to verb-in-isolation and sentence treatments respectively, which informed the content of the SPT and ensured it was based on the best available evidence. These two reviews encompassed 37 studies (182 participants) and 33 studies (126 participants) of Level 4 evidence, employing phonological and semantic cueing (verb-in-isolation) and mapping, predicate argument structure, and VNeST (sentence) treatments. Trained items improved for 80-83% of participants; generalization was minimal for untrained verbs (15% participants) but better for untrained sentences (59% participants), and sentence therapies yielded gains in discourse (70% participants). Whilst treatments were well described, there was no consensus on the dosage that is effective, and fidelity was rarely assessed. This makes it difficult to determine whether treatments were indeed delivered as intended. As such, future research should focus on training verbs but additionally in sentence contexts, using a minimum dosage of at least 20 hours and proactively assess generalization and fidelity. Given the limited generalization findings, these reviews also highlight the central importance of the items chosen for treatment.

The SPT developed for this study targeted personally relevant (PR) verbs and the literature relating to the selection and treatment of PR words (including verbs) is reviewed in the first half of this chapter. The narrative review informed the rationale for the targeting of verbs which were personally relevant, and the process used to select PR verbs. Issues discussed in relation to the targeting of PR words in aphasia treatment include why treating PR words may increase the impact of treatment, and the difficulties of determining what constitutes PR words for any one individual. Finally, the SPT targeted a variety of sentence types (i.e., declarative sentences, imperatives, Yes/No questions and wh-questions) and this has rarely been investigated. The chapter therefore includes a brief discussion of the (limited) literature relating to the treatment of sentence structures other than declarative sentences.

Because the treatment investigated in this study is supplemented by self-managed home practice via computer-based treatment exercises, the second half of the chapter reviews and synthesizes the literature relating to: i) the acceptability to PwA and their carers of computer-based treatments, and the limited research on the views of SLTs in this area (with the aim of incorporating features into the SPT to maximize its acceptability); ii) design accessibility of computer-based aphasia treatments (with the aim of incorporating as many aphasia-friendly features as possible into the PowerPoint slides developed to deliver SPT exercises) and iii) the evidence base for self-managed computer-based aphasia treatments (with the aim of evaluating this approach to delivery and assessing its potential for use in the current study). The chapter finishes with a brief summary of the current uptake of self-managed, computer-based aphasia treatments.

3.2 The treatment of personally relevant verbs.

3.2.i Background.

The few existing reviews of verb treatment studies indicate that treatment effects are limited to *trained* items for c80% of PwA (e.g., de Aguiar et al., 2016; Hickin et al., 2020). The lack of generalization of verb treatments to untrained verbs strongly suggests that treatments should target verbs which are *personally relevant* to participants (e.g., Webster and Whitworth, 2012). However, the systematic reviews reported in Chapters 1 and 2 (Hickin et al., 2020; Hickin et al., 2022) found that, to date, both verb and sentence treatments have very rarely done this: only three studies of verb treatments targeted PR verbs (Carragher et al., 2013; Palmer et al., 2012; Park et al., 2013[9]), as did only three studies of sentence treatments (Goral and Kempler, 2011; Hoover et al., 2015; Webster et al., 2005).

One possible reason for the neglect of PR verbs in the treatment literature is that the process of determining PR items is not straightforward. This is discussed by Thiessen and Brown (2021) in a rare paper on the subject. They discuss the rationale for personalization of treatment in general (e.g., that it may increase compliance), and note that, given the lack of generalization of anomia treatments to untrained items, selection of PR words is imperative if the impact of treatment is to be maximized. However, they also note that it is unclear how PR items may best be selected, and that it may present logistical issues because it is likely to be time consuming.

The lack of clarity about the best way to select PR words is evident from the variety of ways in which PR words were selected in the few studies which did target them. For example, Hoover et al. (2015) selected PR verbs from nine pre-determined functional conversation topics. However, how these topics were deemed functional is not described. Goral and Kempler (2011) treated PR verbs related to their participant's occupation, "home" verbs and "recreation" verbs. These verbs "were selected on the basis of relevance to the participant's life and communication needs" (p.1389) which is a welcome rationale, but again it is not specified how relevance was determined. Webster et al's (2005) participant selected his own verbs and, whilst it is encouraging that NS

[9] Palmer et al (2019) published a study which targeted PR words after the completion of the systematic review reported in Chapter 2, and this is discussed later in this chapter.

appeared to be proactive in choosing his PR verbs, once again it is not reported how this process was facilitated. Palmer et al. (2017) asked their 100 participants to select 100 PR words for treatment and gave PwA and their families time to think about this, which is to be applauded. They also used 18 picture card prompts with possible topic areas to prompt them to think about what was important to them. These cards were informed by topics chosen in a pilot CACTUS study (Palmer et al., 2012) and the content of the StepbyStep programme investigated in a series of studies (Mortley et al., 2004: Palmer et al., 2012, 2019). However, the actual topics covered by the 18 cards are not reported. In summary, identifying the best way in which to select PR words for treatment would be facilitated by reporting this process in replicable detail in future studies, and the process by which the PR verbs were selected for treatment in this study is described below, alongside the literature which motivated this. The process drew upon an eclectic range of research and consequently used a variety of techniques.

Finally, before discussing the process used to select PR verbs, it should be noted that the construct of *personal* relevance has been debated in the literature under different terms including '*functional*' relevance. Whilst these two terms are related, they are not synonymous. Turning to functional relevance (FR) first, the Oxford English Dictionary defines functional as:

> "designed to be practical and useful rather than attractive"
> (https://en.oxforddictionaries.com/definition/functional).

Thus, FR words chosen for treatment should presumably be both practical and useful, however identifying such words has proved to be far from easy. Personal, on the other hand, is defined as:

> "relating or belonging to a single or particular person rather than to a group or an organisation"
> (https://dictionary.cambridge.org/dictionary/english/personal).

Thus, PR words chosen for treatment should have *individual* relevance to a PwA which does not necessarily mean that a PR word will be either practical or useful (e.g., words chosen relating to hobbies and entertainment might be chosen for reasons of pleasure and enjoyment rather than practicality). In other words, a personally relevant word may not be functionally relevant.

3.2.ii The process of identifying personally relevant verbs.

The SPT investigated in this study targeted 20 PR verbs (out of a set of 40 PR verbs). To facilitate the selection of what was hoped would be genuinely PR verbs, firstly the process of selection was delayed until the third week of initial baseline testing to allow the researcher to get to know participants and begin to establish their interests and what were difficult communication situations for them. The process of "getting to know" someone with aphasia to identify PR words warrants more discussion. In a rare paper on the topic, Renvall, Nickels and Davidson (2013a) note that it may not be easy for a PwA to identify PR words, particularly if they have more severe aphasia. One way to assist is to ask a proxy (such as a significant other) to help. However, the involvement of a proxy may compromise the personal relevance of chosen words since there is evidence that the views of significant others as to what may facilitate the communication of a PwA may be different to the views of the PwA themselves. For example, Haley, Womack, Helm-Estabrooks, Lovette and Goff (2013) found only 71% agreement between PwA and their proxies in terms of preferred topics of conversation which would suggest that PwA and their proxies are also likely to differ in their selection of PR words. However, participants in this study were asked if they would like to involve a significant other in the selection process, but, in the event, only one participant (P2) did so. (She and her partner spent a week discussing her PR verbs and presented the researcher with a list). Even if a PwA can be proactive in selecting their PR words, they may select words which have high personal value as opposed to those which have a large impact on communication success (Mason et al., 2011). As Thiessen and Brown note (2021), establishing how to select PR items which maximize the functional impact of treatment should be a high priority for future research.

The lack of detailed guidance on selection of PR words means that clinicians and researchers may use a variety of strategies. These include interviewing a PwA to establish what they like to talk about, and this was done informally throughout the baseline testing phase of the study. Although *not* used in this study because of time constraints, additional strategies could include: establishing a PwA's social network to highlight potential topics of interest (e.g., gardening if someone's social network included fellow allotment plot holders); observing the PwA communicating in real life which has the advantage of ecological validity (Davidson, Worrall & Hickson, 20003); analyzing conversation samples of a PwA to identify sources of breakdown such as being unable to retrieve a PR word (e.g., Perkins, Crisp and Walshaw (1999) who found that lexical retrieval difficulties were the commonest cause of conversation breakdown) and target these PR words in treatment (e.g., Lesser & Algar, 1995); and, asking a PwA and their carer/s to keep a communication diary - potentially time consuming but useful to give insight across a wider time frame and a variety of communication situations than is possible using conversation analysis for example (e.g., Davidson et al., 2003). In summary, it is unclear which of the strategies described above are most effective in selecting truly PR words, and, given that they are potentially time consuming to undertake, establishing their relative effectiveness seems a high priority for future research.

Another strategy used to identify PR verbs in this study was to set goals for participants and use these to inform the selection of PR verbs. The procedure used was a Goal Action Planning (GAP) procedure (Scobbie, McLean, Dixon, Duncan, & Wyke, 2013; Scobbie, Dixon, & Wyke, 2011) with the SMARTER goal setting framework (Hersh, Worrall, Howe, Sherratt & Davidson, 2012)) also influential. Thiessen and Brown (2021) regard client-centred goal setting as crucial to the process of personalizing treatment.

The results of the CETI (Lomas et al., 1989) already carried out with participants (as one of the outcome measures for preliminary efficacy testing) were also used to identify potential PR verbs. Thus, verbs relating to questions rated as difficult (e.g., talking about my emotions) were discussed as potential treatment targets. The CETI was viewed as a useful means of determining PR verbs because it was developed in collaboration with PwA and their carers: they identified communication situations that were important in their day-to-day lives. Through the process of validating the CETI, the number of situations was reduced to those generalizable across people, time, and places. Using the CETI communication scenarios rated as difficult by participants therefore seemed likely to give a valid insight into verbs which might be useful to target in treatment, together with discussion with participants to ensure that verbs were personally (as well as functionally) relevant.

During baseline testing, participants in this study were also asked to identify *people, topics,* and *situations* when the participant found it hard to communicate using prompt cards taken from the interview section (Part B) of the Conversation Analysis Profile for People with Aphasia (CAPPA: Whitworth, Perkins & Lesser, 1997) (see Appendix A). The CAPPA was used because clinical experience indicated that the prompt cards were a useful resource in identifying PR words for treatment. The CAPPA prompts were cross-referenced with the findings of a study carried out by Davidson et al. (2003). Davidson and colleagues observed 15 neurologically healthy older people (NHP) and 15 PwA, communicating in the community. They compared the communication activities carried out by each group and the topics that they discussed. They found that NHP and PwA had many topics and communication activities in common, but also activities and topics unique to each group. All of these were also used to inform the prompt lists. For example, *holidays, pets* and *hobbies* were added to the Topics prompt list as these were common topics for both PwA and NHP, whilst talking to health professionals was added to the People list for the same reason. Discussing *current affairs* was added to the Topics list because it was commonly discussed by NHP but not by PwA and thus appeared to be a fruitful area to target in treatment. The prompt list was also compared to the topics most commonly chosen by 100 PwA who each selected 100 PR words in the Big CACTUS study (Palmer et al., 2017) to identify any topics not covered: no topics were added as a result. Finally, participants in this study were encouraged to consider not only what they *had* to communicate, but also about what they *wanted* or *liked* to communicate in line with questions asked in the Communication Disability Profile (Swinburn & Byng, 2006). The aim here was to identify verbs which were likely to perform an *interactional* role (e.g., discussing hobbies), as well as those likely to perform a *transactional* role in communication (e.g., discussing health problems) and is in line with treatment priorities identified by PwA (e.g., Wallace et al., 2017).

Whilst written prompt lists were used in the process of selecting PR verbs, picture prompts were not. Picture prompts are often used to help PwA identify personally (and/or functionally) relevant words as, for example in the Big CACTUS study (e.g., Palmer et al., 2017) including using commercially available picture libraries e.g. (e.g. Speechmark Colourcard Categories (https://www.routledge.com/Categories-ColorCards-2nd-Edition/Speechmark/p/book/9780815369011https://www.routledge.com/Categories-ColorCards-2nd-Edition/Speechmark/p/book/9780815369011), and this may reduce the time taken to identify potentially relevant words for treatment. However, there is a concern that the use of picture prompts may restrict the selection of words by PwA in terms of both the *grammatical class* of words selected and their *imageability* and *concreteness*. To expand, Renvall et al. (2013a) compared the concreteness and imageability ratings of the words in 10 semantic categories commonly regarded as functionally relevant and therefore typically targeted in treatment (body parts, fruits, vegetables, spices, non-alcoholic drinks, clothing, footwear, sports, furniture and transportation vehicles) to those of the 100 most frequent words in the CELEX and SUBTLEXus databases (both based on spoken word frequency). They found that the 100 most frequent words in the databases were less imageable and less concrete/more abstract, than topics typically chosen in aphasia treatment studies. There is therefore a bias towards treating *highly imageable* and *concrete* words.

A further concern identified by Renvall et al. related to the *grammatical class* of the words depicted in commercially available picture libraries. When Renvall et al. analyzed the CELEX and SUBTLEXus databases to establish the 100 most frequent words in each, they found that the most frequent word class (in both databases) was *verbs* followed by *pronouns, adverbs* and *prepositions*, with *nouns* perhaps surprisingly in ninth place. When the 1000 most frequent words were considered, nouns were the most frequent word class closely followed by verbs, with adjectives in third place and adverbs again appearing in the top four (in fourth place). The frequency of words other than nouns is, as Renvall et al. state, a serious cause for concern given that treatments very predominantly target nouns, with an almost complete neglect of adverbs, adjectives, and prepositions. For example, of the 10,000 words chosen by the 100 participants in the Big CACTUS study only 98 were verbs, whilst only 77 were categorized as descriptive (i.e., adjectives/adverbs with most of these colours). If we are to facilitate PwA to talk not just about the more material aspects of life but also to express their emotions and to "have complex conversations including giving explanations" (Wallace et al., 2017, p. 1370) it is imperative that we treat more abstract words, including more abstract nouns and verbs. For example, targeting verbs such as "know," "think," "mean" and "say" - all of which are in the top 100 words of either SUBTLEX-US or CELEX - might significantly help PwA to express their ideas and opinions. Thus, it was a deliberate decision not to use picture prompts in the process of selecting PR verbs in the hope that this would encourage the selection of more abstract verbs, and, as already noted, this was also supported by encouraging participants to consider not just what they *needed* to communicate but what they *wanted* to communicate, such as communicating emotions and discussing things in depth (scenarios also included in the CETI).

Finally, the verbs selected by participants were compared to the verbs included in the list of 357 words which were the most frequent words in samples of everyday conversations of unimpaired speakers, published in a companion paper to Renvall et al. (2013a) - see Renvall et al., (2013b) for a full description of how the list was developed. Additionally, selected verbs were compared to those amongst the 100 and 1000 most frequent words from both SUBTLEX and CELEX also listed by Renvall et al. (2013a & b). These lists represent an important resource to support the selection of PR words including verbs. The MRC database was used to ascertain the ratings of PR verbs for concreteness, familiarity, frequency, imageability and syllables for the purposes of matching treated and untreated sets. The use of psycholinguistic databases for this and related purposes (e.g., exploring the role of psycholinguistic properties of words in terms of responsiveness to treatment) is important in aphasia research.

3.2.iii The response of personally relevant words to treatment.

It has been speculated that PR words might be more responsive to treatment than words which are not (e.g., Renvall, 2013a & b). A number of theories have been proposed to support this view, and these include that PR words may be i) more motivating and thus will be practiced more resulting in greater treatment effects, ii) that they are more salient, and this will result in more (effective) neuroplastic change, and iii) that PR words are more likely to be used in real life communication resulting in greater generalization of treatment. The evidence in support of these theories will now be summarized briefly.

Getting a sufficient dose of aphasia treatment is increasingly recognized as a likely vital (active) ingredient of treatment (e.g., Baker 2012a; Dignam, Rodriguez, & Copland, 2016). The potential for PR words to be more motivating and thus to increase the amount of practice undertaken is therefore an attribute worthy of serious clinical consideration. There is, however, only indirect evidence that PR words are more motivating, and this comes from a small number of studies investigating the utility of PR material with people with more severe aphasia. These studies are based on the premise that, in the context of severely impaired language, the contextual information provided by PR stimuli will be crucial in supporting language-based tasks (see Van Lancker & Nicklay (1992) for a more detailed discussion). Wallace and Canter (1985) found superior performance with PR stimuli on naming, repetition, reading and auditory comprehension tasks for participants with severe aphasia, and they speculate that "*greater motivation* to respond to information that is more familiar and relevant" could have accounted for the difference in performance (Wallace & Canter, 1985, p. 388). PR words also may be *intrinsically* motivating because they may help a PwA with aphasia to successfully achieve activities of daily living (e.g., by targeting verbs such as "order" and "pay" when someone enjoys eating out). This may mean that PR words are also more likely to be used in everyday communication than words which are not personally relevant and so they will receive more reinforcement outside of treatment sessions also facilitating generalization.

That working on PR words during treatment may be more motivating can also be related to theories of adult learning. Holland and Hopper (2005) discuss adult learning theory in relation to aphasia treatment. They note that it is likely that adults learn best when they experience *success*, *volition* (i.e., they feel they have a choice in how they learn), they *value* what they are learning, and when they find learning *enjoyable*. Whilst the success of learning cannot be guaranteed with PR words (since, as Renvall et al. point out, it is equally possible that a PwA may choose words that they perceive to be difficult to learn but useful, or easy to learn and useful), it seems likely that personally chosen words could enhance volition, value and enjoyment during the learning process, and that this in turn may mean that PwA practice them more and find them easier to learn.

Another factor which may account for the potentially greater responsiveness of PR verbs to treatment is that they may be highly *salient. Salience* is a characteristic of stimuli thought to be important in inducing neuroplastic changes (e.g., Kleim & Jones, 2008; Raymer et al., 2008), and, whilst there is no clear definition of salience (see Renvall et al., 2013a), it relates to the perceived meaningfulness of a stimulus. It seems reasonable to suggest, therefore, that PR words would represent (more) salient stimuli. However, whilst there is evidence which indicates neuroplastic changes are more likely to occur when salient stimuli are used in the rehabilitation of *movement* in animals (see Kleim & Jones, 2008), there is no concrete evidence that this is the case in relation to the rehabilitation of *language* in humans.

Van Lancker and Nicklay (1992) propose an alternative explanation for the potential neuroplastic advantage of PR words over generic (common) words. They speculate that the mechanism underlying the superior performance of PR words relates not to salience, but to the specialist role of the (undamaged) right cerebral hemisphere in processing PR information. That is, the specialist role of the right hemisphere in processing material which has *emotiv*e meaning to an individual may underlie the superior performance of PR words during treatment. Finally, Renvall et al., (2013a) speculate that PR words respond better to treatment because they are more richly represented in the lexicon i.e., that a richer semantic network surrounding personally chosen words may facilitate the strengthening/elaboration of the network which is theorized to underpin treatment (e.g., in SFA and in VNeST).

The factors outlined above suggest that PR words should respond better to treatment, and, in particular, result in greater generalization to real-life communication. However, the results from the few verb and sentence treatment studies which targeted PR *verbs* only partially support this. Goral & Kempler (2011), Hoover et al. (2015) and Palmer et al. (2012) reported a positive impact of treatment on functional communication, whilst Carragher et al. (2013) and Palmer et al. (2019) reported no impact of treatment on real life conversation. The lack of impact on functional communication in Palmer et al. (2019) is likely to be related to extremely limited therapeutic input targeting contextual practice (average of 45 mins across 6 months) and the outcome measure used being insensitive to change (TOMS activity scale in Palmer et al., 2019). In Carragher et al. a dose of only eight hours might not have been sufficient to provoke changes in conversation.

Thus, there is currently little empirical evidence to support the premise that improved retrieval of PR words will preferentially improve everyday communication over improved retrieval of more generic words. Despite the lack of evidence, the belief that PR words are important in achieving generalisation of treatment gains to everyday communication is evident in VNeST (e.g., Edmonds et al., 2015) which currently has the strongest evidence base of all sentence treatments (see Chapter 2). The authors of VNeST encourage the production of PR Agent – Verb –Theme exemplars during treatment (e.g., Edmonds et al., 2009; Edmonds & Babb, 2011) and specifically state that this is because they believe that this maximizes the likelihood of these trained structures being used in real life communication.

Script Training treatment also has a developing evidence base (see e.g., Cherney, 2012) and personal relevance may play a key role in the success of this treatment too. Script training involves rehearsing a script which has personal relevance to the participant (e.g., ordering food in a restaurant) but the degree of personal relevance of each script varies with some studies training scripts deemed to be more generically relevant (such as ordering food) whilst others trained much more personal scripts (such as a wedding speech) (see Cherney, Kaye, Lee & van Vuuren (2015) for a fuller discussion). To investigate the potency of PR words in script training, Cherney et al. (ibid) carried out a study which compared the production of four PR items (words or short phrases) with two-four generic items in trained and untrained scripts. Whilst both PR and generic items improved significantly in trained scripts, only PR items showed significant generalization to untrained scripts. Whilst this finding is encouraging, its significance is tempered by the small number of words (and participants) treated. Clearly the importance of treating PR words in terms of generalization to untreated discourse production requires much more extensive investigation.

3.2.iv Summary.

This discussion has highlighted the importance of treating PR words given that word retrieval treatments predominantly only improve access to *trained* words. The difficulties associated with identifying PR words for treatment have been discussed, and some potential resources to assist in identifying PR words (such as lists of frequently occurring words) pinpointed. Whilst these lists are welcome, caution is necessary as words which are frequently occurring and/or seen as functionally relevant, may not be personally relevant to any one individual. It therefore remains imperative that words included in lists are *discussed* with individual PwA and their families to identify pertinent items. It is also important that picture prompts are not used as the sole/main means of selecting PR words for treatment as these are highly likely to bias selection of words to nouns which are highly imageable. The discussion has also explored the (limited) evidence which suggests that PR stimuli may respond better to treatment than non-PR stimuli and the possible mechanisms responsible for this. It is, however, difficult to untangle the ingredients of treatment which may be unique to (and thus active for) PR stimuli and which might underlie their possible superior response to treatment. To illustrate this, the *motivating* quality of PR stimuli may be important in their response to treatment, but this also likely gives rise to high *salience* and this too may be a key factor. High salience may in turn relate to the rich cerebral representation PR items seem likely to have in terms of both semantic representation and more generally via experiences, associations, and memories. Finally, PR words are *ipso facto* more likely to be used in everyday communication than non-PR words, potentially facilitating generalization of treatment to everyday communication – the holy grail of aphasia treatment.

The SPT targeted the production of PR verbs in a variety of sentence types (i.e., declarative sentences, imperatives, Yes/No questions and wh-questions) and this has rarely been investigated. The first half of the chapter therefore concludes with a brief discussion of the (limited) literature relating to the treatment of sentence structures other than declarative sentences.

3.3 Treatment of different types of sentence structure.

The SPT investigated in this PhD treated declarative statements as well as imperatives, Yes-No questions and Wh- questions. This is unusual for sentence production treatments in aphasia and so the rationale for the inclusion of other types of linguistic structure in the SPT is outlined here. The main motivation to include different types of sentences came from a general meta-linguistic awareness that people do not communicate using only declarative sentences. In addition, extensive clinical experience indicated the need to equip PwA with the ability to produce other types of structures – to ask questions and make requests – not just to make statements. Further justification for this clinical intuition comes from the Longman Grammar of Spoken and Written English (LGSWE: Biber et al., 1999). Biber et al. state that:

> "It is not at all surprising that questions and imperatives, the sentence types that typically elicit a response, are more frequent in conversation than in written language. Questions are three to four times more common . . . and imperatives are five times as common" (p.1045).

Thus inclusion of these sentence structures within the SPT appears important, especially given the aim of facilitating the use of treated verbs and sentences within discourse (RQ 4).

Despite the obvious existence/use of structures other than declarative sentences, of the 33 sentence treatment studies accepted into the literature review, only two studies treated a range of structures with the aim of helping PwA improve their ability to produce each of these different types of structure. Helm-Estabrooks and Ramsberger (1986) report the results of a group study with six PwA using the Helm Elicited Language Programme for Syntax Stimulation (HELPSS) (Helm-Estabrooks, 1981). Silagi, Hirata, Iracema, and de Mendonça (2014) also report a single case study also using the HELPSS. The HELPSS programme treats 11 different types of sentences in a hierarchy of difficulty based on a study by Gleason, Goodglass, Green, Ackerman and Hyde (1975). Gleason et al. analysed the output of eight people with Broca's aphasia in a story completion task and established a hierarchy of difficulty in terms of the types of sentences produced by their participants. This hierarchy is shown in Table 1 (Helm-Estabrooks & Ramsberger, 1986, p.40).

Table 1 The hierarchy of sentence structures treated in the HELPPS (Helm-Estabrooks & Ramsberger 1986, p.40).

1. Imperative Intransitive	Example: 'Wake up.'
2. Imperative Transitive	Example: 'Lock the door.'
3. Wh-Interrogative	Example: 'What are you eating?'
4. Declarative Transitive	Example: 'He paints houses.'
5. Declarative Intransitive	Example: 'He sings.'
6. Comparative	Example: 'He is taller.'
7. Passive	Example: 'The car was towed.'
8. Yes-No Questions	Example: 'Did you buy the paper?'
9. Direct-Indirect Object	Example: 'He gives his son a toy'.
10. Embedded Sentences	Example: 'He wanted him to be rich.'
11. Future	Example: 'He will travel.'

Each type of structure was treated using a short story cue followed by a probe question, at two levels of difficulty. Helm-Estabrooks and Ramsberger state that multiple exemplars of each type of structure were treated. The authors do not report on the improvement of each of the structures individually but do report a significant improvement for the group on both a standardised test of expression and on the Cookie Theft Picture Description task (but they do not use a measure of conversation so the impact of treatment on this is unclear). Silagi et al. (2014) treated sentence structures 1 – 8 in the hierarchy and report the number of treatment sessions required to reach 85% correct responses, with most sentences taking four or fewer sessions to reach this criterion.

3.4. Conclusions.

The first half of this chapter has reviewed the literature pertaining to the selection of PR words and their treatment, and to the treatment of sentence structures other than declarative sentences. It should be acknowledged that the literature is slim, particularly regarding the latter. The two papers relating directly to the *selection* of PR words in aphasia treatment highlighted the need to select and treat words other than nouns because the most frequently used spoken words include verbs, adverbs and adjectives (Renvall et al., 2013a & b). Broader aphasia research also influenced the process used to select PR verbs due to the limited literature pertaining to PR words. One paper on the personalization of aphasia treatment *in general* supported the use of goal setting to guide the selection of PR words (Thiessen & Brown, 2021). Research relating to establishing the *functional communication* situations important to PwA and their carers was also influential (Lomas et al, 1989), as was that investigating the *topics* discussed by PwA compared to their neurologically healthy peers (Davidson et al, 2003). Finally, research relating to *conversation analysis* informed the PR verb selection process (e.g., Whitworth et al., 1997). Treatment of structures other than declarative sentences has also been little researched but was supported by the frequency of, for example, questions and imperatives in everyday spoken conversation (Biber et al., 1999).

The literature relating to what might make PR words more amenable to treatment was also reviewed. The literature on adult learning suggests that PR stimuli are conducive to facilitating successful learning (e.g., Hopper & Holland, 2005), whilst there are also indications that PR stimuli are more richly represented neurologically (Renvall, 2013a; Van Lancker and Nicklay, 1992) and this may make them easier to learn (e.g., McKelvey et al., 2010). This also potentially contributes to greater saliency for PR stimuli which is regarded as an important factor inducing the neuro-plastic changes upon which recovery from aphasia relies (e.g., Dignam et al., 2016; Raymer et al., 2008). It is also speculated that treatment of PR relevant verbs may be important because such words are more likely to be used in everyday communication and thus generalization of treatment effects to discourse is more likely (e.g., Webster & Whitworth, 2012). However, due to the small number of studies which have targeted (at least some) PR words, further research is required. Finally, as imperatives and questions frequently occur in conversation (Biber et al., 1999) these are also worthy targets of treatment, yet they are almost absent from sentence treatment studies. There is a need to build the evidence base in this regard, and thus imperatives and questions are targeted in the SPT.

3.5 Self-delivery of computer-based aphasia treatment.

This PhD study investigates the feasibility of a novel SPT, with a low dose clinician-delivered element supplemented by self-managed, computer-based exercises. To inform the development of the computer-based aspects of the SPT, the relevant literature is now reviewed.

3.5.i Background.

Computer-based treatments for aphasia have become an increasingly important resource for PwA (e.g., Floel, 2019; Godlove, Anantha, Advani, Des Roches & Kiran, 2019; Kurland et al., 2018; Palmer et al., 2019). This change has been driven by several forces. The most recent of these has been the COVID pandemic which has enforced the remote delivery of non-urgent treatment, with increased need for patients to self-manage their treatment. For example, the Royal College of Speech and Language Therapists (RCSLT) conducted a survey of SLTs at the start of the pandemic (23 – 29 April 2020) and found that 60.7% of clinicians were now delivering treatment remotely via phone consultations, and 43.6% via video consultation. The survey also found that 70.7% of respondents would like to see some changes to service delivery continue, with the most common change being offering a choice of methods to provide services, particularly through the use of telehealth (https://www.rcslt.org/wp-content/uploads/media/docs/Covid/RCSLT-Survey---impact-of-COVID---FINAL-(1).pdf). Indeed, self-delivery of aphasia treatment via computer can be considered as a component of the self-management of a chronic health condition, an approach now seen as important but which has been neglected in relation to the management of aphasia to date (Nichol et al., 2019; Wray, Clarke & Forster, 2018).

Prior to the pandemic the most important drivers of the increased use of computer-based aphasia treatments were two-fold. Firstly, there was an awareness of the potential for technology to enhance the treatment of language and communication disorders. Examples include the use of a virtual world (Eva Park) to deliver treatment (e.g., Amaya et al., 2018; Marshall et al., 2016), and the use of avatars to deliver proven face-to-face treatments such as script training (Cherney, Halper, Holland & Cole, 2008) and the complex Treatment of Underlying Forms (with computerised treatment seen as particularly suited to the delivery of complex treatments with good fidelity: Thompson, Choy, Holland & Cole, 2010). A second driver is the more pragmatic consideration of the need to supplement the limited face-to-face treatment available to PwA. For example, Clarke et al. (2018) found that therapists on stroke units spend considerable time in information exchange and non-patient contact activity, and that stroke units were understaffed especially with respect to SLTs indicating a lack of treatment for PwA at the acute stage. Palmer et al., (2018) found a similar picture for PwA in the community: the median amount of treatment received by the 278 participants in the usual care arm of their Big CACTUS RCT was just 6.3 hours in the first 3 months post-stroke, at an average of one 60 min session every two weeks**.** As Floel (2019) states:

> "Given that self-administered computerised SLT offers a low-cost and widely available approach to training, this method might be more compatible with demands of health-care policy makers and insurance companies than face-to-face interventions, when considering

economic constraints and shortages in staff delivering the treatment” (Floel, 2019, p.806)

In addition to the economic argument in favour of self-managed, computer-based aphasia treatment, there is evidence that a larger dose of aphasia treatment is required for it to be effective. For example, Palmer et al. (2018) analysed 27 RCTs of aphasia treatment and found that functional communication improved significantly for PwA who received treatment at high intensity (4 - 15 hours a week), high dose (27 - 129 hours in total), or over a long period of time (up to 22 months), compared with those who received lower intensity or dose. The requirement for larger/more intense doses of aphasia treatment likely has a neurophysiological basis. That is, the neuroplastic changes which underlie (re)learning following brain damage are likely to require intense (*massed*) practice (e.g., Dignam et al., 2016; Kleim & Jones, 2008; Raymer et al., 2008). However, as noted above, PwA are not routinely receiving adequate doses of treatment. There is thus a clear need to supplement publicly funded aphasia treatment, and the literature on computer-based treatments clearly demonstrates that such treatments have the potential to do this (see final section).

Whatever the motivation for the development of computer-based aphasia treatments, there are challenges. One crucial challenge is to ensure that computer-based aphasia treatments are acceptable to PwA, their carers and to clinicians, and the small evidence base pertaining to this will now be reviewed.

3.5.ii The acceptability of self-delivered, computer-based aphasia treatment for people with aphasia, their carers and clinicians.

The acceptability of computer-based treatment to PwA is likely to be an important factor in the uptake of treatment delivered in this way, with the views of carers and clinicians also likely to be influential. However, despite the proliferation of aphasia treatment apps, relatively little is known about how acceptable PwA find them. Kearns, Kelly and Pitt (2019) carried out a literature review of self-reported feedback on the usability, feasibility and acceptability of aphasia treatment delivered by Information and Communication Technologies (ICT: computers, tablets and smartphones and in conjunction with software), and included studies which had any measure of self-reported feedback. The resulting review was of 17 studies, with 14 of the reviewed studies requiring participants to wholly or partly self-deliver treatment.

The findings of the review were grouped under three categories: *perceived gains, usability* and *engagement* with the ICT-based mode of rehabilitation. In terms of *perceived gains*, most participants reported improvements in language skills. However, perceived gains were not always in line with actual improvements in language outcome measures (where these were reported). Other perceived gains were of increased confidence and increased engagement with activities outside of treatment, with the independence and autonomy associated with self-delivered treatment valued, as was the ability to carry out repetitive practice. Regarding *usability,* nine studies investigated this with 83 of 85 participants rating their satisfaction and/or enjoyment of the treatment program as high. However, some negative emotions were also reported - most commonly frustration and this was most often associated with a language difficulty (e.g., not being able to say a word targeted in treatment). In terms of *engagement* with treatment, an overall theme emerged pointing to the desirability of a combination of face-to-face and computer-based treatment, particularly in relation to *training* in the use of the computer-based treatment, and in resolving any technical issues. In summary, the review found that overall participants reported positive experiences and satisfaction with computer-based treatment.

Palmer et al. (2013) and Wade et al. (2003) report on the acceptability of the StepByStep programme. Wade et al. found that participants valued the increased autonomy and opportunity to practice offered by the self-delivered, computer-based treatment, but saw the supervisory role of an SLT as crucial. Palmer et al. also found that PwA appreciated the independence and repetition afforded by StepByStep, although this could cause fatigue and interfere with other commitments. The involvement of carers and/or volunteers was seen as important to motivation, as was the personalization of treatment. Further advantages and disadvantages highlighted in both studies included perceived improvement in several language areas (such as word finding, sentence production and spelling), and improvements in confidence, participation, independence, and motivation. Interestingly, both studies reported that improvements in communication following the computer-based treatment were attributed not just to improvements in language but also to improvements in confidence. Input from an SLT at the *start* of treatment was seen as important to ensuring that exercises of an appropriate level of difficulty were set. The difficulty level of exercises was seen as key to both motivation (i.e., that exercises were not so easy as to be un-motivating), and to reducing anxiety and fatigue which were reported when exercises proved too difficult. Access to *continued* support was also seen as important. Carers and participants gave examples of multiple ways in which they felt supported by volunteers including technical support (to carers as well as the PwA), and visits by volunteers being motivating, giving social opportunities and opportunities to use trained words in real life.

Finally, it is important to note that Palmer et al. (2013) reported that although computer-based treatment was deemed acceptable (with participants and carers acknowledging in particular that it allowed continuation of treatment at the chronic stage of aphasia), overall face-to-face treatment was preferred. Disadvantages of self-delivered computer-based treatment included lack of a perceived effect of treatment (for four participants: n=4), lack of independence in using the treatment (n=4), frustration (n=3) and reduced time for other activities (n=3). With regard to the latter, the comments made by participants indicated that this was related to lack of independent use of the treatment meaning the carers had to assist the PwA rather than carry out their normal daily activities, whilst one participant was attempting to practise much more than requested. This led to the authors recommending that a maximum amount of practice be suggested to participants as well as a minimum. Overall, more advantages than disadvantages of computer-based treatment were reported, and Palmer et al. (2019) confirmed the lack of adverse effects of computer-based treatment in their large scale RCT of StepByStep treatment. No serious adverse events were attributed to participating in the computer-based treatment although 27% of participants reported feeling overtired or anxious/worried as a result of treatment. It seems prudent therefore to follow the advice of Palmer et al. (2019) that consideration is given to fatigue and lifestyle when recommending practice schedules. This principle was applied to the SPT in that, before embarking on the SPT programme, practice schedules were discussed with participants, with any barriers to practice identified and solutions discussed as part of a Goal Action Planning (GAP) process (e.g., Scobbie et al., 2011).

In summary, key themes which emerge from the limited number of studies investigating the acceptability of computer-based treatments are that, overall, PwA regard them as an acceptable alternative to face-to-face treatment. However, there are strong indications that this acceptability is dependent on self-delivered treatment being accompanied by supervision from a clinician with subsequent support from either a clinician or a volunteer. In terms of perceived benefits of computer-based treatments, key themes which emerged were that most participants perceived that their language had improved because of treatment. However, these perceived gains were not always borne out by improvements in language-based outcome measures, but it was not possible to comprehensively explore this relationship since most studies did not report on subjective and objective measures together. Kearns et al. recommend that this approach be taken in the future.

Swales, Hill and Finch (2016) report the only study to date that has investigated the *views of SLTs* on the desirable features of computer-based aphasia treatment programmes (CBATs). They analysed the views of 10 SLTs on three questions: (1) What are the advantages of current CBAT programmes? (no specific programmes were identified); (2) What are the limitations of current CBAT programmes? and (3) If the sky was the limit, what features would you like in your ideal CBAT? (p.318). The views expressed by the participants were analysed using framework analysis. They fell into five themes:

(1) *Therapy activities*. Respondents highlighted the lack of variety of therapy tasks in existing CBATs, and particularly in relation to expressive language. They also pointed to the need for more functional activities such as ordering in a cafe and using the phone.

(2) *Stimuli*. All participants agreed that the ability to personalise therapy stimuli was paramount, in terms of uploading personal photos and videos, plus the facility to choose categories of stimuli that were of personal relevance. Most participants noted the concentration on nouns in therapy programmes and requested a greater variety of word classes, with particular reference to the desired presence of verbs (p. 321).

(3) *Cues*. Participants requested that the cues provided in CBATs were given in a hierarchy that mimicked the one in face-to-face sessions; were multimodal and were individualised i.e., that the clinician could adjust them to suit the client.

(4) *Access*. Clinicians highlighted the importance of CBATs being user-friendly not just from the perspective of the PwA but also from their perspective and that of therapy assistants and carers. Thus CBATs should be navigated easily including when customising therapy activities for PwA. Aphasia-friendly features were requested on the HCI and these were in line with the features identified by Brandenburg, Worrall, Rodriguez and Copland (2013) discussed in the following section. The capacity to use a CBAT on more than one platform was seen as crucial

to facilitate maximum access to PwA, rather than apps which could only be used on Apple or Android devices. This was also seen as important in potentially facilitating practice as PwA would be able to use the CBAT on a variety of devices they might have access to in their daily lives. Expensive CBAT licenses were seen as a significant factor in limiting access as services may not be able to afford the license or have to limit the number of clients using the CBAT.

(5) *Progress data*. Participants' ideal CBAT would have remote monitoring of client performance, including detail of client performance such as level of success, latency, number of attempts and cues used. Graphical reporting of client data was suggested as facilitating quick interpretation of results.

Such elements deserve consideration in the development of future computer-based aphasia treatments. However, there are significant resource implications with regard to the implementation of some of these elements, for example for the clinician supervising self-delivered treatment ((2) and (3)), and in terms of software development for all elements (e.g., the StepByStep programme is now in its fifth iteration to incorporate technological advances such as voice recognition capability (https://aphasia-software.com/)).

In summary, the views of SLTs on the desired features of computer-based treatment very much mirror those of PwA and their carers regarding the features which make computer-based treatment acceptable. Common themes are the need for the HCI to be accessible, for treatment to be personally relevant and to have a treatment programme where the level of difficulty can be adjusted to suit the ability of a PwA. Challenges remain in how to ensure that computer-based treatment incorporates functionally relevant activities, and this in turn may mean that computer-based treatment is more likely to impact on functional communication (e.g., Palmer et al., 2019).

3.5.iii The design of an aphasia-friendly human computer interface (HCI).

Although computer-based aphasia treatments are overall acceptable to PwA, their carers and clinicians, the section above also demonstrates some difficulties with their use, including frustration and difficulties with independent use. In order to minimise such disadvantages, it is important that the HCI of computer-based aphasia treatments is as accessible as possible. The literature relating to the design of aphasia-friendly HCIs is therefore briefly reviewed. This is an emerging field of research with few studies currently existing. Brandenburg et al. (2013) carried out a review of the literature relating to the accessibility of mobile technology for PwA. They identified six key design features that may enhance accessibility (see Brandenburg et al., 2013 for a fuller discussion). These key features are *aphasia-friendly text, multimodality, large buttons,* a *stable interface, simple navigation,* and *visual simplicity.* The latter five of these key features are unique to HCIs. The key design feature of *aphasia friendly text* is derived from the (small) research literature relating to the design of aphasia-friendly *written* materials (e.g., Rose, Worrall, Hickson & Hoffmann, 2012 and 2011; Rose, Worrall & McKenna, 2003; Wilson & Read, 2016). The features of aphasia-friendly text are: i) simplified vocabulary and syntax[10], ii) the use of a large, sans serif font, iii) increased white space and iv) the provision of graphics to support comprehension of text (e.g., Brennan, Worrall & McKenna., 2005; Rose et al., 2003).

With regard to design features unique to HCIs, Wilson, Galliers, Roper, Macfarlane and O'Sullivan (undated) set out nine design principles on designing aphasia-friendly HCIs (http://languagelightux.org/). They state that an HCI should be: i) consistent, ii) give clear feedback, iii) show what's happening (i.e. make areas of the interface which are key to navigating the software obvious), iv) provide prompts (for what to do next), v) let users control the pace, vi)

[10] Brennan et al. (2005) state that simplifying vocabulary means that more frequent words are substituted for less frequent ones, and simplifying syntax included shortening long and complex sentences, making passive sentences active and clarifying the referents of pronouns

minimise interaction (i.e. reduce the amount of effort taken to navigate the HCI), vii) minimise language, viii) keep actions direct (i.e. don't use hierarchical menus) and ix) minimise distractions. There is clear overlap between Brandenburg et al.'s key design features and Wilson et al.'s guidelines (see Table 2 below italicised items) but also some different recommendations (see Table 2 below non-italicised items). These differences likely stem from Wilson et al's guidelines being intended to guide design of a wide range of technologies: they state that the guidelines are intended for designers of software, apps and websites (opening screen). The features identified by Brandenberg et al. on the other hand, relate specifically to the accessibility of mobile technology.

Table 2 Key design features and guidelines for aphasia-friendly HCIs.

Key Design Features (Brandenburg et al., 2013)	**Design Guidelines (Wilson et al., undated)**
aphasia friendly text	*minimise language*
stable interface	*be consistent*
simple to navigate	*keep actions direct*
visual simplicity	*minimise distractions*
multimodality	give clear feedback
large buttons	show what's happening
	provide prompts
	let users control the pace,
	minimise interaction

It is not clear if such guidance has influenced the development of computer–based treatments in aphasia rehabilitation as, although researchers raise access as important, there is little discussion in their papers of the HCI of interventions (Kurland et al., 2014; 2018; Mortley et al., 2004; Palmer et al., 2012; 2013; Routhier et al., 2016). Clearly such features are important considerations in the development of computer-based treatments in the future.

3.5.iv The effectiveness of computer-based aphasia treatments.

Systematic reviews of health treatments represent the highest level of evidence (e.g., the Oxford Centre for Evidence Based Medicine: https://www.cebm.ox.ac.uk/about-us/about/cebm), but there are few reviews of computer-based aphasia treatments. Allen, Mehta, McClure and Teasell (2012) included four studies of computer-based treatments in their systematic review of aphasia treatments initiated after 6 months post onset. They concluded that computer-based treatments significantly improved naming ability and verbal communication. Zheng, Lynch and Taylor (2016) conducted a systematic review relating to computer-based treatments only. They found only seven studies (five RCTs and two case series), suitable to answer their research questions (was computer-based treatment: i) better than no treatment, and ii) as effective as clinician delivered treatment) but reached a positive conclusion in relation to both questions. The most comprehensive review of computer-based aphasia treatment to date was carried out by Lavoie, Macoir and Bier (2017). They reviewed the effectiveness of computer-based *anomia* treatments and included 23 studies. Fourteen of the reviewed studies investigated self-delivered treatment, with a clinician leading treatment in the other nine studies. Fifteen of the studies treated nouns with the remaining eight studies including verbs. Encouragingly, all 23 studies reported significant improvement in trained items, and this was maintained in all the studies where this was assessed (n=17). Generalisation to untrained items (within level) was reported much less commonly, whilst generalisation to discourse/conversation (across level) was assessed in only four studies and detected in three of these. In light of this, Lavoie et al. highlighted the importance of i) targeting PR items in treatment and ii) of assessing the impact of treatment at the level of functional communication and discourse. They also noted the potential of computer-based treatments to be motivating for PwA, to increase the intensity and duration of treatment, and to increase self-esteem as a result of independent administration of treatment. However, the authors also expressed concern that so few computer-based treatments targeted verbs (because of their pivotal role in sentence production).

In addition to the few systematic reviews of computer-based aphasia treatment, there are a limited number of studies of *self-managed* computer-based treatments which also constitute a high level of evidence. Des Roches et al. (2015) and Braley et al. (2021) investigated the effectiveness of Constant Therapy in a (relatively) large scale group study (n=42) and an RCT (n=32) respectively. (Constant Therapy contains a suite of evidence-based language and cognitive exercises which are self-managed: https://constanttherapyhealth.com/constant-therapy/). Both Des Roches at al. and Braley et al. found significant changes on impairment-based assessments of language post-treatment, with improvements in the rating of quality of life also found by Braley et al. Separately, Palmer et al. conducted a pilot RCT (2012) (n =15) and then a full-scale RCT (2019) (n =83) of StepByStep therapy software. (StepByStep also contains a suite of evidence-based exercises for language which can be self-managed: https://aphasia-software.com/.) In both studies, Palmer et al. found significant improvement in naming ability, but no improvement to functional communication or real-life conversation (Palmer et al., 2019). Finally, Kurland, Liu and Stokes (2018) carried out a group study (n=21) of the effectiveness of an iPad-based verb and noun treatment, and also found significant improvements in naming. As such, naming ability can be significantly improved by self-managed treatment whether it is linked to software (e.g., Des Roches et al., 2015) or to a device (e.g., Kurland et al., 2018). Across all four studies there was wide variation in the amount of practice self-managed (from 2.3 sessions per week (Palmer et al., 2019) to 6.2 sessions per week (Des Roches et al., 2015). There were also conflicting findings regarding the relationship between compliance with treatment and degree of improvement. Thus, Des Roches et al. (2015) report that their participants who self-delivered Constant Therapy improved significantly more on standardised measures than the control group who only had a weekly face-to-face treatment session. They state that the only explanation for this appears to be the difference in the amount of treatment received by the experimental and control groups (average 4 hours 49 minutes per week for the experimental group versus average 40 minutes per week for the control group). However, due to the extent of individual variation, they also report no overall relationship between response to treatment and compliance. Kurland et al. (2018) on the other hand, report a positive correlation between compliance and response to treatment, with greater gains on treated items and fewer losses on untreated items for those participants who practiced the most.

To more specifically inform the development of the novel treatment developed in this study, the larger scale studies which investigated *self-managed*, computer-based treatment and *included verbs* (i.e., similar to the SPT) were reviewed and are reported in Table 3.3 alongside three smaller scale studies identified in the systematic review (reported in Chapter 1: Kurland et al., 2014; Mortley et al., 2004; Routhier et al., 2016). Table 3 below presents the effectiveness data for these studies and demonstrates overall that self-managed, computer-based treatment was effective in improving retrieval of trained verbs with significant change reported for 100% of participants (n=34) for whom individual results were reported. (Palmer et al. (2012) also reported significant improvement for their *group* of participants in their pilot RCT (n=15) and in their subsequent large scale RCT (n=83) (Palmer et al., 2019))[11]. Table 3.3 also highlights the need to measure the impact of computer-based treatments on levels of communication other than trained words, an omission addressed by this PhD study.

[11] Palmer et al. (2019) treated a minority of verbs: 98/9,999 chosen PR words. Palmer et al. (2012) treated an unspecified number of verbs: 48 words from the OANB and 48 PR words.

Table 3 The impact of self-managed computer-based treatment that included verbs: on verb and sentence production, functional communication and discourse. (Abbreviations: NA = not assessed; TC = transcortical; TCM = transcortical motor; TCS = transcortical sensory; + = numerical improvement; + = statistically significant improvement; - = no improvement).*

1.Authors	2.Study Aim	3.Number of participants & type of aphasia	4. Trained Verbs in Isolation	5. Untrained Verbs in Isolation	6. Trained Verbs in Sentences.	7. Untrained Verbs in Sentences.	8.Functional Communication	9.Discourse
Kurland, Liu & Stokes (2018)	To establish the effectiveness of a home treatment for noun & verb retrieval delivered via iPad.	n = 21 6 anomic; 3 Broca's; 3 Wernicke's; 2 TCS; 2 TCM; 2 optic; 1 conduction; 1 mixed TC; 1 global	**+*** for all participants	-	NA	NA	NA (anecdotal reports of improvement)	NA
Kurland, Wilkins & Stokes (2014)	To establish the effectiveness of a home treatment for noun & verb retrieval delivered via iPad .	n=5 3 anomia; 1 TCS; 1 Wernicke's	**+*** for all participants	-	NA	NA	NA (anecdotal reports of improvement)	NA
Mortley, Wade & Enderby (2004)	To assess the effectiveness of StepByStep computer-based treatment delivered remotely	n=6 nonfluent	**+*** for all participants	NA	NA	NA	NA (interviews with carers and participants report positive benefits (see Wade et al., 2003)	NA
Palmer et al., (2019)	Full scale RCT to assess the effectiveness of StepByStep computer-based treatment	n = 83 all types of aphasia included; mild, moderate and severe;	**+*** for treated group	NA	NA	NA	-	-
Palmer et al., (2012)	Pilot RCT to assess the feasibility of a full-scale RCT of StepByStep	RCT n=15	**+*** for treated group	NA	NA	NA	NA (interviews with carers and participants report positive benefits (see Palmer et al., 2013)	NA
Routhier, Bier & Macoir (2016)	To assess the effectiveness of semantic-phonological treatment for verb anomia self-delivered by a tablet.	case series n=2	**+*** for both participants	**+** for P1 - for P2	NA	NA	NA	NA

In summary, the evidence base for self-managed, computer-based verb treatments is very small, with no studies of computer-based, self-managed *sentence* treatments. It is also of note that only three studies investigated treatment that was specifically adapted to take into account the unique qualities of verbs. Kurland et al. (2014; 2018) and Routhier et al. (2016) used *video* stimuli for the actions treated in their studies, whilst static *pictures* were used for target nouns. As noted by Hickin et al. (2020), the limited research suggests that video stimuli may facilitate action verb retrieval better than picture stimuli (Blankestijn-Wilmsen et al., 2017), and indeed it may be the case that developing treatments which target the unique features of verbs may make them more potent or active than techniques derived from anomia treatments. There is therefore an urgent need to implement treatments specifically developed for verbs (such as semantic feature analysis: Wambaugh & Ferguson, 2007; Wambaugh et al., 2014) - and sentences (such as mapping treatments: e.g., Jones, 1986; Nickels et al., 1991) - using computer-based exercises which can be self-delivered. (See Chapter 4 for a detailed description of the SPT exercises and the rationale for each).

The need for *theory-based,* computer-delivered verb and sentence treatments is also underlined by a search of the Aphasia Software Finder (https://www.aphasiasoftwarefinder.org/advanced-software-search). Searching the Finder for 'Talking in Sentences' identified 17 treatment apps/software, 14 of which targeted spoken verb or sentence production. However, only seven of these 14 apps were deemed to be theory-based on the Software Finder Checklist, and for these it was unclear if/how treatment was informed by the verb and sentence treatment literature (e.g., there was no evidence that VNeST had informed any of the treatments although this has a good level of evidence, see Hickin et al, 2021).

Finally, despite the limited (if growing) evidence for the effectiveness of self-managed computer-based treatments, their use has proliferated over the last decade. For example, the Constant Therapy website reports that 150 million Constant Therapy exercises have been completed, with 400,000 downloads and 37,000 clinician users (https://constanttherapyhealth.com/constant-therapy/ accessed 4 October 2021). Indeed, Godlove et al. (2019) present data from 2013 – 2017 on 3,686 users of Constant Therapy, and Munsell, De Oliveira, Saxena, Godlove and Kiran (2020) report on data from 2850 users of Constant Therapy who, between October 2016 and January 2019, completed an average of 8.6 weeks of treatment at a frequency of 1.5 days a week. Godlove et al. demonstrated similar outcomes of treatment for participants who either used Constant Therapy in the clinic only versus those who practiced at home only, but that home users were able to reach targets more quickly because of the greater amount of treatment they were able to self-deliver. Munsell et al. report on the factors which influenced compliance with self-delivered treatment. They found that older participants (>50 years) self-delivered more treatment than younger ones, and that people who lived in rural areas (where it was harder to access treatment) were more engaged than people who lived in urban areas. Use of remote treatments has almost certainly proliferated further due to the pandemic (as noted earlier in the RCSLT survey of clinicians). Braley et al. (2021) demonstrate that it is feasible to deliver self-managed, computer-based treatment (Constant Therapy) entirely remotely, noting that this is especially important given the ongoing COVID 19 pandemic.

As noted earlier, Constant Therapy has an evidence base, as does, for example StepByStep and Sentence Shaper (e.g., Linebarger, Romania, Fink, Bartlett & Schwartz, 2008; McCall, Virata, Linebarger & Berndt, 2009). However, many other commercially available computer-based aphasia treatments do not have an evidence-base but are being widely used. Kurland et al. (2018) lament that:

> “the evidence base has taken a back seat to the proliferation of commercially available aphasia-specific apps.. (which).. seem to be routinely utilised in clinical practice” (p.1140).

Cann (2021) conducted a survey of how PwA (n =70), their family and friends, and clinicians use technology, published in the Bulletin of The Royal College of Speech and Language Therapists. Aphasia treatment apps/software listed by respondents included evidence-based software such as StepByStep and Constant Therapy, and also apps such as TACTUS, Aptus and CueSpeak, with, as Cann remarks, few of the latter having an evidence base. Indeed, Cann found that PwA were much more likely to use free or single payment apps (80%) than subscription apps, suggested that cost was a prime consideration in choosing an app as opposed to evidence of its effectiveness. In order to assist PwA to choose evidence-based apps, she suggests that a more robust system for evaluating apps/software needs to be developed, including the possibility of evaluating and kitemarking relevant apps. Indeed, the latter would also assist clinicians: the RCSLT states that SLTs providing self-managed computer programmes to service-users must be aware of the evidence base behind the programme, and that a complete assessment, support and evaluation is still required for a self-managed programme (https://www.rcslt.org/members/delivering-quality-services/self-managed-computer-therapy/self-managed-computer-therapy-guidance/#section-3). This PhD study attends to a number of pillars of underpinning evidence and literature that make the development of novel treatments involving computers complex.

3.5.v Conclusions.

In summary, the few studies investigating self-managed, computer-based verb treatments indicate that they can effectively improve production of trained items. However, the evidence-base is very small, and non-existent for self-delivered *sentence* treatments. There is thus an urgent need for further research, particularly of *theory-based* verb and sentence treatments delivered via computer. Future research also needs to investigate the impact of treatment on untrained items and at other levels of communication, particularly functional communication and discourse. Further work is also required to ensure that self-managed computer-based treatments are maximally accessible and acceptable to PwA, their carers, and to SLTs prescribing computer-based treatments. Despite a limited (if growing) evidence-base, the use of self-managed, computer-based treatments has proliferated to supplement limited publicly funded aphasia treatments, and the need for treatments to be delivered in this way has increased exponentially due to the COVID 19 pandemic underlining the vital importance of pursuing this field of aphasia research.

Chapter 4. Methodology.

This chapter begins by setting out the research questions for this feasibility study. It then describes the research design used for the study and the rationale for this. The outcome measures (OMs) used to address the research questions are described, with those used to investigate the feasibility of the SPT treatment (research questions (RQs) 1 – 4) described first, and the outcome measures used for preliminary efficacy testing of the SPT reported last (RQ5).

4.1 Research Questions and Study Design.

The study aimed to answer the following research questions:

1. Is it *feasible* to deliver and self-manage personalised sentence production treatment (SPT) by computer? Specifically, is it feasible to a) recruit and retain suitable participants to the SPT, b) self-deliver the SPT using a computer, and c) is it feasible to select a set of PR verbs?

2. Is the SPT *acceptable* to the participants with aphasia and their significant others (SOs)?

3. What factors influence *compliance* with self-delivered computer treatment?

4. Were treatment procedures carried out during the SPT administered with acceptable *fidelity*?

5. Preliminary efficacy testing of the effect of the SPT on the production of i) trained and untrained verbs, ii) untrained nouns, iii) sentence production using trained and untrained verbs, iv) verb and sentence production in discourse and v) in functional communication as perceived by i) the participants with aphasia themselves and ii) their significant others?

The study design employed to investigate the RQs above was a pre-post intervention design as defined by the Single-Case Reporting Guideline in BEhavioural Interventions (SCRIBE) Statement (Tate et al., 2016). A pre-post intervention design is not categorized as a single case experiment by SCRIBE because it does not take repeated measurements of the dependent variable *during* the intervention phase (unlike, for example, multiple baseline designs). However, a pre-post intervention design was chosen for this study as this design is suitable for providing *indicative* information about efficacy, answering the question “can it work” before undertaking a full trial to assess “does it work” (Bowen et al., 2009 – see discussion below). A pre-post design was also the predominant experimental design found in the literature reviews which informed the study. Of the 37 *verb-in-isolation* treatment studies reviewed in Hickin et al. (2020), a pre-post intervention design was used in 13 studies, with the next most common being a withdrawal/reversal design (with multiple probing during both pre/post and treatment phases) used in 11 studies. Of the 33 *sentence* treatment studies reviewed in Hickin et al. (2022), a pre-post intervention design was also most common (in 19 studies), with the next most common again being a withdrawal/reversal design (12 studies). There was a clear geographical influence on the choice of experimental design, with a withdrawal/reversal design very predominantly chosen in American studies, whilst British studies predominantly chose pre-post designs. A possible reason for this is that, for ethical permission to be granted, UK studies must demonstrate that they minimize the potential for harm to research participants, including in terms of assessment burden. As noted by Howard, Nickels and Best (2015), multiple probing of performance may be onerous or even not feasible for some participants. The systematic review of verb treatments (Hickin et al., 2020) found that some studies did indeed report that multiple probing resulted in fatigue effects (e.g., Wambaugh et al., 2004) meaning that the number of items probed and/or the frequency of this had to be reduced (e.g., Rodriguez et al., 2006; Rose & Sussmilch, 2008). This was a major consideration underlying the choice of a double baseline pre-post treatment design (as opposed to multiple probes) for this study, as ethical approval required that the assessment burden was minimised. Additionally, multiple probing involves repeated exposure of untrained items which may itself result in improved performance (e.g., Nickels, 2002), and also limits the size of the set of treated items (e.g., to 10) which may lead to over estimation of treatment effects (Howard et al., 2015). It is, however, acknowledged that not using multiple baselines before and during treatment reduces the rating of the quality of the design according to the SCED rating scale (Tate et al., 2008) and the RoBiNT (Tate et al., 2013). Not employing multiple baselines means this study may be vulnerable to not controlling confounding factors (e.g., spontaneous recovery).

Whilst n of 1 study designs have limitations, they are considered an appropriate design for feasibility studies such as this one (e.g. Bowen et al., 2009; MRC guidance on Developing and Evaluating Complex Interventions, 2019: https://mrc.ukri.org/documents/pdf/complex-interventions-guidance/). Bowen et al. (2009) highlight the important role of preliminary, feasibility studies in ensuring that funding for full scale evaluation of interventions is directed at those which are *most likely* to be efficacious. Bowen et al. discuss the circumstances when a feasibility study is appropriate, and these include when there are few existing studies or existing data relating to a specific intervention technique. This is the case for verb and sentence treatments self-delivered by computer and supports the need for the feasibility study carried out for this PhD. Bowen et al. propose that there are eight areas which should be addressed by feasibility studies and six of these are addressed by this study: *acceptability (RQ2), implementation* (similar to fidelity: how far intervention can be implemented as planned – RQ4), *demand* (or the extent to which the intervention is used – RQ3), *practicality (RQ1)*, *adaptation* (the need for the intervention to be modified – in this study for delivery by computer instead of face-to-face delivery, described in section 4.6 below and addressed by RQ2) and *limited efficacy testing* (RQ5)[12]. Bowen et al. suggest that feasibility studies should address the questions of *can* an intervention work which was the main aim of this research study, with the question of *does* it work answered by larger scale efficacy and effectiveness studies. The use of single cases also allows initial exploration of who might and might not benefit at the early stage of developing and testing a novel intervention (e.g., Howard et al., 2015).

The need for feasibility studies is of particular importance in the development of *complex* interventions (MRC Guidance on Developing and Evaluating Complex Interventions (2019: (https://mrc.ukri.org/documents/pdf/complex-interventions-guidance/) such as the SPT investigated in this study. Complex interventions are defined as such because of the number of - and interactions between – components (ibid), and, as was apparent from the scoping reviews of verb-in-isolation and sentence treatments reported in Chapters 1 and 2, both types of treatments are complex, with sentence treatments particularly so. Interventions can also be complex because of the number and complexity of *behaviours* which may be required of *recipients* during treatment, and this was the case for the SPT investigated in this study because it was predominantly self-delivered by computer. In summary, complex interventions may be both more difficult to *implement* and to *evaluate* and this underlines the importance of establishing their feasibility prior to large scale efficacy and effectiveness studies (ibid).

---

[12] The remaining two areas relate to aspects of health promotion interventions which are the focus of Bowen et al's paper: *integration* (the degree to which an existing infrastructure needs to change) and *expansion* (the use of an existing intervention in a different setting or with a different population).

This feasibility study followed the MRC guidance on the *development* phase of complex interventions in that the intervention (the SPT) was based on a *systematic (scoping) review* of the evidence (for both verb and sentence treatments) to ensure it was based on the *best available evidence*. This also suggested it was *likely to be effective* (based on face-to-face implementation) and ensured that there was a "*coherent theoretical basis*" to the SPT (as described in section 4.7 and in the treatment manual included as supplementary material). The MRC guidance suggests that the *piloting and feasibility* phase of a complex intervention should address whether an intervention can be delivered *as intended* and whether *recruitment and retention* rates suggest that the treatment should go forwards to the next phase of efficacy and effectiveness testing: these issues were addressed in this study (RQs1 and 4).

Robey and Schultz (1998) also discussed phases of research with specific reference to treatment for communication disorders. This feasibility study of a novel SPT is categorised as a Phase I research study because, for example, it seeks to *establish a treatment effect* and the *safety* of treatment. Robey and Schultz state that single cases are appropriate designs to accomplish the objectives of Phase 1 research (along with case reports and small group studies). Indeed, the prevalence of single case studies (and case series) in the systematic reviews of both verb and sentence treatments testifies to the developmental nature of these fields.

In summary, the design for this study used a pre-treatment double baseline (8 weeks apart) with one immediate post-treatment assessment and one follow-up assessment also eight weeks later (see Table 4) employed across a series of single case studies[13]. The study design is categorised

[13] The initial study design submitted for ethical approval was a longer study with baseline, treatment and follow up phases of 12 weeks each. However, feedback from the ethics committee required that this be shortened to reduce the commitment required from the participants.

as Approach B by Howard, Nickels and Best (2015) as it has the following characteristics: i) both baseline and therapy phases continue for a predetermined number of trials, ii) performance on treated and untreated items is assessed sparsely - usually only pre- and post-treatment, iii) the effectiveness of treatment is assessed using inferential statistics to compare performance pre- and post-treatment for treated and untreated items, and iv) the number of items in the treated and control sets is larger (n = 20–50). Twenty PR verbs were treated in this study guided by the verb-in-isolation treatment review where this was the most common number of verbs treated (plus 20 matched, untreated control verbs). The most common number of verbs treated in sentence treatments was less (10; Hickin et al., 2021), however Webster and Whitworth (2012) in their review of both types of treatment suggest that treating a larger number of verbs is more likely to result in functional gains which was an aim of this study.

Table 4 Study Design.

Study Design

Ax1	**Pre-Tx Phase**	Ax 2	**Tx Phase**	Ax 3	**Maintenance Phase**	Ax 4
	8 weeks		8 weeks		8 weeks	

4.2 Outcome Measures.

The outcome measures (OMs) used to answer each research question are described below with the OMs used to address aspects of feasibility (RQs 1 – 4) described first. The OMs are also summarised in Table 5 towards the end of the chapter. The chapter finishes with a more detailed description of the assessments used to address RQ 5 (preliminary efficacy testing) together with a rationale for the selection of each assessment.

1. Is it *feasible* to deliver and self-manage personalised sentence production treatment (SPT) by computer? Specifically, is it feasible to a) recruit and retain suitable participants to the SPT, b) self-deliver the SPT using a computer, and c) is it feasible to select a set of PR verbs?

The approach taken to assessing the feasibility of the SPT programme was informed by the limited literature in the area (e.g., Bowen et al., 2009) as discussed above. Thus, participant *recruitment* and *retention* were monitored *as were compliance, acceptability* and *fidelity*. Other studies investigating the feasibility of computer-based aphasia treatment have adopted a similar approach (e.g., De Cock et al., 2021; Palmer et al, 2012; Pitt, Theodorus, Hill & Russell, 2019; Woolf et al., 2016).

For ease of reading, the recruitment process for participants to this study will now be described (RQ1a).

4.2.i Recruitment Procedures for participants with aphasia.

Potential participants were alerted to the study by several different routes. Firstly, the City Aphasia Research Register held by the Division of Language and Communication Science (LCS) was accessed to identify potential participants. (Ethical permission for new research studies to access the database was granted by the City University of London Senate Research Ethics Committee in September 2016). Potential participants or their SOs were contacted either by email or phone and given details of the project. If they expressed interest in participating, as much information as possible regarding their suitability for the project was gained at this point to avoid participants undertaking screening assessment unnecessarily. (e.g., whether potential participants were more than six months post stroke and whether they had access to a computer). Potential participants were also contacted via aphasia self-help groups who were given (aphasia-friendly) information about the research. If a group (or a member of a group) expressed interest the research student visited the group to give further information about the study. Finally, potential participants came from relatives of PwA directly contacting the Division of Language and Communication Science regarding any research studies that might be suitable for their loved one.

4.2.ii Inclusion and Exclusion Criteria.

The rationale for the inclusion and exclusion criteria used in this study was informed by the systematic scoping reviews of verb and sentence treatment studies reported in Chapters 1 and 2 and this rationale will now be described.

In order to minimise the screening of unsuitable participants, the following exclusion criteria were established *prior to face-to-face screening* whenever possible (e.g., during initial contact emails/phone calls):

- **No access to a computer (desktop, laptop, tablet, or iPad):** because treatment was delivered via computer and the research did not have funding to supply potential participants with devices.
- **Aphasia not due to a cerebrovascular accident (CVA):** aphasia due to other neurological damage (e.g., a brain tumour) can be accompanied by concomitant factors which may influence response to treatment (e.g., recurrence of tumour).
- **Less than 6 months post stroke:** to exclude the effects of spontaneous recovery which might have masked treatment effects.
- **Multiple CVAs and cognitive impairment additional to that associated with ageing:** Participants with multiple CVAs and/or other concomitant neurological disorders (e.g., dementia) may have additional cognitive deficits. Such deficits were likely to affect the ability of participants to respond to treatment and as such would act as a confounding variable making interpretation of the response to treatment difficult. Participants therefore had to perform within normal limits on the Semantic Memory and Recognition Memory

subtests of the Comprehensive Aphasia Test (CAT) Cognitive Screening Battery (Swinburn et al., 2004).

- **Limited proficiency in English as a second language:** If a potential participant spoke more than one language, they could be included providing they identified English as their main language (because the treatment was delivered in English and aimed to improve sentence production in English).

The following exclusion criteria were established via a *face-to-face screening session* which took a maximum of 40 minutes**.**

- **Insufficient comprehension to understand and consent to inclusion in the study:** Participants with impaired comprehension of *both* spoken and written material were deemed unlikely to be able to give truly informed consent to participate in the research study. Participants' comprehension was therefore screened using the CAT (Swinburn et al., 2004). Those who scored less than 26/30 on Spoken Word Picture Matching *and* less than 28/30 on Written Word Picture Matching subtests were excluded.
- **Severe apraxia of speech:** A likely active ingredient of sentence production treatment is the ability to repeat spoken cues (e.g., Conroy, Sage & Lambon-Ralph, 2009) thus participants were expected to score at least 75% on a test of repetition (CAT Repetition of Words n =16). This criterion was used by Conroy et al., in a series of successful verb treatment studies using spoken cues (Conroy, Sage & Lambon-Ralph, 2009a, b and c).
- **Severe naming impairment:** Participants scoring less than 15% (i.e., 3/20) of 20 nouns and 20 verbs from the (Object Action Naming Battery (OANB), Druks & Masterson, 2000) were

excluded, as they were deemed unlikely to respond to treatment (as indicated by Palmer et al. (2012) and Conroy et al. (2009)).

- **Mild naming impairment:** Participants had to score no more than 80% on a naming screen of 20 verbs taken from the OANB (Druks and Masterson, 2000; Edmonds et al., 2015). An upper cut-off of 80% was proposed to reduce the likelihood of ceiling effects.
- **Insufficient manual dexterity to use the computer delivered treatment exercises**: Potential participants' manual dexterity was assessed using a form of dynamic assessment (Caute, personal communication, 2017). The procedure assessed the degree of support potential participants required when opening and using example treatment slides (e.g., hand-over-hand versus spoken instructions alone), as well as their ability to learn from instructions given during the assessment. Those participants who could use the treatment slides independently were accepted into the research study, as were those who responded adequately to instruction during assessment.
- **Mobility.** Participants who had the ability to travel were seen at the university clinic. Participants who were otherwise suitable but unable to travel had to live within an hour's commute of the university or of the researcher's home so that domiciliary visits could be made to deliver treatment.

4.2.iii Recruitment Procedures for SOs.

If a participant with aphasia had a SO who was willing to participate in the study the following inclusion criteria were applied to them:

- **Length of relationship with participant with aphasia:** a SO must have known the participant for at least 6 months prior to the start of the study. This was to ensure that they had a reliable view of their friend/relative with aphasia's ability to communicate prior to their treatment, in order to complete the CETI questionnaire (Lomas et al., 1989) in relation to their friend/relative at 4 four points during the study: twice before and twice after treatment.
- **Proficiency in English as a second language:** if a SO spoke more than one language, they could be included providing they identified English as their main language (because the exit interview and questionnaire they were required to complete were in English). This was also important to ensure that SOs understood the information they were given about the study so that truly informed consent was attained. Participants with aphasia who did not have a suitable/consenting SO were *not* excluded from the study since it was a pilot, and the presence/absence of a SO was a factor of interest in terms of the compliance of participants and the success of treatment.

4.3 The feasibility of self-delivering the SPT.

The feasibility of self-delivering the computer-based SPT was assessed by logging any technical difficulties experienced by participants in clinical notes: the nature of the difficulty, whether it was resolved, and the type and amount of help required to do this (RQ1b). Assessment of technical feasibility was informed by a review of the literature relating to computer-based aphasia treatments - reported in the latter half of Chapter 3 (see e.g., Pitt et al., (2019) who also used a literature review to inform the technological development of their novel computer-based aphasia treatment, in this case an online group intervention). Thus, for example, the literature review indicated that aphasia treatment to be self-delivered by computer should be compatible with different platforms (Swales et al., 2016) and so this was monitored in the study (type of device on which the SPT was used).

4.4 The feasibility of selecting a set of PR Verbs.

Another aspect of feasibility investigated in relation to RQ1, was the feasibility of *selecting* a set of *PR* verbs for treatment (RQ1c). This was investigated because a review of the literature relating to the selection of PR words in aphasia treatment (reported in the first half of Chapter 3) revealed that this is not a straightforward process (e.g., Renvall et al., 2013a & b) and that the process of selecting PR verbs was not reported in replicable detail. Despite the potential difficulty, making computer-based aphasia treatment personally relevant to PwA has been highlighted by both clinicians and PwA as important (e.g., Swales et al., 2016; Palmer et al., 2013). The process by which PR verbs were chosen will now be described in detail.

4.4.i The process by which PR verbs were chosen.

The SPT targeted 20 verbs (out of a set of 40) which were identified by the participant as being important and/or useful to them - “personally relevant” (PR) verbs. The process of selecting these PR verbs drew upon an eclectic range of research, used a variety of techniques and is described in detail in Chapter 3.2.ii. For convenience, it is also summarised here. The process was informed by the results of the CETI (Lomas et al., 1989) already carried out with participants i.e., verbs relating to *questions rated as difficult* (e.g., talking about my emotions) were discussed as potential treatment targets. Participants (and their consenting SO if the participant wished the SO to be involved) were also asked to identify *people, topics,* and *situations* when the participant found it hard to communicate and this process was supported by prompt cards from the interview section (Part B) of the CAPPA (Whitworth, Perkins & Lesser, 1997) (see Appendix A). Additionally, the CAPPA prompts were cross-referenced with the findings of a study carried out by Davidson et al. (2003). Davidson et al. observed neurologically healthy older people and PwA communicating in the community. The communication activities carried out by each group and the topics that they discussed were used to inform the prompt lists. The prompt list was also compared to the topics most commonly chosen by 100 PwA who each selected 100 PR words in the Big CACTUS study (Palmer et al., 2017) to ensure that these topics were covered. Finally, participants were encouraged to consider not only what they *had* to communicate, but also about what they *wanted* or *liked* to communicate.

Once 40 personally relevant (PR) verbs had been identified, images were chosen to represent each verb. Participants were encouraged to provide personal photographs to represent their selected verbs, but if they were unable to do so, the research student selected pictures from Google Images for approval by participants. These pictures were copied into PowerPoint slides and presented for naming at assessments 1 and 2 (i.e., at the beginning and end of the pre-treatment baseline). This enabled baseline naming performance to be established to i) assess the impact of treatment on verb naming and ii) to allow treated and control sets of verbs to be matched on the basis of naming performance. The 40 PR verbs were divided into sets of treated and control verbs (n = 20) also matched (as far as possible) for frequency, familiarity, imageability, concreteness and number of syllables (based on data from the MRC Psycholinguistic Database (https://websites.psychology.uwa.edu.au/school/MRCDatabase/uwa_mrc.htm). That is, the MRC database was not used to select PR verbs but was used in the process of matching treated and control sets.

As well as the verbs themselves being personally relevant, throughout treatment consideration was given to the personal relevance of the *sentences* (using the PR verbs) to be practised during Phase 2 of the SPT, and to the narratives practised in generalisation exercises during Phase 3 (Phase 1 focussed on production of verbs in isolation). This data was captured for each participant partly by noting the context of verb use for participants during the process of choosing their PR verbs. It was also informed by discussion/interaction during weekly visits when stepping participants up to the next level, and collaborating with them about what their target sentences and script/narrative would be. So, for example, the verb “try” was chosen as a PR verb by P2 and her partner because she wanted to convey her determination to continue to improve her speech. During Phase 2 “try” was targeted as a declarative sentence (“I am trying to improve my speech”) and as a Yes/No questions (“Are you trying?”), whilst in Phase 3 the previously practised declarative sentence (I am trying to improve my speech) was included as part of a script exercise for a review appointment with P2’s Stroke Consultant (designed to increase her independence during these consultations – one of P2’s goals).

The process of selecting PR verbs also incorporated a Goal Action Planning (GAP) procedure (Scobbie et al, 2013; Scobbie et al., 2011). GAP has three stages – Goal Negotiation, Goal Setting and Action Planning. Goal Negotiation and Goal Setting stages were used to identify each participants’ goals, and these also informed the selection of PR verbs. A secondary aim of the GAP procedure was to help participants comply with the self-delivered amount of SPT requested (two hours per week) – an important component of the SPT. The Action Planning stage of GAP involves asking patients (i.e., participants) how confident they are that they can achieve their goals and includes the formulation of an action plan and a coping plan (to achieve goals), if this is felt to be necessary. The action plan involves asking the participant “What are you going to do? *(to achieve your goals)* When will you do it? How often will you do it?” (p.477, Sykes et al., 2011), and the coping plan involves asking “What might get in the way of carrying out the action plan? How can this be avoided?” (ibid). This process allowed the identification of potential barriers to participants complying with treatment and consequently the formulation of potential solutions.

4.5 Is the SPT acceptable to the participants with aphasia and their SOs?

The second aspect of feasibility investigated in this study was that of the acceptability of the SPT to participants (RQ2). Thus, on completion of the SPT, the researcher administered a questionnaire to participants (and separately to their SOs) which investigated their views of the self-delivered, computer-based treatment. The questionnaire was based on that used by Palmer et al. (2013) who also investigated their participants' views of self-delivered, computer-based treatment (the StepByStep programme). The questionnaire was composed of six questions with a 5-point rating scale from 0 to 4, 4 being the most positive rating. The questions were:

1. Did you have previous experience of using a computer?
2. How did you find doing treatment on a computer?
3. How much help did you need with the computer treatment?
4. Did it work?
5. Have you used the words practised on the computer in your daily life?
6. Would you use it again?

Brief post-treatment interviews based on the topic guide used by Palmer et al. (2013) were carried out with the participants with aphasia and with their SOs (if the latter had consented). Participants with aphasia were asked three open-ended questions:

- Have there been any benefits of computer treatment?
- Are there any disadvantages of computer treatment?
- What would influence your decision to do computer treatment?

The interviews with SOs were guided using the following topics:

Prior expectations of computer therapy:

- Ease/acceptability of using computer
- Benefits of computer therapy for the participant

- Disadvantages of computer therapy for the participant
- Limitations of computer therapy
- Perceived change in talking
- Changes in daily activity
- Benefits of computer therapy for SO

4.6 What factors influenced compliance with the self-delivered component of treatment (RQ3)?

Ideally, the amount of SPT self-delivered by participants would have been automatically recorded by the SPT programme itself (like, for example, the StepByStep programme e.g., Palmer et al., 2012 & 2019). However, the SPT is a prototype that utilised Microsoft PowerPoint and thus does not have this capacity. Therefore, the amount of SPT self-delivered by participants was recorded manually, by participants completing an aphasia-friendly diary of the amount of practice they had done. Diary sheets were provided to participants at the end of each session and recorded a) the exercises completed, b) the length of time spent on each exercise, and c) how often each exercise was completed (see Figure 1 below). Although not ideal because self-reported data may not be completely accurate, two of the four studies of self-delivered computer-based verb treatment reported in Hickin et al. (2020) also captured data in this way (Kurland et al., 2014; Routhier et al., 2016). Both studies note the limitation that the self-reported data may not be 100% accurate.

Computer Exercise Diary

Participant Code **Treatment Week:**

Date	Exercise Name	Exercise Name	Practiced for:
Example	Ex 2 Light Verbs	Ex 3 Statements & Requests	
Monday 29th February	✓		20 minutes
		✓	30 minutes
Tuesday 16th October			minutes
			minutes
			minutes
Wednesday 17th October			minutes
			minutes
			minutes
Thursday 18th October			minutes
			minutes
			minutes
Friday 19th October			minutes
			minutes
			minutes
Saturday 20th October			minutes
			minutes
			minutes
Sunday 21st October			minutes
			minutes
			minutes
Monday 22nd October			minutes
			minutes
			minutes

Figure 1 Computer Exercise Diary

Diary sheets were collected during each face-to-face session, and if the participant had been unable to complete the target amount of practice this was discussed. Individual data was charted to explore the relationship between a number of variables and compliance with treatment. These were: i) the ability to use a computer (as measured by a bespoke Dynamic Assessment of Ability to use the SPT (based on Caute, 2017 personal communication)) ii) the severity of the overall profile of aphasia (as measured by the CAT) and iii) the presence/absence of a significant other (SO).

4.7. Were treatment procedures carried out during the SPT administered with acceptable fidelity (RQ4)?

The fidelity of *treatment delivery* relates to how far treatment was delivered *as intended* (e.g., Brogan et al., 2019; Conlon et al., 2020). Before describing how the fidelity of the delivery of the SPT was assessed, the development of the novel intervention will be described. This is particularly important for complex interventions such as the SPT: the MRC Guidance for the Development of

Complex Interventions (2019) recommends that they are informed by systematic reviews to ensure that they have a coherent rationale. The following section starts with a brief summary of the SPT programme, followed by a description of how first, the narrative review of computer-based aphasia treatments (reported in Chapter 3) informed the *design of the human computer interface (HCI)* of the computer-based programme, and second, the systematic reviews of verb and sentence treatments (reported in Chapters 1 and 2 respectively) informed the *content* of the SPT treatment exercises.

4.7.i Summary of the SPT programme.

The SPT comprised of a low-dose clinician delivered programme supplemented with self-managed computer-based treatment. Clinician delivered treatment took place in the university clinic if participants were able to travel there or in participants' homes if they were not. All clinician-delivered treatment was provided by the researcher who is a qualified speech and language therapist with over 30 years' experience of working with PwA. Clinician delivered sessions took place on a weekly basis for 8 weeks and lasted for a minimum of 45 minutes and a maximum of one hour 15 minutes depending on the researcher's clinical judgement as to the maximum amount of treatment a participant could benefit from during a session, or if the participant indicated that they were tired and wished to finish the session. The 45 – 75 minute treatment sessions were comprised entirely of treatment exercises i.e. activities such as greeting the participant and setting up the computer took place *before* treatment time was measured. Participants were additionally requested to carry out at least 2 hours self-directed homework each week (between face-to-face sessions) using the treatment exercises which had been introduced in that week's face-to-face session, during which treatment exercises were uploaded onto the participant's computer. Treatment was delivered using PowerPoint slides with the aim of increasing accessibility for PwA alongside several other features aimed to ensure that the HCI of the SPT programme was as aphasia-friendly as possible. These aphasia-friendly features will now be described.

4.7.ii The aphasia-friendly design of the HCI of the SPT programme.

The aphasia-friendly design of the HCI of the SPT programme was informed by a review of the small literature pertaining to this (Chapter 3). The few studies that investigated the acceptability of computer-based aphasia treatment were also reviewed and these too informed the design of the HCI.

First, PowerPoint was used to deliver the SPT slides because it is generally already uploaded onto computers at the point of purchase and does not therefore need to be purchased. This removed potential financial barriers to PwA accessing the SPT (e.g., Menger, Morris & Salis, 2016). Using PowerPoint slides to deliver treatment also allowed the slides to be designed according to aphasia friendly principles (Brandenberg et al., 2013; Language Light Interaction Design Guidelines for Aphasia available at http://languagelightux.org). The aim here was to enable easy navigation of the HCI for PwA so that they could independently deliver their own treatment. Thus, the slides incorporated all six of the aphasia-friendly features identified by Brandenberg et al. in the following ways:

- the written instructions on the slides used an aphasia-friendly *sans serif* font and were *large* (at least 14 point (Rose et al., 2012; 2011)
- the slides are *multimodal* in that they provide *visual* cues in the form of (PR) photos to represent target verbs and sentences, plus *written* cues, and *auditory* cues (accessed by clicking on a speaker icon) for each target verb or sentence.
- the only *button* to press is a speaker icon to access the audio cue, and this is *large* to make it easy to see and click on.

- the *interface is stable* in that the layout/content of each slide within an exercise is predictable (i.e., the photo of the target verb/sentence is in same place, as are the audio and written cues). Cues are also consistent between exercises and treatment phases (e.g., a written cue for a verb is always green, for an agent it is red and for a theme, black). Additionally, when animation is used in an exercise, the animation is consistent (e.g., the sequence in which the written cues for a sentence are revealed is always the same: verb first, agent second and theme third).
- the slides are *simple to navigate* in that there is no hierarchical menu structure to negotiate: a participant only needs to open the slides in PowerPoint in presentation mode (i.e., in Slide Show). (Simple navigation is also helped by the fact that many PwA are at least a little familiar with PowerPoint which was one key factor in deciding to use PowerPoint for the SPT – see next section).
- The slides were kept as *visually simple* as possible by always using a white background and keeping as much white space as possible.

Figure 2 shows four examples of slides taken from the SPT programme to illustrate these aphasia-friendly HCI principles: Phase 1 Verb Treatment (slide a), Phase 2 Sentence Treatment (slides b and c) and Phase 3 Generalization (slide d). (NB The voice bubble icons are included to illustrate the audio cues available from the speaker icons and are not on the actual exercise slides).

Of the additional principles recommended by Wilson et al. (n.d) *minimising interaction* is incorporated into the SPT in that participants could click anywhere on a treatment slide to reveal cues and to move from one slide to the next. (Only audio cues required the participant to click on the speaker icon itself, and the icon was made as big as possible to assist with this, whilst also trying to keep the icon small enough so that slides were not too busy thus violating the principle of visual simplicity). Users could also *control the pace* at which they went through the SPT slides (e.g., they could access cues in a slide as many times as they liked before moving onto the next). *Feedback* could not be included into the slides because their aim was to improve spoken production and the slides did not have speech recognition capability. The principles of *showing what is happening* and *providing prompts* relate to helping a PwA navigate a menu and since the SPT slides did not have a menu these were not relevant.

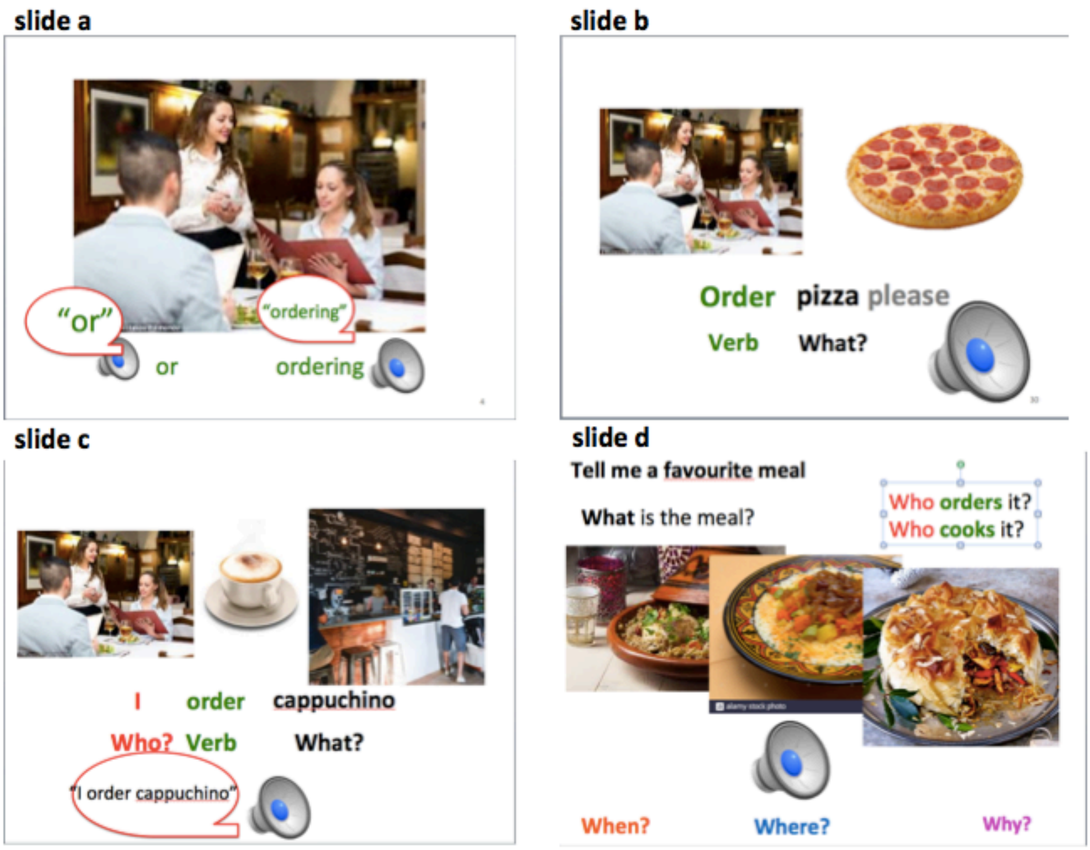


Figure 2 Examples of exercise slides from the SPT illustrating aphasia-friendly principles of HCI design.

To further maximise the accessibility of the SPT to the participants, their ability to use the SPT was assessed during the pre-treatment baseline phase to identify if any cognitive, sensory, or physical factors affected this. Thus, their ability to open PowerPoint and to open and navigate an exercise slide was assessed simultaneously with their ability to use cues to assist with these difficulties (e.g., verbal instructions, hand over hand demonstration). Any difficulties were noted, and training given prior to starting the SPT programme during baseline assessment sessions. If any difficulties were experienced during treatment itself, participants could phone the research student for advice. The research student also designed aphasia-friendly written instructions for participants to keep at home if required. These were tailored to each participant's needs as recommended by Kelly, Kennedy, Britton, McGuire and Law (2016) (e.g., using pictures of the screen on their own computer).

The views of participants and SOs on the acceptability of self-delivered computer-based treatment (Kearns et al., 2019; e.g., Palmer et al., 2013; Wade et al., 2003) were taken into account by:

- a weekly face-to-face visit by the research student during the eight weeks of self-managed treatment to i) update exercises, ii) address any technical difficulties identified by a participant or SO and iii) to discuss practice and progress (because support was identified as important to acceptability).
- the targeting of personally relevant verb and sentences during treatment (to maximize the interest level of treatment and thus motivation).
- the use of the GAP approach (Scobbie et al., 2013; Scobbie et al., 2011) to discuss how confident a participant was in being able to self-manage their treatment, any barriers to this and how these might be solved and to ensure that the SPT aligned with participants' goals (e.g., Marshall et al. 2013).
- clear advice to the participant each week that they should discontinue practice if they felt tired, frustrated, or anxious and that the research student could be contacted by phone to discuss this at any time to minimise the likelihood of any adverse effects of the SPT (Palmer et al., 2013; 2019).

The design of the SPT is also grounded in *clinicians'* perspectives of the ideal characteristics of computer-based aphasia treatment (Swales et al., 2016):

- All the *therapy activities* in the SPT targeted *expressive language*. They also included *functional* activities which simulate for example ordering in a cafe and using the phone.
- The *stimuli* In the SPT are *personalized* in that personal photos can be uploaded[14]. The SPT also targets *verbs* as desired by clinicians.
- The c*ues* in the SPT are *multimodal* and they can be *individualized* in that the PwA using the SPT can decide whether to use them, in what order and how many times to use them.
- *Access*. It is anticipated that the SPT would be *easy for clinicians to navigate* as it uses PowerPoint with which they are extremely likely to be familiar. The HCI of the SPT is *aphasia-friendly* as discussed above. The SPT can be used on *different devices* - either Android or Apple devices including PCs, laptops, tablets, iPads, and mobiles.
- The SPT is *financially accessible* as PowerPoint is usually already present on devices.

The only aspect that could not be addressed in the SPT is that of storing *progress data*. This would have required the development of new software and unfortunately funding was not available to pay for this.

[14] It was decided not to include PR video stimuli because this would add to the complexity of interacting with the SPT thus violating an aphasia friendly principle of an HCI. Experience of piloting the SPT in a student led university clinic (Hickin, Caute, and Woolf, 2012) had also demonstrated that including videos within the PowerPoint slides could sometimes lead to technical difficulties. Given that technical difficulties were cited as a disadvantage of computer-based treatment by PwA (Wade et al., 2003; Palmer et al., 2013) this was further grounds for not including video stimuli in the SPT).

4.7.iii The development of the SPT treatment exercises.

The SPT comprised of three phases. Phase 1 contained Verb Treatment Exercises, Phase 2 Sentence Treatment Exercises, and Phase 3 Generalization Treatment Exercises (i.e., to facilitate use of practised verbs and sentences in real life communication). All three phases of treatment drew from the evidence base for both verb and sentence production treatments, as identified by the results of the systematic scoping reviews reported in Chapters 1 and 2. The SPT was designed to be delivered over eight weeks which was motivated by investigating a regime of treatment which could conceivably be delivered within the constraints of NHS service provision. The SPT Manual contains a detailed description of the treatment protocol for all the SPT exercise and is provided as supplemental material.

The three phases of treatment and the exercises in each phase are shown in Figure 3 below.

1. VERB PHASE TREATMENT
Exercise 1. Sound and Letter Cueing Exercises.
Exercise 2. Sentence Closure Exercises.
Exercise 3. Verb Networks A Exercises.
Exercise 4. Verb Generation Exercises.
2. SENTENCE PHASE TREATMENT
Exercise 1. Simple Statements, Requests And Yes/No Questions
Exercise 2. Light Verb Exercises
Exercise 3. Verb Networks B Exercises
Exercise 4. Longer Statements and Wh- Questions.
3. GENERALISATION PHASE TREATMENT
Exercise 1. Personal and News Stories
Exercise 2. Scripts
Exercise 3. Story and Video Retell

Figure 3 Treatment Phases and Exercises.

The SPT programme will now be summarised starting with the *general principles* which guided treatment and were implemented at all stages. These principles were:

1. The PowerPoint slides for each treatment exercise were introduced by the research student

at the start of each face-to-face session. She then demonstrated how to use the slides for each exercise. Next participants were encouraged to *independently navigate* the slides during all face-to-face sessions. This was to both assess and facilitate participants' ability to navigate the slides independently as this would be required for self-delivery of treatment. Any assistance needed such as verbal prompts or guiding the participant's hand was given and noted on the appropriate Treatment Record Sheet.

2. *Gesturing* of a verb was encouraged when it was difficult to retrieve verbally. This was based on i) the evidence that gesture and associated verbs are strongly connected in terms of their neural representation and that therefore producing a gesture will facilitate retrieval of the corresponding verb (e.g., Marangolo et al., 2012) and ii) the evidence suggesting the effectiveness of gestural treatment for verb retrieval deficits (see Rose (2013) for a review). Any assistance needed to produce a gesture was again noted on the Treatment Record Sheet.

3. *Feedback* was always given regarding participants' responses during treatment whether these were correct or incorrect, and when incorrect an explanation of the nature of the error was given. Such feedback may be an *active ingredient* of treatment as it may facilitate learning for people with aphasia (e.g., Edmonds et al., 2009; Nickels et al., 1991).

4. *Colour coding*. The use of colour coding is reported very frequently in the sentence production treatment research literature (see Chapter 2) and was used throughout the SPT programme. It was introduced gradually in a "drip feed" manner, with the colour coding for verbs (green) introduced first, and then red for a verb's agent, followed by black for its theme and so on.

5. Participants were encouraged to practise *saying sentences at least as much* as they *listened* to spoken cues. This is based on the suggestion that when PwA use computer delivered treatment for sentence production, they learn most effectively by spending more time *saying*

sentences than *listening* to the spoken cues for these sentences (Cherney, 2012).

6. When participants produced a verb or sentence with *incorrect tense or morphology* this was *not corrected* as this was not the aim of treatment.

All the SPT exercises used the following *types of cue* as these were widely reported in the literature reviews (Chapters 1 and 2) and thus they very likely constitute the "active ingredients" of treatment:

- *Written cues* were used almost without exception in the reviewed studies. Although not explicitly stated, the rationale for this is presumably that written cues will be an effective support for spoken sentence production for participants whose ability to read aloud is reasonably intact. Written cues may also be used because they offer a permanent cue as opposed to the transient nature of spoken cues. Examples of the use of written cues in sentence production treatment include Baastianse, Hurkmans, & Links (2006), Links, Hurkmans & Bastiaanse, (2010), Webster, Morris, & Franklin (2005) and Webster & Gordon (2009).
- The use of *Wh- questions* to highlight (or teach) the thematic roles played by syntactic constituents of sentences was reported by Jones (1986) in mapping treatment used to successfully treat her client in an early clinical single case study. Her approach was emulated in a number of subsequent studies (e.g., Le Dorze et al., 1991; Takizawa et al., 2015).
- The use of *colour cues* to scaffold sentence production (and comprehension) was also widely reported in the reviewed studies. Colour cueing in the form of specific colours for particular word classes (e.g., green for verbs, red for nouns) is reported by, for example, Byng, Nickels and Black, (1994) and Newton et al. (2017). Colours have also been used to distinguish the thematic roles played by syntactic constituents in sentences. For example, Marshall, Chiat and Pring (1997) used colours to distinguish the thematic roles of source and goal in their

treatment which targeted production of sentences containing change of possession verbs. Byng et al. (1994) and Carragher et al. (2015) used a colour coded sentence frame to scaffold production of SVO(A) sentences. Schwartz et al. (1994) used colour coded wh- questions (see below) to highlight the roles of the verb, actor and theme. A hierarchy of difficulty with regard to the use of colour coding was introduced in some studies by gradually withdrawing the colour cues (e.g., Byng et al., 1994; Marshall et al., 1997).

- Many reviewed studies *combined* cueing techniques. For example, Nickels et al. (1991) used colour coded sentence constituents, a colour coded sentence frame and wh- questions in their treatment. In their implementations of PAS treatment, Whitworth Webster and Howard (2015) combined written cues, wh- questions and a sentence frame, and Biran and Fisher (2015) used shape cues, a sentence frame and wh- questions. The SPT is similarly eclectic in that a variety of cueing strategies are used. This was both based on the evidence but also gave participants the opportunity to use cues which worked best for them (and to ignore those that did not work) in their self-delivered treatment with the aim of facilitating effective and efficient learning (see Ball et al., 2018). The *exercises* comprising each phase of treatment will now be summarised with additional detail provided in the Treatment Manual.

Phase 1 of the SPT - **Verb Treatment** - aimed to improve verb production in isolation and comprised four exercises: 1) Sound and Letter Cueing; 2) Sentence Closure; 3) Verb Networks A; and 4) Verb Generation. Exercise 1 Sound and Letter Cueing was based on the evidence from anomia treatment and that demonstrating the effectiveness of these cues in improving verb retrieval (e.g., Conroy et al., 2009a, b and c; Conroy & Scowcroft, 2012; Raymer et al., 2007; Rose & Sussmilch, 2008). Exercise 2 Sentence Closure was based on the hypothesis that provision of a sentence supports verb production by priming the syntactic information stored as part of the verb's lemma, and was based on the evidence from the following studies (Conroy et al 2009a; Edwards and Tucker 2006; McCann & Doleman, 2011; Wambaugh et al., 2004). Exercise 3 Verb Networks A was based on Semantic Feature Analysis as adapted for the treatment of verb retrieval (Carragher et al., 2013; Knoph, Simonsen & Lind, 2017; Knoph, Lind & Simonsen, 2015; Rose & Sussmilch, 2008; Wambaugh et al., 2014; Wambaugh & Ferguson, 2007). Finally Exercise 4 Verb Generation was designed to facilitate transition away from production of a single verb in response to a picture stimulus. Thus, it required production of a verb in response to either i) a picture of an *object* with the prompt "What can you do with this? (e.g., a book to prompt production of the verbs "read" and "write") or ii) a *scenario* presented as an audio cue (e.g., "You have invited friends over for dinner- what will you do?" to prompt the production of the verbs, "clean, cook, eat, drink, talk"). This exercise was based on that used by Marshall et al. (1998).

Phase 2 - **Sentence Treatment** - aimed to improve verb production in sentences and comprised four exercises: 1) Simple statements, Requests and Yes/No Questions; 2) Light Verb Exercises; 3) Longer Statements and Wh- Questions; and 4) Verb Networks B Exercises.

Exercise 1) Simple statements, Requests and Yes/No Questions represents the intermediate stage of a syntactic hierarchy used in the SPT: Phase 1 Verb Treatment targeted production of *one* syntactic element (verbs in isolation), and Phase 2 Exercise 1 targets production of *two* elements in Requests (VO) and Yes/No Questions (SV with S V inversion), and then *three* elements in Simple (declarative) Statements (SVO). The rationale for the syntactic hierarchy is based on a number of studies which have shown that a syntactic hierarchy of increasing sentence length is effective in improving the production of declarative sentences of SV(O) structure (e.g., Bazzini et al., 2015; Biran & Fisher, 2015; Carragher, Sage & Conroy, 2015; Jones, 1986; Le Dorze, Jacob & Coderre, 1991; Newton, Finch & Bruce, 2017; Nickels, Byng & Black, 1991; Springer, Huber, Schlenck & Schlenck, 2000).

Exercise 2 targets the production of simple statements and requests (imperatives) including using five light verbs (do, be, go, make and have). The rationale for the inclusion of light verbs in the SPT programme arises from the frequency of light verbs in daily discourse. For example, Cruice, Pritchard and Dipper (2014) found that nearly half of the verbs used in discourse were light. Conrad and Biber (undated) report that twelve verbs account for 45% of the verbs produced in conversation, seven of which are light. Carragher et al. (2013), Goral and Kempler (2009) and Jones (1986) also all treated light verbs as part of treatment aimed more generally at increasing heavy verb and sentence production. It was hoped that because light verbs are so frequent, their inclusion in the SPT program would facilitate generalisation of treatment to real life communication, and in particular because light verb requests are so useful (e.g., make a cup of tea please, go to the shops, do the washing up).

Exercise 3 targets the production of Longer Statements (SVOA) and Wh- Questions. The production of four constituent (longer) sentences represents the final stage in the syntactic hierarchy of increasing sentence length used throughout the treatment programme, which, as noted for Exercise 1, has been shown to be effective in a number of studies, whilst Wh- questions are targeted in stage 3 of the HELPSS (e.g., Helm-Estabrooks & Ramsberger, 1986).

Exercise 4 Verb Networks B is based on Verb Network Strengthening Treatment (VNeST) developed by Edmonds and colleagues in a series of studies (Edmonds Obermeyer & Kernan, 2015; Edmonds, Mammino & Ojeda, 2014; Edmonds & Babb, 2011; Edmonds, Nadeau & Kiran, 2009). In particular these exercises are adapted from VNeST-C, a computerised version of VNeST which aimed to improve written (typed) verb naming (Furnas & Edmonds, 2014). VNeST requires the production of multiple exemplars of SVO structures using target verbs and encourages the production of personally relevant examples and therefore lent itself well to the SPT. Edmonds et al.

(2015) report improved production of verbs, sentences and/or discourse in all 11 participants who were treated with VNeST in the series of studies listed above.

Phase 3 of the SPT - **Generalisation Treatment** - aimed to help participants use the verbs and sentences treated in Phases 1 and 2 in real life communication and was composed of three exercises: Exercise 1) Personal and News Stories; Exercise 2) Scripts; and Exercise 3) Story and Video Retell. The exercises in this final stage of treatment included simulation of communication scenarios identified by participants as difficult during goal planning to facilitate the use of treated verbs and sentences in everyday life (e.g., P2 practised telling her stroke story in Phase 3 Exercise 1 to increase independence during consultations with health professionals. She then practised a script in relation to review appointments with her stroke consultant in Exercise 2 with the same aim). The overall rationale for the inclusion of generalisation exercises in the SPT arose from the literature review where a number of studies indicated that a generalisation phase of treatment may be necessary to enable PwA to transfer improved communication skills demonstrated in the clinical context, to real life communication (e.g., Byng, Nickels & Black, 1994; Newton, Finch & Bruce, 2017; Nickels, Byng & Black, 1991).

Exercise 1) Personal and News Stories aimed to facilitate transition from producing single sentences (Phase 2 Sentence Treatment) to producing narratives by using the same photographs as Verb and Sentence Treatment, but this time in picture sequences to facilitate the production of a narrative. Newton, Finch and Bruce (2017) state that exercises using single picture stimuli (as in their treatment) do not allow practice of narrative skills which they believe may be important in facilitating the transition away from picture description towards daily discourse. Nickels, Byng and Black (1991) and Byng, Nickels and Black (1994) also used picture sequences during a final generalisation stage of treatment and both reported improvements in narrative and real-life conversation. Telling personal stories and news/current affairs were chosen as topics in Exercise 1 as they are a frequent communicative activity for PwA and shared with non-aphasic individuals (Davidson et al., 2003). They have also been used in previous treatment studies with a positive impact on daily communication (Hoover et al. 2014; Jones, 1986; Marshall, 1998).

Exercise 2 (Scripts) used two different treatment protocols. Initial tasks were based on the treatment protocol used in the HELPSS (e.g., Helm-Estabrooks & Ramsberger, 1986; Helm-Estabrooks, Fitzpatrick, & Barresi, 1981). The HELPSS uses two levels of script to elicit target sentences, with the first level containing the targets whilst the second does not. Exercise 2 thus had a PowerPoint slide for each treatment level for target sentences. Subsequent tasks in Exercise 2 were based on the protocol used in Computerised Script Training for PwA developed by Cherney and colleagues (Cherney, 2012; Cherney Halper, Holland & Cole, 2008).

Exercise 3 (Story and Video Retell) aimed to improve the ability of participants to tell stories, important for daily communication (e.g., Davidson et al., 2003). This exercise is similar to exercise 1 (Personal and News Stories) but was hypothesised to be more challenging because it used *unfamiliar* material to stimulate story production i.e., picture sequences and video clips. It therefore seemed likely that the cognitive demands of Exercise 3 might be higher (i.e., because participants had to first work out what the story of the video was and then how they were going to tell it (see Carragher et al., 2015). Initial tasks in Exercise 3 used cartoon-like picture sequences (sourced from Teaching English as a Foreign Language material) to support narrative production (Hickin, Mehta & Dipper, 2015), and which has been proven effective (Byng et al., 1994; Goral & Kempler, 2009; Maul et al., 2014; Nickels et al., 1991). Subsequent tasks in Exercise 3 used Video stimuli which included short Pixar films (e.g., For The Birds https://www.youtube.com/watch?v=pWIVoW9jAOs) as well as clips of cooking, gardening and DIY TV programmes when these were of interest to participants. Video narrative tasks have been used previously to facilitate generalisation (Carragher et al., 2015; Goral & Kempler, 2009; Schwartz et al.

1994); and there is emerging evidence which tentatively suggests that the use of (dynamic) video cues may be more effective than (static) picture cues in treating verb retrieval deficits in aphasia (Blankestijn-Wilmsen, 2017).

4.7.iv Assessment of the fidelity of the SPT.

The treatment manual summarised above (and provided as supplemental material) was used to develop a *fidelity checklist* to assess whether intervention was delivered as intended in the manual. A checklist was developed for each of the 11 SPT exercises, and Appendix B provides an example of a blank checklist here for Phase 1 Verb Treatment, Exercise 2: Sentence Closure. The process of developing the checklists was informed by the emerging literature in the field. The checklists constitute indirect assessment of treatment fidelity (as defined by Kaderavek and Justice (2010)) because they are self-report of compliance with intended behaviours. The development of the checklists was influenced by the fidelity checklist contained in Appendix A of Kaderavek and Justice's paper (ibid). Also influential was the framework proposed by Baker (2012a & b) to determine the active ingredients of treatment. Thus, behaviours monitored by the checklist were divided into clinician behaviours (or "therapeutic interventions" as defined by Baker) and participant behaviours (or "client acts"). This approach was also adopted in the fidelity checklists developed by Behn and colleagues in monitoring the fidelity of a peer befriending intervention for people with aphasia (SUPERB e.g., Behn, Moss, McVicker, Roper, Northcott, Marshall, . . . Hilari, K., 2021). Behaviours were rated as either present or absent - an approach taken in the Kaderavek and Justice checklist and recommended by Behn (personal communication). Finally, the treatment protocol for VNeST reported by Edmonds (2014) was also instructive in considering behaviours which needed to be monitored in the SPT.

The SPT checklist was completed by the research student at the end of every session conducted with each participant, with the fidelity data comprising counts of present/absent behviours. The research student also audio recorded every face-to-face treatment session and for 25% of sessions the audio recording was reviewed by the research student to check adherence to the treatment manual and hence the accuracy of the fidelity rating.

The methodology used to answer the final research question relating to preliminary efficacy testing will now be described.

4.8 Preliminary efficacy testing of the effect of the SPT on the production of i) trained and untrained verbs, ii) untrained nouns, iii) sentence production using trained and untrained verbs, iv) verb and sentence production in discourse and v) in functional communication as perceived by i) the participants with aphasia themselves and ii) their significant others?

To carry out preliminary efficacy testing of the SPT, several measures of language production were taken pre-treatment at (T1 and T2) and post=treatment (at T3 and T4). A more detailed description of each assessment is given in section 4.7.i including number of items, scoring procedure, psychometric data, and rationale for use in this study.

i) Improvement in the production of both treated and untreated PR verbs i*n isolation* was assessed by asking participants to describe what was happening in a prompt picture using **one word**. Improvement in *untreated* personally *irrelevant* verbs was assessed using a subset of verbs (n=43) taken from the Object Action Naming Battery (OANB)[15] (Druks &

[15] P1 and P2 were the first participants in the research study and named 50 verbs and 50 nouns from the OANB which were not matched. Matched sets were used for subsequent participants following advice given at the research student's upgrade assessment.

Masterson, 2000) also used by Edmonds, Obermeyer & Kernan, (2015) to assess the impact of VNeST[16]. Responses made within five seconds were credited as correct, mirroring the protocol used in the CAT naming tests (Swinburn et al., 2004). The significance of the effect of treatment was calculated using a McNemar Test. The stability of pre-treatment assessments (time points 1 and 2: T1 and T2) was also checked using the Wilcoxon Two Sample Test. Where no significant difference was found, T2 was used as the point of comparison for post treatment scores (T3 and T4). When a significance difference was found, whichever of the two pre-treatment assessments was higher (T1 or T2) was used as the point of comparison for T3 and T4, to minimise over-inflating any change effect.

ii) Improvement in untreated nouns was assessed using nouns from the OANB subset (Edmonds et al., 2015) and these matched the untreated personally irrelevant verbs assessed in i) above (allowing assessment of the impact of treatment on noun retrieval and a comparison of a participant's ability to retrieve nouns and verbs). Edmonds and colleagues report improvement in noun retrieval as a result of Verb Network Strengthening Treatment (VNeST) and assert that improvements in lexical retrieval are at the heart of improvement in sentence production as a result of VNeST (ibid). It was therefore deemed important to measure noun as well as verb retrieval)[17.]

[16]Verbs in the OANB were checked against participants' PR verbs and any common (i.e., treated) items were removed from the analysis.

.

[17] Nouns in the OANB were checked against participants' PR verbs and any nouns likely to have been treated (e.g., as a theme or instrument of a PR verb in a practised sentence) were removed from the analysis.

iii) Sentence production using treated verbs was assessed by testing (picture) naming of the treated set of PR verbs (n=20). (Participants were asked to describe what was happening in the picture using **a sentence**)[18]. Sentence production using untreated verbs was assessed by testing a) (picture) naming of the untreated set of personally relevant verbs (n=20) following the procedure described above and b) sentence production using untreated (personally *irrelevant*) verbs was assessed using the Sentence Construction Test of the Verb and Sentence Test (VAST: Bastiaanse, Edwards & Rispens, 2002). The number of verbs, agents, objects, and complements produced in each sentence were counted separately to measure whether retrieval of not only verbs had improved post treatment, but also whether there had been any impact on production of the arguments associated with each verb, both obligatory and non-obligatory (no distinction was made between obligatory and non-obligatory arguments following the method of Webster, Franklin & Howard (2007)). To be credited as correct, agents, verbs and objects had to be produced within 10 seconds of picture stimulus presentation. The significance of any change was assessed using a Wilcoxon Two Sample Test. Change in the production of i) verbs alone and ii) of a verb with its arguments (1, 2 and 3) was analyzed as these were the structures targeted in the Verb Phase and the Sentence Phase of the SCT respectively, and the significance of any change was assessed using a χ^2 Test. (Predicate argument structure (PAS) was analysed using the method adopted by Webster et al. (ibid)).

iv) The CAT Picture Description Task was used to assess discourse. It was scored using the procedure described in the CAT Manual (Swinburn et al., 2004) to attain an overall score.

[18] To avoid priming effects, assessment of sentence production using personally relevant verbs took place in a different session to that of verb production in isolation (RQ 3i)

The overall score is calculated using the number of Appropriate Information Carrying Words (ICWs) produced, and the number of Inappropriate ICWs produced is then subtracted from this to create a combined score. The picture description is then rated on a six-point scale in terms of a) syntactic variety, b) grammatical well-formedness and c) speed of production. Each of these three scores is added to the combined ICW score to calculate the overall score. The significance of any change in the overall score was assessed via Friedman's Test. In addition, production of verbs (types and tokens) was counted. PAS was assessed via the same method used to analyse PAS in sentence production with the aim of facilitating comparison of performance between the two tasks.

v) Discourse in terms of narrative production was assessed via samples elicited using the Autobiographical Memory Interview (AMI: Kopelman & Baddeley, 1989). Samples were analysed using the same procedure as for sentences and picture description (from Webster et al., 2007). Additionally, samples were scored against three of the criteria used to score the CAT picture description (syntactic variety, grammatical well-formedness, and speed) to facilitate comparison between the two discourse genres.

vi) Improvement in functional communication was assessed using the Communicative Effectiveness Index (CETI: Lomas et al., 1989) which was given to participants and their SOs (if the latter had consented) at each assessment point. The CETI uses a rating scale and the significance of pre/post treatment changes in scores was therefore assessed using Freidman's Test. Change was also regarded as significant if it exceeded the retest standard error of the mean of 5.2 established during CETI development (Lomas et al., 1989) following the methods used by Edmonds et al. (2014) who used the CETI to assess the impact of VNeST on functional communication.

A summary of the outcome measures used to assess each of the research questions is given in Table 5 below.

Table 5 Research Questions, outcome measures (OMs) and timing of OMs.

Research Question	Outcome Measure	Administration of OM
1. **Is it feasible to self-deliver SPT by computer?**	ease of participant recruitment; participant attrition; log of technical difficulties; self-reported compliance with treatment; logging of selection of PR verbs	Throughout study
2. **Is the SPT acceptable *to* the participants with aphasia and their SOs**.	TA of interview with participants and their SOs	After treatment completed for each participant and their SO
3. **What factors influence compliance with self-delivered computer treatment?**	exploration of any potential relationships using descriptive statistics	After treatment completed for each participant
4. **Were treatment procedures carried out during the SPT administered with acceptable fidelity?**	self-rating by research student of face-to-face treatment sessions using fidelity checklist	During and after treatment phase
5. **Is the SPT efficacious i.e., does it result in statistically significant improvement in the production of:** **i) treated and untreated verbs**	picture naming score (using single word) for PR treated and untreated verbs; verb subtest of the OANB	Ax1, 2, 3 and 4
ii) untreated nouns	noun subtest of the OANB	Ax1, 2, 3 and 4
iii) sentence production using	Picture naming score (using a sentence) of PR treated and untreated verbs;	Ax1, 2, 3 and 4

treated and untreated verbs	Sentence Construction Test of the VAST	
iv) verb and sentence production in discourse	Analysis of the AMI: verb production and percentage of V plus 1/2/3 argument production; syntactic variety, grammatical well-formedness and speed	Ax1, 2, 3 and 4
v) *functional* communication as perceived by i) the participants with aphasia themselves and ii) their SOs?	Participant CETI score; SO CETI score	Ax1, 2, 3 and 4

4.8.i Assessments used to investigate preliminary efficacy testing of the SPT.

The Comprehensive Aphasia Test (CAT) (Swinburn et al., 2004)

The Cognitive and Language Batteries of the CAT were administered to all participants to give “a quick but comprehensive profile of the language performance of the person with aphasia” (CAT Manual p.5). Whilst the CAT does not allow type of aphasia to be established, it enables deductions to be made regarding the psycholinguistic deficits underlying each participant’s verb and sentence production difficulties, as well as providing a measure of the severity of their aphasia. (The CAT is standardised based on data from 113 PwA and 27 non-aphasic subjects which enables T Scores to be calculated for each subtest). The relationship between each participant’s CAT profile and their response to treatment was explored to identify any potential prognostic factors or contraindications regarding the suitability of the SPT.

The Object and Action Naming Battery (OANB) (Druks & Masterson, 2000)

Confrontation naming of verbs and nouns was assessed using matched subsets taken from the OANB. The OANB was developed to facilitate exploration of the psycholinguistic differences between verbs and nouns and has 162 action pictures and 100 verb pictures matched for age-of acquisition, frequency and imageability (with 93% agreement on the rating of the aforementioned variables). The literature review established that the OANB was widely used in both verb and sentence treatment research (see Chapters 1 and 2), and its use in this research therefore facilitates comparison of the results of this study to that of the existing research. The matched subset used was developed by Edmonds et al., (2015) to investigate the impact of VNeST, and comprises 42 verbs and 42 nouns matched for age-of acquisition and imageability (as far as possible given the lower imageability of verbs compared to nouns). Verb and naming scores on the OANB will allow exploration of the relationship between each participant's degree of verb and noun impairment and their response to the SPT.

The Verb Comprehension Test (VCT) from the Verb and Sentence Test (VAST) (Bastiaanse, Edwards & Rispens, 2002)

The VAST is a linguistically based assessment battery which was developed to allow "a thorough investigation of verb and sentence comprehension and production, and give directions for therapy" (VAST Manual, p.4). The VAST contains ten subtests and is not intended for use in its entirety but rather individual subtests can be administered to investigate the aspects of verb and/or sentence processing of interest.

The VCT was selected to assess the integrity of participants' semantic representations of verbs and thus the likely contribution of any deficit to problems with verb retrieval. The VCT contains 40 verbs controlled for frequency, transitivity, and phonological relatedness to a noun. The test procedure requires a participant to select the correct drawing from a selection of four: i) the target verb, ii) a semantically related verb, iii) an object semantically related to the target verb or iv) to the semantic distractor verb. The VCT has good reliability established via testing on a sample of both aphasic and non-aphasic subjects, and reasonable validity (compared to the Token Test) ($r = -0.41$, $p < 0.10$).

The Sentence Construction Test (SCT) from the Verb and Sentence Test (VAST) (Bastiaanse, Edwards & Rispens, 2002)

The SCT was selected to assess participants' ability to produce sentences. It uses a picture description task which the authors regards as "one of the best ways to estimate the ability to make sentences in daily life . . . because one has a good idea of the target sentence, which is not always the case in spontaneous speech." (VAST Manual p.13). The SCT has 20 items controlled for frequency, transitivity and reversibility, and it allows assessment of the ability to retrieve lexical items and produce a grammatical sentence. The SCT has good reliability established via testing on a sample of both aphasic and non-aphasic subjects. Its validity (compared to the Token Test) was the lowest of the ten subtests of the VAST with a correlation of -0.36 ($p > 0.10$).

The Communicative Effectiveness Index (CETI) (Lomas et al., 1989)

The CETI was used to assess the impact of treatment on functional communication. The CETI is a questionnaire which asks SOs to rate a PwA's ability to communicate in sixteen situations, using a visual analogue scale. The sixteen situations were identified in consultation with PwA and their SOs giving the CETI good face validity (Lomas et al., 1989). The CETI also has good construct validity (by comparing CETI scores to relevant subtests of the WAB) and good test-retest and inter-rater reliability (see Lomas et al., for detailed discussion). Finally, although the functional impact of both verb and sentence treatments has rarely been assessed (see Chapters 1 and 2), the CETI has been used with the most participants in the existing research (19 participants across two studies) – an additional justification for its use as an OM in this study.

The Autobiographical Memory Interview (AMI) (Kopelman, Wilson & Baddeley, 1989)

The AMI was developed to assess the autobiographical memory and personal semantic memory of patients with amnesia due to a variety of aetiologies, including stroke. The AMI was chosen as an OM in this study not to assess participants' memory but because the procedure used to test autobiographical memory elicits a *personal narrativ*e using a structured interview procedure. Because personal verbs and sentences were targeted in the SPT, it was hypothesised that a personal narrative would be more sensitive to changes in discourse production than discourse samples more commonly used in the verb and sentence treatment literature (such as a Cinderella narrative). The AMI was also demonstrated to have good reliability in terms of scoring memory using a scale for the amount of information recalled in each personal narrative. Whilst this rating scale was not used in this study, this finding nevertheless suggests that the AMI is able to collect reliable data. The scoring procedures from the CAT Picture Description Task were applied to the samples of narrative obtained via the AMI to facilitate comparison of the two discourse genres.

Chapter 5. Results for research questions 1 - 4: the feasibility of the SPT.

The results for research questions 1 – 4, relating to the feasibility of the SPT, are reported in this chapter, and the results relating to the preliminary efficacy testing of the SPT (research question 5) are reported in the subsequent chapter (6). To identify indicative influential factors, feasibility results for the SPT are reported for the group. Preliminary efficacy results are reported as single cases only (in chapter 6).

The research questions relating to the feasibility of the SPT were:

1. Is it *feasible* to deliver and self-manage personalised sentence production treatment (SPT) by computer? Specifically, is it feasible to a) recruit and retain suitable participants to the SPT, b) deliver and self-manage the SPT using a computer, and c) is it feasible to select a set of PR verbs?
2. Is the SPT *acceptable* to the participants with aphasia and their significant others?
3. What factors influence *compliance* with self-managed computer treatment?
4. Were the treatment procedures that were carried out during the SPT administered with acceptable *fidelity*?

5.1 Feasibility of the SPT.

5.1.i Recruitment.

Regarding recruitment, twelve PwA were screened for their eligibility for the research project between the end of January 2018 and the beginning of April 2019. Eight individuals were eligible, of whom six consented to participate in the study. Participants were recruited in stages (to manage the research student's workload), with P1 and P2 recruited first (in January 2018), P3, P4 and P5 May – July 2018, and P6 at the beginning of April 2019.

The reasons for *not* recruiting six individuals who were screened are reported in Table 6 below, together with their recruitment source.

Table 6 Reasons screened participants did not participate in the research project.

P	Gender	Reason for non-participation	Recruited via:
7	F	Repetition <75% on CAT Repetition of Words	City Aphasia Lab Research Register
8	M	Verb naming <15% on OANB screening	Referred by wife
9	F	Verb naming >80% on OANB screening	City Aphasia Lab Research Register
10	F	Suitable but decided not to participate as secured a volunteer post with aim of resuming employment	Aphasia self-help group
11	M	Suitable but unable to participate due to a 4 month sailing holiday	Aphasia self-help group
12	F	Verb naming >80% on OANB screening	City Aphasia Lab Research Register

Two people fell below the inclusion criteria, one in terms of repetition (P7) and the other (P8) had a global aphasia and would potentially have fallen below many, if not all, of the inclusion criteria in terms of naming, repetition and ability to comprehend language sufficiently well to give his informed consent. (He was not subjected to screening tests as the research student recognized his aphasia was too severe). Two people (P9 and P12) were above the inclusion criteria in terms of verb naming. Two people (P10 and P11) were suitable for the research project but did not participate for personal reasons (see table above).

In terms of sources of recruitment, the most successful source was the City Aphasia Lab Research Register (i.e., participants who had completed previous research projects with the Language and Communication Science Research Centre and had given their consent to be contacted about future research projects that may be of interest). Five of the six participants in the SPT research project were recruited via this route with the remaining participant being referred by her sister who emailed the Division of Language and Communication Science. Of the six participants who were screened but did not participate, three had participated in previous research and three were recruited via contact with aphasia self-help groups. In terms of the geographical location of participants, three participants lived in Greater London, one in Hertfordshire, one in Middlesex and one in Northamptonshire. All six participants were retained for the duration of the project i.e., no participants dropped out. However, some assessment data could not be collected due to participant illness (the AMI at T3 for P1) and to participant fatigue (the SCT for P6 at T3). In terms of the assessment data from SOs, P1's partner completed all CETIs, P2's sister forgot to complete the CETI at T2 but completed all others, and P4's wife did not complete the CETI at T3 and then at T4 mislaid both her CETI and that of P4 when they moved house.

The demographic information and baseline assessment profile for each of the six recruited participants will now be reported.

Demographic information for each participant is given in Table 7. and the results of background assessment using the CAT (Swinburn et al., 2004) are given in Table 8 (Cognitive Battery), Table 9 (Language Comprehension) and Table 10 (Expression)

Table 7 Demographic data for participants.

	Gender	**Age**	**Previous Occupation**	**Years post onset**	**Aetiology**	**Aphasia Type**	**Physical Impairment**	**Consenting SO**	**Involved in previous research**	**Community involvement**
P1	M	53	Accountant	10	CVA (embolic)	Non-fluent + mild dysarthria	Dense R hemi-paresis	No	Yes	Attended stroke groups
P2	F	54	Primary School Teacher	3	CVA (embolic)	Non-fluent	None obvious	Yes	Yes	Attended stroke groups & volunteer
P3	F	49	Operating Theatre Nurse	3	CVA (haemorrhagic)	Non-fluent + mild apraxia	Dense R hemi-paresis	Yes	No	Attended 1 stroke group
P4	M	57	Ex-Serviceman	**7**	**CVA** (embolic)	Fluent	None obvious	Yes	No	Attended stroke group

			and wine merchant							& ex-servicemen group
P5	F	81	Office Manager	4	CVA (embolic)	Fluent/anomic	None obvious	No	Yes	Attended stroke groups
P6	M	59	Worked in The City	2	CVA (embolic)	Non-fluent	Dense R hemi-paresis	**No**	**No**	Attended stroke group

Table 8 Comprehensive Aphasia Test for participants (Cognition).

	P1		**P2**		**P3**		**P4**		**P5**		**P6**	
	SCORE	**T SCORE**	**SCORE**	**T SCORE**	**SCORE**	**T SCORE**	**SCORE**	**T SCORE**	**SCORE**	**T SCORE**	**SCORE**	**T SCORE**
Cognition												
Line Bisection	-1	53	1	53	-2	44	0	66	0	66	2	44
Semantic Memory	9	51	10	60	10	60	10	60	9	51	10	60
Word Fluency	7	51	15	59	13	57	12	56	7	51	0	37
Recognition Memory	10	59	10	59	10	59	10	59	9	48	10	59
Gesture Object Use	10	55	7	45	10	55	12	68	9	51	10	55
Arithmetic	3	49	5	57	5	57	5	57	5	57	4	53
Cognitive Total	**32**		**32**		**35**		**37**		**32**		**34**	

Table 5.3.ii Comprehensive Aphasia Test Results for Participants (Language Comprehension).

Table 9 Comprehensive Aphasia Test results for participants (Language Comprehension).

	P1		P2		P3		P4		P5		P6	
	SCORE	T SCORE	SCORE	T SCORE	SCORE	T SCORE	SCORE	T SCORE	SCORE	T SCORE	SCORE	T SCORE
Language Comprehension												
Comprehension of spoken words (n=30)	28	58	28	58	28	58	26	53	28	58	30	65
Comprehension of spoken sentences (n=32)	25	57	28	61	24	56	21	53	23	55	28	61
Comprehension of spoken paragraphs (n=4)	4	60	4	60	3	49	3.5	41.5	1.5	42	3.5	41.5
Spoken Comprehension Total	**57**	**57**	**60**	**60**	**55**	**55**	**50.5**	**52**	**52.5**	**53**	**61.5**	**61.5**
Comprehension of written words (n=30)	26	51	25	50	28	58	23	43.5	30	64	28	58
Comprehension of written sentences (n=32)	13	47	19	53	19	53	12	46	20	54	23	57
Written Comprehension Total	**39**	**48**	**44**	**51**	**47**	**53**	**35**	**46**	**50**	**56**	**51**	**57**

Table 10 Comprehensive Aphasia Test results for participants (Expressive Language).

	P1		P2		P3		P4		P5		P6	
	SCORE	T SCORE	SCORE	T SCORE	SCORE	T SCORE	SCORE	T SCORE	SCORE	T SCORE	SCORE	T SCORE
Expressive Language												
Repetition of words (n=32)	30	57	32	65	24	50	24	50	28	55	32	65
Repetition of complex words (n=6)	6	62	3	49	0	38	2	43	4	57	6	62
Repetition of nonwords (n=10)	10	67	7	55	3	48	2	46	9	64	0	38
Repetition of digit strings (n=14)	4	43	4	43	6	46	4	43	12	59	6	46
Repetition of sentences (n=12)	6	48	6	48	10	56	6	48	12	63	10	56
Repetition Total	**56**	**54**	**52**	**52**	**43**	**50**	**38**	**47**	**65**	**58**	**54**	**53**
Naming Objects (n=48)	35	55	43	61	42	61	16	47	48	74	18	48
Naming Actions (n=10)	3	49	4	50	10	69	2	47	6	54	6	54
Word Fluency	7	51	15	59	13	57	12	56	7	51	0	37
Naming Total	**45**	**53**	**62**	**59**	**65**	**60**	**30**	**49**	**61**	**58**	**24**	**48**

Spoken Picture Description	*18.5*	*52*	*24*	*56*	*20*	*53*	*18.5*	*52*	*27*	*58*	*7*	*48*
Reading words (n=48)	33	51	46	62	44	60	22	48	46	62	11	45
Reading complex words (n=6)	2	51	4	57	6	67	0	40	6	67	0	40
Reading function words (n=6)	6	62	6	62	6	62	4	49	6	62	0	35
Reading nonwords (n=6)	0	40	2	49	9	64	0	40	7	58	2	49
Reading Total	**41**	**50**	**58**	**57**	**65**	**65**	**26**	**47**	**65**	**63**	**13**	**45**
Writing copying (n=27)	25	50	26	52	27	61	26	52	27	61	27	61
Writing picture names (n=21)	17	56	18	58	19	60	7	47	19	60	21	67
Writing to dictation (n=28)	13	50	19	53	28	68	6	47	25	59	26	61
Writing Total	**55**	**52**	**63**	**55**	**74**	**64**	**39**	**48**	**71**	**60**	**74**	**66**
Written Picture Description	10	58	10	58	11	59	2.5	52	10	58	12	60

5.1.ii Participant 1.

Participant P1 was a 53 year old man who was 10 years post stroke at the start of the study. He had a dense right-sided hemi-paresis but was mobile with a walking aid and used his non-dominant left hand to write. P1 was a senior accountant at the time of his stroke but had to resign from his position because of his aphasia. He lived alone and was estranged from his family. He had a pet cat and participated in a number of community-based stroke groups. He had also participated in research projects prior to this one. P1 did not identify a SO or friend to participate in this research study. He did however talk quite frequently about a friend who he asked to help him on one occasion when he was having difficulty using his SPT exercises. P1 had an iPad which he used to email and browse the Internet and which he used to carry out his SPT exercises. P1 received his SPT from the beginning of June 2018 – mid August 2018.

P1 had a moderate non-fluent aphasia with an additional mild dysarthria which affected his intelligibility at times. P1 had good understanding of both spoken and written single nouns. Regarding comprehension of verbs, on baseline testing (T1 and T2) P1 scored an average of 35/40 (87.5%) (on the VCT), choosing either the semantically related verb distractor or the related noun suggesting a mild semantic impairment. He had difficulty understanding both spoken and written reversible sentences suggesting difficulty mapping thematic roles onto syntactic structures. In relation to written sentences, his low score reflected delayed responses in general, in addition to difficulty with reversible sentences. P1 had good repetition skills and reasonably preserved ability to read aloud words, but this was affected by frequency, imageability and length. He was unable to read aloud non-words suggesting significant damage to grapheme–phoneme conversion rules. P1's scores at baseline on key outcome measures are reported in Table 11 below. P1 showed stable performance with the exception of OANB Verbs and PR Verbs in sentences. In summary, in terms of lexical retrieval, P1 showed moderate impairment of both nouns and verbs, with verbs more impaired. He showed a small syntactic boot strapping effect when producing verbs in sentences, and in terms of sentence production, he most commonly produced verbs with one or two arguments. However, narrative production revealed that P1 could produce longer and more complex verb phrases, although these were infrequent.

Regarding P1's goals and PR verbs, his goal was to improve his speech in general, and despite encouragement to refine this goal he maintained that he did not want to do this. His PR verbs were predominantly related to football as he reported that following his favourite team (Arsenal) was his main interest. P1 was content for the research student to choose his PR verbs which he could then approve or not. The researcher therefore suggested that verbs related to his aphasia and stroke would also be useful to practise (e.g., cough, drink, read, talk) particularly as P1 usually did not have anyone to advocate for him during health consultations, as well as verbs relating to his daily activities (e.g., texting, feeding (his cat)) and P1 agreed. (See Appendix C for P1's PR verbs (and those chosen by all the participants)).

Table 11 P1 Baseline performance.

Outcome Measures - Verb & Sentence Production	T 1	T 2
OANB Verbs (n=50)	30	23
OANB Nouns (n=50)	36	37
PR Verbs Treated (n=20)	7	7
PR Verbs Untreated (n=20)	8	6
PR Verbs in Sentences Treated (n=20)	12	8
PR Verbs in Sentences Untreated (n=20)	7	7
VAST Sentence Construction Test (personally irrelevant verbs) (n=20)	10	n/a
Outcome Measures - Discourse		
CAT Spoken Picture Description Total Score	18	19
Verb Types	3	2
Verb Tokens	3	3
Verb + 1 argument	5	4
Verb + 2 arguments	2	4
Verb + 3 arguments	0	1
Verb + complex phrase	0	0
Autobiographical Memory Interview (AMI)		
Verb	0	1
Verb + 1 argument	3	5
Verb + 2 arguments	12	10
Verb + 3 arguments	5	0
Verb + complex phrase	3	0
Outcome Measures - Functional Communication		
CETI PWA	76.25	n/a

5.1.iii Participant 2.

P2 was a 54-year-old woman who was 3 years post stroke at the start of the study. She had no obvious hemi-paresis but used her non-dominant left hand to write. P2 lived with her partner. She was close to her three grown-up children who she saw regularly, and her first grandchild was born during the study. P2 was an award-winning primary school teacher at the time of her stroke – a job she loved - but had to resign from her position because of her aphasia. P2 was volunteering for local charities and participating in two community-based stroke groups whilst involved in this study. She had participated in research projects prior to this one. P2's partner consented to participate in the research. He helped P2 to choose her PR verbs and showed a keen interest in her treatment, though P2 did not ask him to help her with her treatment and he respected her wishes. At the start of the study P2 was a competent user of her Android tablet which she used to carry out her SPT exercises. P2 received her SPT from mid-May 2018 to the end of August 2018.

P2 had a moderate non-fluent aphasia. She had good understanding of both spoken and written single nouns, and of spoken verbs scoring 37/40 (92.5%) on the VCT. P2 had difficulty understanding reversible sentences in both spoken and written modalities suggesting problems mapping thematic roles onto syntactic structures, with responses more delayed in general to written sentences. P2 had excellent repetition skills. She also had very good ability to read words aloud but had limited ability to read aloud non-words suggesting significant damage to grapheme–phoneme conversion rules. P2's scores at baseline on key outcome measures are reported in Table 12 below. P2 Showed instability on some measures. For example, the SCT and the CETI both improved at T2 and this may reflect the impact of contact with the researcher during the baseline period, although it is not clear why this did not impact all OMs at T2. In summary, in terms of lexical retrieval, P2 showed moderate impairment of both nouns and verbs with verbs more impaired. She showed a small syntactic boot strapping effect when producing verbs in sentences, and in terms of sentence production, she most commonly produced verbs with one or two arguments. Narrative production revealed that P2 could produce longer and more complex verb phrases.

For P2, her goals were:

To communicate better in formal situations especially with strangers: “to stop feeling like I am going to get stuck”.

To be able to communicate better on the phone in formal conversations e.g., with energy providers.

To gain some independence in conversations with health professionals.

P2’s PR verbs were chosen by P2 and her partner and were related to cooking (e.g., whisking), holidays (e.g., rent, save), her family (e.g., love) and her aphasia (e.g., read, talk). Interestingly P2 also choose a number of mental verbs relating to her experience of aphasia (e.g., struggle, suffer) as well as to her aspirations to improve (e.g., hope, try).

Table 12 P2 Baseline performance.

Outcome Measures - Verb & Sentence Production	T 1	T 2
OANB Verbs (n=50)	34	37
OANB Nouns (n=50)	39	43
PR Verbs Treated (n=20)	7	7
PR Verbs Untreated (n=20)	7	7
PR Verbs in Sentences Treated (n=20)	11	7
PR Verbs in Sentences Untreated (n=20)	14	12
VAST Sentence Construction Test (personally irrelevant verbs) (n=20)	6	16
Outcome Measures - Discourse		
CAT Spoken Picture Description Total Score	26.5	27
Verb Types	10	8

Verb Tokens	12	13
Verb + 1 argument	4	5
Verb + 2 arguments	4	6
Verb + 3 arguments	0	1
Verb + complex phrase	0	0
Autobiographical Memory Interview (AMI)		
Verb + 1 argument	10	8
Verb + 2 arguments	18	19
Verb + 3 arguments	5	1
Verb + complex phrase	2	1
Outcome Measures - Functional Communication		
CETI PWA	38.25	50.25
CETI SO	29.063	35.375

5.1.iv Participant 3.

P3 was 49 years old at the start of the research study. She had a left haemorrhagic stroke at the hospital where she was a theatre nurse three years prior to her involvement in the study. P3 had to give up the profession she loved because of her aphasia and a dense right-sided hemiparesis. She was mobile with a walking aid and lived with two of her three adult children as she had recently separated from her husband of 25 years. She was supported by carers on a daily basis. P3 was referred to the project by her sister who visited her regularly, consented to participate in the research and drove her to her face-to-face treatment sessions. P3 attended a local stroke group but was unable to access any aphasia support groups and was socially quite isolated. She was a competent IT user, using her smart phone and her iPad to text, email, and shop online. P3 received her SPT from the beginning of October 2018 to the beginning of December 2018.

P3 had a moderate – severe non-fluent aphasia together with a mild apraxia of speech. Speech was very effortful for her resulting in long pauses and high levels of frustration when P3 attempted sentences. P3 had good understanding of both spoken and written single nouns, but had difficulty understanding both spoken and written reversible sentences (which was worse for written sentences) indicating problems with mapping. P3's ability to repeat words was affected by the number of syllables, but she had very good ability to read words aloud. P3's scores at baseline on key outcome measures are reported in Table 13 below. She showed a degree of instability which may reflect the impact of the distress she felt at times due to her imminent divorce, as well as recovering from a brief hospitalisation due to a very severe nosebleed. In summary, in terms of lexical retrieval, P3 showed a mild – moderate impairment of nouns and verbs to a largely equivalent degree. P3 was worse at retrieving verbs in sentences than in isolation indicating that syntactic difficulties contributed significantly to her sentence production problems. In terms of sentence production, she most commonly produced verbs with one or two arguments, with narrative production revealing some ability to produce verbs with three arguments, but with no evidence of complex verb phrase production.

For P3, her goals were:

To be more independent when talking with doctors (especially her GP)

To be more confident to use the phone

To be able to discuss personal issues with family more (particularly in relation to her divorce)

P3's PR verbs were chosen by P3 in collaboration with the research student. P3 predominantly chose verbs related to communication (e.g., talk, text and write), feelings and senses (e.g., enjoy, hope, struggle) and cooking (e.g., boil, fry and eat).

Table 13 P3 Baseline performance.

Outcome Measures - Verb & Sentence Production	T 1	T 2
OANB Verbs (n=42)	28	33
OANB Nouns (n=42)	37	33
PR Verbs Treated (n=20)	12	10
PR Verbs Untreated (n=20)	13	9
PR Verbs in Sentences Treated (n=20)	9	5
PR Verbs in Sentences Untreated (n=20)	7	6
VAST Sentence Construction Test (personally irrelevant verbs) (n=20)	8	11
Outcome Measures - Discourse		
CAT Spoken Picture Description Total Score	16	14
Verb Types	4	5
Verb Tokens	4	6
Verb + 1 argument	2	3
Verb + 2 arguments	2	1
Verb + 3 arguments	0	0
Verb + complex phrase	0	0
Autobiographical Memory Interview (AMI)		
Verb + 1 argument	2	1
Verb + 2 arguments	6	10
Verb + 3 arguments	2	2
Verb + complex phrase	0	0
Outcome Measures - Functional Communication		
CETI PWA	55.125	52.9375
CETI SO	49.125	n/a*

5.1.v Participant 4.

P4 was 57 years old at the start of the research study having had a left CVA seven years earlier. P4 lived with his wife and the youngest of his four children who was taking her A levels. (Although P4's wife consented to participate in the research, she ultimately found this difficult because midway through the study, due to the compulsory purchase of their house, they had to move, and P4's wife had the burden of responsibility for organising this). P was an ex-soldier: he was a navy diver and saw active service during the Falklands War including being on board his ship when it was hit by an Exocet missile, although he was not hurt during the incident. After leaving the army, P5 met his wife in the USA. They went on to live in Spain and France where P4 earned a living selling wine, becoming fluent in French and Spanish as a result. P4 and his wife moved back to the UK following his CVA. P4 was a very competent user of IT and was an avid online shopper. He predominantly used a PC to carry out his treatment exercises but also used an iPad when he was on an intensive, residential physiotherapy course during the research study, and when he and his wife went to visit his terminally ill mother for three weeks. P4 received his SPT from mid-April 2019 to late June 2019.[19]

P4 had a moderate - severe fluent aphasia. P4 had impaired spoken and written word comprehension (scoring 26/30 and 23/30 on CAT spoken and written word-to-picture matching respectively) reflecting slow processing, with requests for repetition in the case of spoken comprehension. His spoken and written sentence comprehension was also impaired (he scored 21/32 and 12/32 on CAT spoken and written sentence-to-picture matching respectively), with his errors demonstrating slow processing (especially in relation to written sentences), as well as problems with comprehension of reversible sentences in both modalities reflecting a difficulty with

[19] Although P4 was recruited just before P5, his treatment was delayed because his wife suffered a panic attack whilst driving and could not drive him to the clinic whilst she waited for treatment. Therefore, P4 started his treatment after P5.

mapping thematic roles onto syntactic structures. Additionally, P4 had severely impaired noun retrieval scoring 16/48 on the CAT naming test, with his errors including both semantic and phonological paraphasias). P4's ability to retrieve verbs was slightly better (see Table 14 below) but was still severely impaired, as was his repetition. In relation to the latter, he commented that "right away I can say it but I leave it I can't" suggesting problems with working memory. Finally, P4's written language was also severely impaired with his written output demonstrating semantic paraphasias and neologistic production, as well as evidence of preserved knowledge of irregular spelling (e.g., YATCH for yacht in spelling to dictation).

P4's scores at baseline on key outcome measures are reported in Table 5.7 below. P4 also shows a degree of instability in his baseline performance. This may sometimes reflect difficulty comprehending the assessment tasks, although every effort was made to explain these using simple spoken and written instructions, and practice examples when allowed. In summary, P4 had severely impaired lexical retrieval, but with verb retrieval better than that of nouns. In this regard, it should be noted that P4 produced a number of ambiguous noun/verb responses during action naming (such as "drink" for drinking) for which he was not credited. If these were credited his naming of verbs would be much superior to his naming of nouns increasing to 18.5/42 (44%). P4 shows a syntactic bootstrapping effect in relation to verb retrieval in sentences when he most commonly produced verbs with one or two arguments. Narrative production revealed his ability to produce verbs with three arguments and complex verb phrases, though these were infrequent.

P4 did not find it easy to formulate his goals but ultimately settled on ones which reflected his desire to be able to chat more easily to fellow ex-servicemen at the weekly group he attended, and to better reminisce about his extensive travels across the world:

To be able to express himself more easily when talking about his service in the Royal Navy.

To be able to talk about his extensive travels/holidays better.

The PR verbs chosen by P4 related to travel (e.g., fly) and his army service (e.g., fight, shoot) but also included several verbs related to gardening (e.g., dig) which was his main (if somewhat reluctant) occupation since his stroke, as well as verbs relating to communication (e.g., write) and shopping (e.g., pay).

Table 14 P4 Baseline performance.

Outcome Measures - Verb & Sentence Production	**BL 1**	**BL 2**
OANB Verbs (n=42)	11	6
OANB Nouns (n=42)	7	9
PR Verbs Treated (n=20)	6	4
PR Verbs Untreated (n=20)	7	3
PR Verbs in Sentences Treated (n=20)	11	11
PR Verbs in Sentences Untreated (n=20)	11	10
VAST Sentence Construction Test (personally irrelevant verbs) (n=20)	15	10
Outcome Measures - Discourse		
CAT Spoken Picture Description Total Score	12.5	17.5
Verb Types	4	4
Verb Tokens	4	6
Verb + 1 argument	1	3
Verb + 2 arguments	2	1
Verb + 3 arguments	1	0
Verb + complex phrase	0	1
Autobiographical Memory Interview (AMI)		
Verb + 1 argument	6	6
Verb + 2 arguments	19	15
Verb + 3 arguments	5	2
Verb + complex phrase	3	1

Outcome Measures - Functional Communication		
CETI PWA	46.1875	45.375

5.1.vi. Participant 5.

P5 was 81 years old at the start of the research study. She had a left CVA in March 2016 and was 28 months post stroke at the start of the study. She had no obvious hemi-paresis and wrote with her dominant hand although this was quite slow due to arthritis. P5 lived with her 84-year-old husband of 60 years. She had three children and a total of 26 grandchildren/great grandchildren. P5's family were the cornerstone of her life, and her house displayed many pictures of the different generations of her family. However, P5 was also very proud of her career: she had worked as a secretary for the same company for 26 years and remembered this fondly. P5 was a competent user of computers when working but had not used them regularly since retiring. P5's husband used a tablet, and he bought P5 a laptop for her to carry out her SPT exercises. P5 was a keen gardener and cook but reported that she had not been able to resume cooking as much as she did before her stroke as she found it hard to remember recipes. P5 attended a stroke group weekly during her participation in the study and had taken part in a previous research study. P5 received her SPT from mid-October 2018 to mid-February 2019.

P5 had a mild anomic aphasia. P5 demonstrated difficulty in understanding reversible sentences in both written and spoken modalities. She had good repetition skills. In terms of spoken picture naming P5 showed a mild impairment. However, she was only able to produce seven nouns in the CAT test of word fluency, suggesting significant semantic priming from the presence of a picture. P5's spoken picture description also demonstrated her word finding difficulties as well as possible syntactic difficulties with a limited variety of syntactic structures used. P5 had excellent reading aloud skills scoring 100% on all tasks except reading aloud of nonwords (70%). Her written skills were also very good at the single word level. However, her written picture description demonstrated agrammatism with functions words omitted and a very limited variety of syntactic structures produced (only NP VP NP structures).

P5's scores at baseline on key outcome measures are reported in Table 15. P5 again showed some instability in baseline performance. For her, this likely reflected a reduction in her linguistic anxiety as she got to know the researcher by T2. (See later discussion in Chapter 7.6.v). In summary, in terms of lexical retrieval, P5 showed a mild impairment, but with a slight syntactic bootstrapping effect in relation to verbs. In sentence production she most commonly produced verbs with two arguments, in both constrained sentence production and in picture description. However, narrative production revealed P5's significant ability to produce verbs in complex verb phrases. The contrast between P5's performance in constrained assessment tasks and in a less constrained discourse task led to the hypothesis that P5's aphasia likely reflected an interaction between mild linguistic impairment and linguistic anxiety as defined by Cahana-Amitay et al. (2011). This interaction will be addressed in detail in the discussion of P5's results (Chapter 7).

P5's goals were:

To improve her fluency when talking to family and friends making it easier/less hard work.

To increase her independence when making phone calls that she had to make.

To increase her fluency and confidence talking about things that she enjoyed.

P5's PR verbs predominantly related to food and drink reflecting her love of cooking (e.g., bake, weigh) and also of gardening (e.g., grow and water). She also chose several verbs relating to her feelings (e.g., enjoy, love) in relation to her third goal above.

Table 15 P5 Baseline performance.

Outcome Measures - Verb & Sentence Production	BL 1	BL 2
OANB Verbs (n=42)	31	37
OANB Nouns (n=42)	38	37
PR Verbs Treated (n=20)	8	8
PR Verbs Untreated (n=20)	8	9
PR Verbs in Sentences Treated (n=20)	10	14
PR Verbs in Sentences Untreated (n=20)	14	11
VAST Sentence Construction Test (personally irrelevant verbs) (n=20)	15	17
Outcome Measures - Discourse		
CAT Spoken Picture Description Total Score	29	29
Verb Types	2	4
Verb Tokens	2	7
Verb + 1 argument	0	3
Verb + 2 arguments	7	7
Verb + 3 arguments	1	2
Verb + complex phrase	0	0
Autobiographical Memory Interview (AMI)		
Verb + 1 argument	7	6
Verb + 2 arguments	36	43
Verb + 3 arguments	12	17
Verb + complex phrase	11	11
Outcome Measures - Functional Communication		
CETI PWA	59.125	61.4375

5.1.vii Participant 6.

P6 was 59 years old at the start of the study having had a left sided CVA 21 months earlier. He had a dense right-sided hemi-paresis but was mobile with a walking aid and used his non-dominant left hand to write. P6 had worked in the financial sector at the time of his stroke as a specialist in reinsurance. He had resigned from his position because of his aphasia. He lived alone but was in regular contact with his family and with friends from work. He also attended a community-based stroke group and had participated in a research project prior to this one. P6 did not identify a SO or friend to participate in this research study. P6 was a very competent IT user: he had a laptop which he used to download music, email, and browse the Internet, and to carry out his SPT exercises. P6 received his SPT from mid-June 2019 to late August 2019.

P6 had a severe expressive non-fluent aphasia in the context of mildly impaired comprehension. In terms of his functional communication, P6 was sometimes able to produce a noun but this was extremely effortful, meaning that P6 relied very heavily on writing key nouns and the skill of his communication partner in order to communicate. P6's spoken naming was severely impaired and was affected by imageability and length. He was however, extremely phonemically cueable, often producing the cued word almost simultaneously with the cue. This suggests significant impairment in mapping from semantics to phonology. P6's written naming was much less impaired (he scored nearly 100% on the writing subsection of the CAT) and, as noted above, he frequently wrote down words he could not say. However, he was unable to self-cue as he could not convert graphemes to phonemes. Reliance on the lexical route was also evident in his repetition as P6 had excellent ability to repeat words but was completely unable to repeat nonwords. His ability to read words aloud was severely impaired, and was affected by frequency, regularity, and length. P6's ability to read aloud nonwords was also poor. However, P6 had very good understanding of both spoken words and sentences. He also had very good understanding of written words, but written sentence comprehension was impaired for reversible sentences indicating problems with mapping.

P6's scores at baseline on key outcome measures are reported in Table 16 below. In summary, in terms of lexical retrieval, P6 was severely impaired, with verbs being more severely affected. He showed a small syntactic bootstrapping effect in relation to verb retrieval in sentences but was completely unable to produce any verbs together with an argument likely due to the severity of his word retrieval difficulties, in possible combination with a deficit in syntactic processing for output. The severity of his expressive difficulties meant that he could not complete the AMI and it was distressing for him. The AMI was not therefore used as an OM for P6.

For P6, his goals were:

To improve his ability to talk in general

To enhance his ability to discuss his interests/hobbies

To increase his independence when he needed to communicate (e.g., on the phone with his GP)

P6's PR verbs were chosen by P6 in collaboration with the research student. P6 predominantly chose verbs related to cooking such as blend, steam, and poach (he was a very keen and accomplished cook prior to his stroke), verbs related to entertainment such as play, score and putt (he was an avid heavy rock and Chelsea FC fan, and a keen golfer), and finally verbs related to communication (e.g., email, phone, and talk).

*Table 16 P6 Baseline performance (*the AMIs were not completed for P6 as his expressive impairment was so severe they were distressing for him to attempt.*

Outcome Measures - Verb & Sentence Production	**BL 1**	**BL 2**
OANB Verbs (n=42)	9	9
OANB Nouns (n=42)	15	20
PR Verbs Treated (n=20)	1	4
PR Verbs Untreated (n=20)	1	3
PR Verbs in Sentences Treated (n=20)	2	8
PR Verbs in Sentences Untreated (n=20)	3	5
VAST Sentence Construction Test (personally irrelevant verbs) (n=20)	0	2
Outcome Measures - Discourse		
CAT Spoken Picture Description Total Score	7	7
Verb Types	0	1
Verb Tokens	0	6
Verb + 1 argument	0	0
Verb + 2 arguments	0	0
Verb + 3 arguments	0	0
Verb + complex phrase	0	0
Autobiographical Memory Interview (AMI)		
Verb + 1 argument	n/a*	n/a
Verb + 2 arguments	n/a	n/a
Verb + 3 arguments	n/a	n/a
Verb + complex phrase	n/a	n/a
Outcome Measures - Functional Communication		
CETI PWA	62.8125	60.3125

5.2 The technical feasibility of self-delivering aphasia treatment by computer.

One factor which could potentially affect the feasibility of successfully self-delivering the SPT was that participants would need to use whatever devices they already owned, since there was no funding to purchase a standard device for each participant. Encouragingly no person was excluded from participating in the research because they did not have access to a computer. The main technical feasibility issue was thus whether the SPT treatment package would work on desktop computers, laptops, tablet devices and smartphones, and whether it would work on both Windows based and Apple devices.

A range of devices was used to self-deliver treatment (see column 3 Table 17 below). These included both Apple and Windows based devices, with the SPT working on all devices without technical problems. The most used device was an iPad: P1 and P3 used this as the only/main way to self-deliver treatment respectively, whilst P4 used it only when he could not access his desktop (see below). P2 used an Android tablet, and P5 and P6 used Windows based laptops. P3 sometimes used her Smartphone to practise and reported doing this successfully although, as she commented, the display was small. P4 used an old iPad to self-deliver treatment (rather than his usual desktop) when he was attending a residential physiotherapy course for five days and could not make the SPT audio cues work but other than this, the SPT exercises worked as intended for all participants.

Table 17 The technical feasibility of the SPT.

	1. Age	2. Sex	3. Device used to self-manage SPT	4. Frequent computer user	5. Access to significant other who could help with IT issues	6. Needed help to solve IT issues
P1	53	M	iPad	Yes	No	Yes
P2	54	F	Android Tablet	Yes	Yes	No
P3	49	F	iPad & Smart Phone	Yes	Yes	No
P4	57	M	Windows Desktop & iPad	Yes	Yes	Yes
P5	81	F	Windows Laptop	No	No	Yes
P6	59	M	Windows Laptop	Yes	No	No

5.2.i. Participants' ability to use technology.

Another factor that could affect the feasibility of self-delivery of the SPT was participants' ability to use their computer independently. This was therefore screened prior to treatment (during baseline testing) using a form of dynamic assessment. Four participants were immediately able to follow instructions and access a trial SPT exercise. Two participants (P1 and P5) were identified as needing more help, but both were able to respond successfully to step-by-step instructions, with the addition of pointing, and accessed the trial exercise. Because P1 and P5 were identified as needing more help, trial slides were demonstrated to them prior to starting the treatment phase, to give them practice with the SPT exercises.

Four participants did not experience any significant difficulty accessing and using the SPT exercises, however two participants (P1 and P5) did. P1 was unable to self-deliver treatment for the first week of treatment because he could not access the exercise slides independently. The researcher demonstrated how to do this again in week 2 of the treatment phase, and he subsequently practised independently. P5 experienced significant difficulties with being able to self-deliver treatment. She was not a frequent computer user at the start of treatment, unlike the other five participants. She reported that she used a computer all the time at work (prior to her retirement some 20 years ago) and subsequently at home, but she had not used a computer at all since her stroke three years earlier. Additionally, her husband had purchased a new computer for her to practise her exercises and this also caused difficulties as neither she nor her husband was familiar with the equipment. Despite being given repeated demonstrations of how to access the SPT exercises at the start of each face-to-face session, and leaving written, aphasia-friendly instructions for her, P5 experienced significant difficulties throughout the study and could not self-deliver the expected amount of treatment (2 hours per week). A fuller discussion of the amount of treatment self-delivered by each participant is given in section 5.6

5.2.ii. Feasibility of uploading the SPT exercises.

The SPT exercises were made available to participants for home practice in two ways. For P1, P2 and P3 they were sent by email just prior to each face-to-face session. P2 and P3 downloaded their exercises independently, whilst P1 waited for the researcher to arrive and downloaded them with her help during the face-to-face session. For P5 and P6 the exercises were transferred onto their laptops using a flash drive. P6 did this independently after the first session, whilst P5 always relied on the researcher to do this. For P4, when the face-to-face session was at the university clinic, exercises were subsequently emailed to his wife to upload onto his personal desktop for home practice. When the face-to-face session was at his home, exercises were transferred by flash drive onto P4's desktop, with this being completed independently by P4.

5.3. Feasibility of delivering the low dose face-to-face sessions.

In terms of the delivery of the face-to-face sessions, these took place at home for five of the six participants (P3 received treatment at a university[20] local to her). Although a room was available at the university clinic for participants to receive treatment, this was difficult to access for all of them: P2, P3, P4 and P6 were all unable to drive since their strokes (P1 and P5 did not drive), whilst travel by public transport was also difficult because of the length of the journey from participant's homes and/or the length of the walk from stations to the clinic.

5.4. The feasibility of selecting a set of PR verbs.

All participants successfully chose a set of 40 PR verbs (half of which were targeted in treatment and half were control items). Only one significant other was involved in selecting PR verbs: P2's partner helped her generate a list which was presented to the researcher in the session following the initial discussion about PR verbs. For the other five participants, the researcher was more involved in the process of selecting participants' PR verbs. P1 accepted the researcher's suggestions almost without question, with football related verbs being the most commonly selected (18/40). For the other four participants, it was a process of negotiation with the majority of the researcher's suggestions accepted but with occasional verbs also rejected (e.g., P3 rejected "cuddle" (suggested in relation to her children, pet dog and cat) because it reminded her of the divorce she was going through).

In total, participants chose 120 different verbs. (These are reported in detail in Appendix D (column 1). The number of participants choosing each verb is given in column 2 and the topic (or

[20] P3 was unable to drive and also unable to access public transport to access the clinic at City University. The researcher therefore travelled to P3's house to carry out pre and post treatment assessments. However, because each journey took over two hours, it was agreed that P3's sister would drive her to Bedford University for the face-to-face treatment sessions as this was midway between the researcher's home and that of P3.

category) into which each verb fell is given in column 3 (following the method used by Palmer et al., (2017). Finally, the subtopic of each verb is given in column 4). The most frequently chosen verbs (i.e., by at least three of the six participants (50%) are summarised in Table 18 below: there are 27 verbs which represent 22.5% of the 120 verbs chosen in total.

Table 18 Personally relevant verbs chosen by at least 50% of participants (n = 6).

PR Verbs	No of P's choosing PR Verb	Verb Topic (after Palmer et al., 2017 re categorisation of nouns)
Talk	6	Communication mediums & modes
Understand	6	Communication mediums & modes
Eat	6	Food/drink
Buy	6	Shopping
Pay	6	Shopping
Phone	5	Communication mediums & modes
Read	5	Communication mediums & modes
Feel	5	Feelings & senses
Think	5	Feelings & senses
Visit	5	Travel
Write	4	Communication mediums & modes
Watch	4	Entertainment
Enjoy	4	Feelings & senses
Cook	4	Food/drink
Drink	4	Food/drink
Walk	4	Health/travel
Shop	4	Shopping
Drive	4	Travel
Work	4	Work
Listen	3	Communication mediums & modes
Hope	3	Feelings & senses
Try	3	Feelings & senses
Help	3	Feelings & senses
Boil	3	Food/drink
Clean	3	House
Catch	3	Nature & gardening
Grow	3	Nature & gardening

The *topics* to which the total 120 verbs chosen by the six participants belonged are given in Table 19 below. The most common topic for the PR verbs chosen were verbs relating to *entertainment and hobbies* (e.g., watch (TV), score, play) and *food and drink* (e.g., eat, drink, cook, boil) with 20% of verbs chosen belonging to each of these topics respectively (or 48 of the total 240 verbs chosen by participants). The next most common verbs were those relating to *nature* (including pets) *and gardening* (e.g., stroke, grow, water) (12.5%), followed by *feelings and senses* (e.g., think, feel, enjoy) (11.7%), and *communication mediums and modes* (e.g., read, talk, write) (10% of verbs). The other eight topics to which selected verbs belonged each represented less than 10% of the total verbs chosen.

Table 19 Topics into which personally relevant verbs fell.

Verb Topic	Number of verbs chosen (n=120)
Entertainment/Hobbies	24 (20%)
Food & Drink	24 (20%)
Nature & Gardening	15 (12.5%)
Feelings & Senses	14 (11.7%)
Communication Mediums & Modes	12 (10%)
Work	9 (7.5%)
Health	7 (5.8%)
Travel	6 (5%)
Shopping	5 (4.2%)
House	2 (1.7%)
Money	2 (1.7%)
Housework	1 (0.8%)
Personal Care	1 (0.8%)

Finally, Table 20 presents the topics into which personally relevant verbs fell for each individual participant. The reader will note that the total number of verbs adds up to 122 rather than 120 because two verbs (swim and walk) fell into two categories (health and travel) and so were counted twice. Furthermore, where n>40 for a participant this is because a verb chosen by them fell into two categories e.g., for P1 walk was included in both health (because he had significant problems walking) and under travel as he walked to and from public transport for all his activities outside of his home.

Table 5.13. Verb topics of the 40 verbs chosen by each participant in this PhD study.

Table 20 Verb topics of the 40 verbs chosen by each participant in this study..

	Number of verbs (n=40)					
	P1 (M)	**P2 (F)**	**P3 (F)**	**P4 (M)**	**P5 (F)**	**P6 (M)**
Entertainment/ Hobbies	16 (40%)	2 (5%)	2 (5%)	2 (5%)	1 (2.5%)	10 (25%)
Nature & Gardening	4 (10%)	3 (7.5%)	3 (7.5%)	8 (20%)	7 (17.5%)	0
Housework	0	1 (2.5%)	1 (2.5%)	0	1 (2.5%)	0
Food & Drink	2 (5%)	6 (15%)	6 (15%)	3 (7.5%)	15 (37.5%)	12 (30%)
Personal Care	0	1 (2.5%)	1 (2.5%)	0	0	0
Health	5 (12.5%)	2 (5%)	2 (5%)	2 (5%)	0	1 (2.5%)
Feelings & Senses	1 (2.5%)	12 (30%)	8 (20%)	4 (10%)	6 (15%)	3 (7.5%)
Communication Mediums & Modes	8 (20%)	6 (15%)	8 (20%)	5 (12.5%)	5 (12.5%)	6 (15%)
Travel	1 (2.5%)	3 (7.5%)	3 (7.5%)	6 (15%)	1 (2.5%)	5 (12.5%)
Work	1 (2.5%)	0	2 (5%)	6 (15%)	1 (2.5%)	2 (5%)
House	0	0	1 (2.5%)	2 (5%)	0	0
Shopping	3 (7.5%)	3 (7.5%)	4 (10%)	4 (10%)	3 (7.5%)	2 (5%)
Money	0	2 (5%)	0	0	0	0

5.5. Participants' and significant others' views on the acceptability of the SPT.

5.5.i. Participants' views on the acceptability of the SPT.

Participants' responses to the exit interview questions are reported in Figure 4. All participants rated themselves as *very experienced* at using a computer (all rated themselves at four on question 1). P2, P3 and P6 found doing the therapy on the computer very easy (question 2) and needed no help (question 3); whereas P1, P4 and P5 found it less easy and needed a little or some help. (P4 elaborated upon this by saying that he occasionally needed help with the Wh-question exercise and with the Verb Networks B exercise which was based on VNeST). All participants perceived change in their communication as a result of treatment, with four perceiving a *lot* of change (P2, P3, P5, P6) and two *some* change (P1, P4) (question 4), and all had used their words in daily communication either every day (P1, P3, P4, P5) or most days (P2, P6) (question 5). Finally, three participants indicated that they would *definitely* use the SPT again (question 6), whilst three rated this as *very likely* (P1, P3, P5). P3 indicated that the reason she rated this question at three was because she had found the exercises relating to Wh- questions hard; others gave no reason.

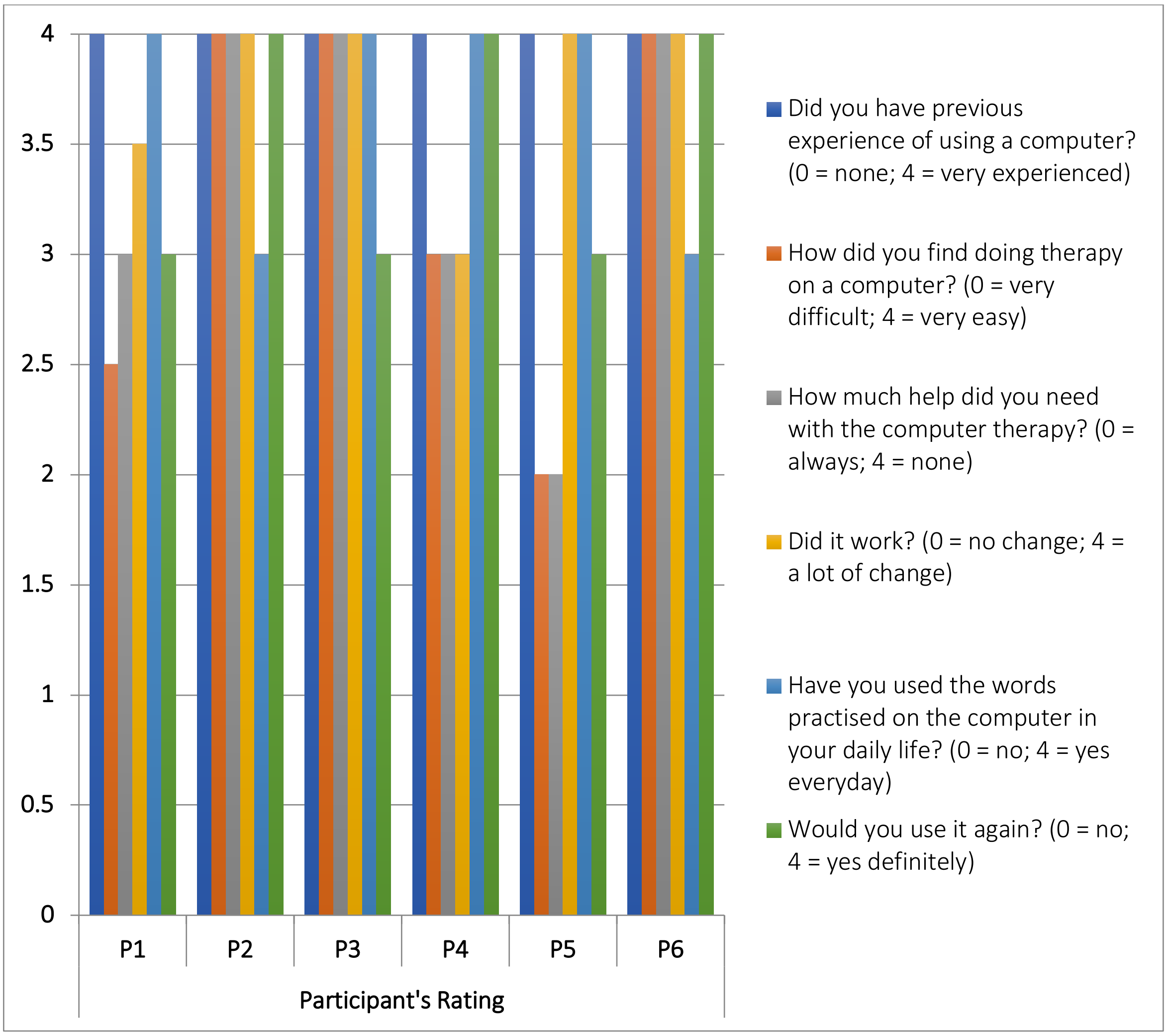


Figure 4 Participants' responses to rating scale questions on the acceptability of the SPT.

In addition to the closed rating scale questions, participants responded to questions regarding benefits and disadvantages of computer treatment and influences on decision-making to engage. The interview relating to these questions was kept brief to comply with the ethical approval for the research, and the responses were not sufficient to merit content analysis thus direct quotes from participants are reported.

Benefits reported by participants included increases in *confidence*: P1 said his confidence increased "because you try this here" (pointing to his iPad). P3 and P6 also felt that doing the SPT had improved their confidence but did not expand on this. When asked the same question, (about any benefits of the SPT) P2 said “Yes because I timetable once a day an hour in the computer that’s it” implying that self-managing her aphasia treatment enabled her to exert some control over her treatment – to be proactive in managing it. P4 said “I found it difficult. Have to think. Correct them now but from before nothing not anything. But a little bit yes” implying that he felt that he had benefitted from treatment. When prompted to expand on this (he was asked “Did it improve your confidence?”) he said “I can’t write it down. I can see it alright alright” again implying benefit but it was not clear what this was. P5 responded that she felt she had “sometimes” benefitted from the SPT and did not expand on this.

Improvements in *talking* were reported by P2 who said “sometimes . . . twenty doing words and I say . . talking or listening in my head . . . I can do it . . . in my mind I can’t say . . . foremost is in my head the doing words I concentrate on the doing words verbs and use it.” Additionally, P2 reported that friends had commented her speech had improved. Improvements in talking were also reported by P3 and although she did not expand on this, her significant other (her sister) gave specific examples of this (see below).

In response to question 2, (Are there any disadvantages of computer treatment?), very few disadvantages were mentioned by participants. Technical issues were cited by P1 and P5. P4 said “if you can all day [do treatment] it will be better. Once a day is good but it’s not the same as everyday” meaning he appreciated the opportunity to practise daily that self-management of treatment gave him, but he still would have liked more intensive face-to-face treatment (similar to an intensive residential physiotherapy course he had just attended). P2 and P3 felt there were no disadvantages to computer treatment and P6 was very eager to express his opinion that face-to-face visits were not necessarily needed to supplement computer delivered treatment.

In response to the question 3 (What would influence your decision to do computer treatment?), P1 said “you’ve got to do it because it’s confidence.” P2 said “In the future a new . .

can't say it . . is er level is higher" meaning that she would have liked some more difficult exercises. She also said she would like to be able to check her understanding (of the exercises) in the future. P3 indicated that time was an issue (i.e., having the time to practise) and that the lack of face-to-face treatment meant that computer-based treatment was the only option for her (and by implication for many PwA). P4 talked about the importance of a computer to his life: "to have a computer is important. If I didn't have it I think afraid I'd be nothing." P5 commented that she felt her age was an issue in relation to doing computer-based treatment: she said she was a "bit old". P6 indicated that improving his talking would be the main influence on his decision to do computer-based treatment and was adamant that he did not see self-delivered treatment as second best.

5.5.ii. Significant others' views on the acceptability of the SPT.

Two significant others (SOs) took part in a brief exit interview administered face-to-face with P2's partner (face-to-face) and by phone with P3's sister. The interview used the topic guides from Palmer et al. (2013) and related to the: ease/acceptability of using a computer; benefits, disadvantages, and limitations of computer therapy to the participant and SO; and perceived changes (talking, daily activity).

In relation to the ease/acceptability of using a computer, P2's partner said:

"*I think it was good. I think once she had mastered how to find the PowerPoints, and of course there was a number of little icons to touch on the screen so erm I think she was really confident with it. You know because you had put icons where it gives a word and then it gives a phrase and then it gives a sentence, I think she bypassed certain ones after a while because she knew what was coming and what was needed. So I think she felt quite confident about accessing what you'd put on.*"

P3 SO's comments (on the ease/acceptability of the computer–based SPT) related to the realities of the availability of SLT services: she said if P3 hadn't done the computer-based treatment she wouldn't get (and hadn't got) any treatment in her area.

In terms of any disadvantages of computer therapy P2's SO said:

> "*Well if there was someone sitting in front of her you know giving her those cues and interacting with her that would have been so much better but obviously that can't happen*."

The latter comment again referred to the limited face-to-face services available. The SO of P3 also made reference to the perceived superiority of face-to-face treatment in that she said P3 believed she would have gained more from face-to-face treatment. However, neither P2's or P3's significant other made any comments in relation to perceived disadvantages of the computer-based treatment. Indeed P2's SO said:

> "*I think it's possibly the best it could have been . . . I think as working as an independent activity working on her own it was quite interactive*."

Additionally, both SOs were positive about the impact of the SPT on talking and daily activity: P2's SO stated that her talking had improved (with the caveat that he had seen steady improvement since her stroke so could not be certain this was due to the SPT). However, he also said that P2 had started to use words which were "unusual for her" and that she had "more words swimming around inside". P3's SO felt that there was definite improvement in her speech: "I think she's talking in more structured sentences" and that she was "more fluent" when she spoke.

In terms of changes in daily activity, P3's SO commented on P3's increased confidence. Specifically, that she had started to answer the phone when previously she wouldn't and that she had rung the SO (i.e., her sister) rather than texting as previously. P2's SO commented that P2 had started going to the gym, and also referred to improvement in her confidence.

5.6. What factors influence compliance with self-delivered computer treatment?

Any conclusions regarding the potential factors that might influence compliance with self-managed computer treatment must be very tentative because of the small number of participants in this study. With this caveat in mind, scrutiny of Table 5.14 gives an indication of the likely influential factors. Columns 4, 5 and 6 report on whether participants were frequent computer users prior to the start of the project (column 4), whether they had access to a significant other who could help them with IT issues (5), and whether they needed any help (6). Column 9 reports the total amount of treatment self-managed by each participant. Comparison of column 9 with columns 4, 5 and 6 reveals that the two participants who completed the least amount of self-managed treatment (and failed to reach the target amount) were P1 and P5. Both of these participants needed help with IT issues, and neither had access to someone who could help. P1 lived alone and, whilst P5's 84 year old husband did use a computer from time-to-time, he was not able to assist her. P4 also needed help with IT issues but unlike P1 and P5 he had someone who could help him (his wife) and he self-delivered (more than) the requested amount of treatment. Also of interest is that P5 was the only participant who was not a frequent computer user at the start of the study. Additionally, health may have played a role in compliance with the SPT: P5 was hospitalised twice during the treatment phase, and P1 once. The combination of lack of recent experience with a computer together with not having ready access to someone who could help with IT issues and periods of poor health likely combined to make it difficult for P1 and P5 to complete the requested amount of self-managed treatment. Finally, it should be noted that whilst P3 self-managed just over the expected amount of treatment (at 17 hours 40 minutes), this is likely an underestimate as P3 only completed her exercise diary for 5 weeks but reported that she practised for a similar amount of time during the weeks she did not complete her diary.

5.7. Fidelity of the SPT: the amount of face-to-face and self-managed treatment delivered.

The intention for the SPT was that each participant would receive one hour of face-to-face treatment per week for eight weeks, delivered by the researcher (either at a participant's home or at the university clinic), supplemented by at least two hours treatment per week self-delivered via computer (via either desktop, laptop, tablet, or smartphone) also for eight weeks. Table 21 reports on the fidelity of how treatment was delivered (columns 7 and 8).

Table 21 The fidelity of the SPT delivery.

	1 Age	**2 Gender**	**3 Device used to self-deliver SPT**	**4 Frequent computer user**	**5 Access to carer who could help with IT issues**	**6 Needed help to solve IT issues**	**7 Amount of sessions delivered F2F**	**8 Duration over which treatment delivered (weeks)**	**9 Total amount of treatment self-delivered**	**10 Average amount of self-delivered treatment per week**	**11 Range of amount of self-delivered treatment per week**
P1	53	M	iPad	Yes	No	Yes	8	11	11 hours	1 hour	0- 2 hours
P2	54	F	Android Tablet	Yes	Yes	No	8	11	59 hours 6 mins	5 hrs 23 mins	2 hrs 56 mins - 7 hrs 55 mins
P3	49	F	iPad & Smart Phone	Yes	Yes	No	8	9	17 hours 40 mins*	1 hrs 46 mins	2 hrs 55 mins – 4 hrs
P4	57	M	Windows Desktop & iPad	Yes	Yes	Yes	6	8**	28 hours 40 mins***	2 hrs 36 mins	2 hrs - 3 hrs 30 mins
P5	81	F	Windows Laptop	No	No	Yes	8	17	6 hours 47 minutes	24 mins	0 - 2 hrs 13 mins
P6	59	M	Windows Laptop	Yes	No	No	8	11	45 hrs 29 mins	4 hrs 8 mins	1 hr 46 mins - 9 hrs 1 min

*P3 only completed her exercise diary for 5 weeks but reported she practised for a similar amount on the other weeks. Therefore, this is very likely an underestimate of the amount of self-administered treatment.

**P4 only received 6 face-to-face sessions as he was visiting his terminally ill mother for planned sessions 7 and 8.

***P4 only completed his exercise diary on 5 occasions but self-reported that he had completed at least 2 hours on two other weeks, and this was confirmed by his wife and is therefore included in the total amount of treatment.

The intention was to deliver the SPT over eight weeks. However, the duration over which the face-to-face and self-managed treatment was delivered was longer than this for all participants meaning that treatment was not delivered as intensively as intended (column 8). Treatment was delivered over 9 weeks for P3, and over 11 weeks for P1, P2, P4 and P6. P5's treatment was delivered over 17 weeks as she was hospitalised twice during the treatment as well as having a flu-like cold on two occasions meaning that her husband postponed treatment sessions. P1 was also in hospital for 1 week during the treatment phase, whilst P1, P3 and P4 also experienced illness at home during the treatment phase which meant they cancelled sessions. P2 went on holiday during the treatment phase (but continued to self-manage treatment during these two weeks) and P2 also cancelled a treatment session as she was too hot to concentrate during the extreme heat of summer 2018. The duration of P6's treatment was extended because of personal reasons for the research student on two occasions and problems travelling to P6's flat to deliver treatment due to rail problems on the third. However, P2, P4 and P6 carried out enough home practice during these weeks to ensure that both the target *total* amount of treatment was self-managed and to the recommended weekly dosage (at least two hours per week) – see discussion below[21].

[21] It should be acknowledged here that the research student was very flexible in accommodating participants' schedules, and this very likely contributed to both retention of participants and how much treatment they self-delivered – see Chapter 7 Discussion 7.2.i.

The amount of treatment delivered *face-to-face* will now be discussed (See Table 21 column 7). The intended eight sessions were delivered to five of the six participants. P4 only received six face-to-face sessions as he went to visit his terminally ill mother for the final scheduled two sessions of treatment. His last face-to-face session (six) was therefore extended to 90 minutes and used to introduce P4 to all the exercises for the final three sessions so that he could self-manage these. His wife also attended the whole of this session so that she could assist him if required.

In terms of the amount of treatment that was *self-delivered,* it was intended that each participant practise for at least two hours per week over the eight-week treatment phase resulting in a total of at least 16 hours of self-delivered treatment. Four of the six participants met this target, with three of these exceeding it by a considerable margin: P2 self-managed over 59 hours of treatment, P6 over 45 hours and P4 approximately 29 hours (see column 9 of Table 21). The two participants who did not meet the target (P1 and P5) were the least competent in terms of their computer skills (as discussed earlier). In addition, neither had a significant other on hand who could assist them if they did have difficulty using their computer-based exercises, and this impacted their ability to administer treatment independently. This meant that P1 did not self-deliver any treatment in the first week because he could not open the SPT exercises. P1 also did not complete his exercise diary for sessions five and six and it was unclear if he had practised or not but, on these occasions, it did not appear to be because of IT issues. P5 had difficulty opening not just her exercises but also her computer itself, and this impeded her ability to self-deliver the treatment repeatedly. She was only able to practise on four occasions, despite the research student demonstrating how to open both the computer and the exercises at each visit and leaving written instructions of how to do this. However, P5 persisted and was able to self-manage an additional 7 hours and 44 minutes of treatment during the maintenance phase. It should also be noted that three of the six participants continued to use the SPT during the maintenance period (between T3 and T4). (It was a deliberate decision not to restrict participants' access to the SPT during the maintenance phase and they were neither encouraged nor discouraged from practicing). P2 reported that she continued to self-manage a significant amount of treatment during the maintenance phase but did not record the amount of this. P5 completed an additional 7 hours and 44 minutes practice and P6 completed an additional 12 hours and 22 minutes. Thus, for these three participants this is not a true maintenance phase, and this is taken into account in the statistical analysis reported in the following sections.

The fidelity of treatment delivery will now be discussed in terms of tailoring to individual participants' needs (with reference to the TIDieR guidelines (Hoffman et al., 2014) although tailoring was not determined a priori as TIDieR guidelines expect). The SPT comprised 11 exercises to be introduced during eight face-to-face sessions (and then self-delivered between sessions), giving the potential, if there were no tailoring of the SPT, to deliver 66 exercises in total to the six participants over a total of 48 face-to-face sessions. However, in order to target each participant's particular pattern of impairment, and to adapt to the rate of progress they were making, treatment delivery was tailored thus reducing the number of *exercises* delivered to a total of 53 exercises (delivered during 46.5 sessions - see Table 22 below). The reduced number of *sessions* was due to only 6.5 sessions being delivered to P4 (instead of eight) because of his mother's illness. For example, P6, who had the most severe expressive impairment of the participants, had his treatment tailored such that he spent only one session in the generalisation phase of treatment. In comparison P5, who had the mildest impairment, spent three sessions on generalisation exercises. The exercises chosen in the generalisation phase were also adapted to suit the differing technological skills of P5 and P6. Thus, because P5 had difficulty accessing the SPT, she only carried out the generalisation exercise relating to personal and news stories as this required the least interaction with the HCI, and in particular she did not do the video retell exercise as she found it hard to open up the hyperlink to the videos and also to close the YouTube window. This on the other hand presented no problem to P6 who did all three generalisation exercises. At the end of the SPT, all of the participants were given a revision SPT exercise which contained a sample of PowerPoint slides from each exercise they had practised during the treatment programme.

Table 22 Tailoring of the SPT exercises for each participant in face-to-face sessions. (Number of sessions spent on each exercise).

	P1	P2	P3	P4	P5	P6	Average
Verb Phase Treatment.							
Exercise 1. Sound and Letter Cueing.	1.5	1.5	0.5	1	1.5	1	1.17
Exercise 2. Sentence Closure.	0.5	0.5	0.25	0	0	1	0.38
Exercise 3. Verb Networks A.	1	0	0.75	1	0.5	0	0.54
Exercise 4. Verb Generation.	0.5	1	0.5	0.5	0.25	0.5	0.54
Total	**3.5**	**3**	**2**	**2.5**	**2.25**	**2.5**	**2.63**
Sentence Phase Treatment.							
Exercise 1. Simple Statements, Requests and Yes/No Questions	0.5	0.75	1	1.5	0.5	1	0.88
Exercise 2. Light Verb Exercises	0.5	0.25	0.5	1	0	0.5	0.46
Exercise 3. Verb Networks B	1	1	1	0	1	2	1.00
Exercise 4. Longer Statements and Wh- Questions.	0.5	0.5	0.5	0.25	0.25	0	0.33
Total	**2.5**	**2.5**	**3**	**2.75**	**1.75**	**3.5**	**2.67**
Generalisation Phase Treatment.							
Exercise 1. Personal and News Stories	1	1	1.25	0.5	3	0.33	1.18
Exercise 2. Scripts	0.25	0.5	0.25	0	0	0.33	0.22
Exercise 3. Story and Video Retell	0.25	0.5	1	0.5	0	0.33	0.43
Total	**1.5**	**2**	**2.5**	**1**	**3**	**1**	**1.83**
Revision	0.5	0.5	0.5	0.25	1	1	0.63
Grand Total	**8**	**8**	**8**	**6.5**	**8**	**8**	**7.75**

Table 23 below reports on the fidelity with which the SPT was delivered during face-to-face sessions. This was assessed using treatment fidelity checklists based on the SPT manual and developed prior to the start of treatment (see Appendix B for a sample checklist). Checklists were completed by the research student after sessions and 25% of sessions were cross-checked by listening to the audiotape of the session with 100% agreement. The fidelity of treatment delivery ranged from 92% - 100% with an average fidelity of 98%. Items on the checklist which were always delivered included explaining and demonstrating the use of cues on the SPT slides and participants interacting with the researcher during sessions. Items more variably delivered included discussing the current exercise in the context of the whole treatment programme and discussing any barriers to practice as this was not deemed necessary for some participants in every session.

Table 23 Fidelity of face-to-face treatment delivery. (Emboldened entries cross-checked).

Exercise	**P1**	**P2***	**P3**	**P4**	**P5***	**P6***
Verb Phase Treatment.						
Sound and Letter Cueing Exercises.	100%	**92%**	100%	100%	100%	100%
Sentence Closure Exercises.	100%	**100%**	n/a	n/a	n/a	**100%**
Verb Networks A Exercises.	**92%**	n/a	100%	**92%**	100%	n/a
Verb Generation Exercises.	100%	100%	100%	92%	**100%**	100%
Sentence Phase Treatment						
Simple Statements, Requests and Yes/No Questions	100%	100%	100%	100%	**100%**	98%
Light Verb Exercises	100%	100%	100%	**98%**	n/a	100%
Verb Networks B Exercises.	100%	100%	100%	n/a	100%	**100%**
Longer Statements and Wh- Questions.	100%	100%	**94%**	100%	100%	n/a
Generalisation Phase Treatment.						
Personal and News Stories.	100%	100%	**100%**	100%	100%	100%
Scripts.	**100%**	100%	n/a	n/a	n/a	100%
Story and Video Retell	**100%**	100%	100%	100%	n/a	100%

5.8 Summary.

In terms of the feasibility of the SPT, six participants were successfully recruited to the study. Preliminary findings are that the SPT is also technically feasible as 5/6 participants were able to self-deliver it independently, with prior experience with technology a likely influential factor. Only three participants were entirely independent in *downloading* the SPT exercises however, with two participants dependent on the researcher to do this, whilst one participant was assisted by his wife. The SPT was used successfully on a variety of devices. The treatment was found to be acceptable to participants with very few negative aspects reported. Four participants complied with the weekly target amount of self-delivered treatment, with prior experience of technology again a likely influential factor, together with physical health. In terms of treatment fidelity, individual face-to-face sessions were conducted with good fidelity. However, there was wide variation in the length of time taken to deliver treatment and in the amount of self-delivered treatment as a result of multiple factors.

Chapter 6. Results for research question 5: preliminary efficacy testing.

The results relating to the preliminary efficacy testing of the SPT are reported below. Specifically, for each participant, the impact of treatment on the production of i) trained and untrained verbs, ii) untrained nouns, iii) sentence production using trained and untrained verbs, iv) verb and sentence production in discourse and v) in functional communication as perceived by a) the participants with aphasia themselves and b) their SOs is reported. A descriptive summary of each participant's results is given first, followed by a detailed description of the effect of treatment at each level of communication (and this section includes statistical analysis).

6.1. The impact of the SPT treatment on the production of verbs in isolation, in sentences, in discourse and on functional communication for P1.

6.1.i. Summary of P1's Results.

P1 had a moderate non-fluent aphasia with moderately impaired noun and verb production. He had good comprehension of both spoken and written words, with moderately impaired comprehension of spoken sentences and a more severe impairment of written sentence comprehension. He had good word repetition and was mildly impaired in terms of reading aloud (see Chapter 5.1.i for a fuller summary). P1 had eight face-to-face treatment sessions over 11 weeks and self-managed a total of 11 hours of treatment during this period. P1 did not continue to practise between T3 and T4 so this represents a true maintenance period for him.

P1's production of *treated* PR verbs *in isolation* improved at T3 (which was maintained), and production of *untreated* PR verbs improved at T4 (within level generalisation). There was no evidence of generalisation to untreated personally *irrelevant* verbs (on the OANB). Production of PR treated and untreated verbs *in sentences* improved at T3 which was maintained, with additional gains in production of objects. There was some evidence of change in the production of untreated personally irrelevant verbs and objects in sentences (on the SCT) at maintenance only. In terms of *syntax (PAS),* there was an improvement in the production of verbs with two arguments for treated and untreated PR verbs, with a trend (non-significant) for untreated personally irrelevant words. In terms of discourse production, there was no significant treatment effect on either the CAT Picture Description Task or the AMI, although P1's score on the Picture Description Task improved at T4 (driven by improvements in syntactic variety and grammatical well-formedness). There was a substantial negative change in perceived functional communication (CETI) at T3 which returned to baseline at T4 and this likely reflects P1's lack of engagement when completing the CETI questionnaire which he disliked.

6.1.ii. The effect of the SPT on P1's production of treated and untreated PR verbs in isolation.

The impact of treatment on the production of treated, untreated and the total set of verbs in isolation for P1 is reported in Table 24 below. P1 showed a stable performance[22] regarding verb production in isolation prior to treatment. Therefore, post-treatment comparisons are with T2. P1 showed a statistically significant improvement in *treated* verb production at both T3 (p=0.000) and T4 (p=0.0327) and a significant improvement in *untreated* verbs (within level generalisation) at T4 only (p=0.0352). The improvement in untreated verbs included improvement in verbs which were semantically related to treated verbs and those which were not. Thus, football related verbs were the most common verb topic for P1 (comprising 18/40 of his PR verbs) but improvement in untreated verbs was not restricted to these verbs: 4/12 untreated football related verbs which were not produced correctly at T2 were correct at T3 and/or T4, as were 3/8 non-football related verbs). There was a significant improvement in the production of the *total* set of verbs at T3

[22] Stability of performance was assessed by applying a McNemar Test to scores attained at baseline T1 and T2. If performance was stable, T2 was used as the pre-post treatment point of comparison. If performance was unstable, the higher score from T1 and T2 was used.

(p=0.0002) and this remained significant at T4 (p=0.0022) and this was driven by the improvement in the treated verbs. (P1 did not practise after the termination of treatment and so for him this represents a true maintenance phase).

*Table 24 P1's production of PR verbs in isolation pre and post treatment. (Significant change (McNemar Test): *=p<.05; ** =p<.01; ***p=<.001).*

		T1	T2	T3	T4
P1	**Treated**	7	7	**16****	**14***
	Untreated	8	6	9	**12***
	Total	15	13	**25*****	**26****

6.1.iii. The effect of the SPT on P1's production of untreated (personally irrelevant) verbs and on noun retrieval.

P1's scores on the OANB verb and nouns subsets are reported in Table 25 below. P1's production of personally irrelevant verbs was unstable at baseline. No other significant changes in production of untreated, personally irrelevant verbs or in noun production were found.

*Table 25 P1's on a subset of verbs and nouns from the OANB. (P1: verbs n=46; nouns n=50). Significant change compared to T2 (McNemar Test: *=p<.05).*

		T1	T2	T3	T4
P1	**OANB Verbs**	30	**23***	33	30
	OANB Nouns	36	37	38	39

6.1.iv. P1's sentence production using treated and untreated PR verbs and untreated personally irrelevant verbs.

Production of untreated and treated PR verbs, agents and objects in *sentences* for P1 is reported in Table 26 below. (Production of untreated personally *irrelevant* verbs in sentences is reported via the VAST SCT in the table below).

*Table 26 P1's production of verbs in sentences.(Significant change compared to T2 (McNemar Test): *=p<.05; ** =p<.01; ***p=<.001).*

Sentence Production	T1	T2	T3	T4
PR Sentences Treated (n=20)				
Agent	16	**11***	16	18
Verb	12	8	**16***	**16***
Object	7	4	**11****	**11***
PR Sentences Untreated (n=20)				
Agent	15	13	**18***	**19***
Verb	7	7	**16****	**14****
Object	8	5	**11***	**12****
PR Sentences Total (n=40)				
Agent	31	**24***	34	37
Verb	19	15	**32*****	**30*****
Object	15	9	**22*****	**23*****
VAST Sentence Construction Test (n=16)				
Agent	13	n/a	16	15
Verb	10	n/a	15	**15***
Object	3	n/a	6	**10***

P1 demonstrated an unstable performance at baseline regarding production of agents and so T1 was the point of comparison. There were significant improvements in the production of trained verbs and objects immediately post treatment (verbs: $p=0.0107$; objects: $p=0.0078$) and at maintenance (verbs: $p=0.0107$; objects: $p=0.0195$). Production of untrained verbs, agents and objects improved significantly both immediately post-treatment ((verbs: $p=0.0020$; agents: $p=0.0313$; objects $p=0.0352$) and at maintenance (verbs: $p=0.0078$; agents: $p=0.0352$; objects: $p=0.0078$). The significant improvement in trained and untrained verbs and objects resulted in significant improvements in the total set of verbs and objects both immediately post-treatment (verbs: p=0.0000; objects: p=0.0005) and at maintenance (verbs: p=0.0001; objects: p=0.0003). There was a significant improvement in the production of untrained personally irrelevant verbs and objects in the in objects at maintenance only ((verbs: p=0.0313; objects: p=0.0195).

Analysis of P1's PAS using PR trained and untrained PR verbs is reported in Table 27 below. The significance of change in the production of i) verbs alone and ii) of a verb with its arguments was analysed as these were the structures targeted in the Verb Phase and the Sentence Phase of the SCT respectively. There was a significant improvement in P1's production of two argument structures for both trained and untrained verbs both immediately post-treatment and at maintenance (trained verbs: T3, χ^2 = 5.3846, p-0.0203; T4, χ^2 = 6.8267, p=0.0090; untrained verbs: T3, χ^2 = 5.1042, p-0.0239; T4, χ^2 = 6.5473, p=0.0105), leading to a significant improvement in the total set of verbs T3 (χ^2 = 12.0000, p-0.0005; T4, χ^2 = 15.0521, p=0.0001).

*Table 27 Analysis of P1's PAS using PR treated, untreated and the total set of verbs. (UTS = undetermined thematic structure. Significant change compared to T2 (χ^2 Test): *=p<.05; ** =p<.01; ***p=<.001; italics = borderline significance. Verbs alone and a verb + 3 arguments were not analysed due to the small number of instances).*

	No output <10 seconds	**UTS**	**Verb**	**Verb + 1 argument**	**Verb + 2 arguments**	**Verb + 3 arguments**	**Verb + complex phrase**
	Treated PR Verbs						
T1	4	4	0	7	4	1	0
T2	5	7	0	5	3	0	0
T3	3	1	0	5	**11***	0	0
T4	1	2	0	5	**12****	0	0
	Untreated PR Verbs						
T1	4	5	0	2	9	0	0
T2	6	8	0	2	4	0	0
T3	2	3	0	3	**12***	0	0
T4	0	3	1	3	**13***	0	0
	Total PR Verbs						
T1	8	9	0	9	13	1	0
T2	11	15	0	7	7	0	0
T3	5	4	0	8	**23*****	0	0
T4	1	5	1	8	**25*****	0	0

There was also some evidence of improved PAS structure involving untreated personally irrelevant words (in the SCT), with the production of two argument structures again increasing post- treatment (NB: T2 recording failed) however this change was not significant (χ^2 = 2.216, p=0.137, see Table 28 below).

Table 28 Analysis of P1's PAS using personally irrelevant verbs in the Sentence Construction Test (Bastaanse et al., 2002).

	No output <10 seconds	**UTS**	**Verb**	**Verb + 1 argument**	**Verb + 2 arguments**	**Verb + 3 arguments**	**Verb + complex phrase**
SCT T1	2	3	0	8	3	0	0
SCT T2	n/a	n/a	n/a	n/a	n/a	n/a	n/a
SCT T3	0	1	0	8	7	0	0
SCT T4	1	0	0	6	8	0	1

The improvements in P1's PAS reflected increased *speed* of production. Thus, before treatment P1 was able to begin sentence production within 10 seconds for 75% of his PR verbs, whilst after treatment (at maintenance) he did so for all but one (of his 40) PR verbs (see Table 6.1.v. column 1). P1's faster sentence production also meant that he was able to produce *more* of a sentence within 10 seconds: before treatment P1 was often only able to produce the agent within 10 seconds or a sentence with an intransitive verb, whereas after treatment P1 showed significant improvement in production of a verb with 2 arguments within this time. P1's production of sentences using personally *irrelevant* verbs mirrored this pattern (of increased production of a verb with two arguments) with additional, tentative, evidence that the use of both longer and more complex verb phrases was starting to emerge (i.e., "the man is *relaxing and smoking pipe*" and "the boy is *crying because the dinner was awful").*

6.1.v. P1's discourse production: picture description.

The impact of the SPT on complex picture description was assessed using the CAT Spoken Picture Description task. The total CAT Picture Description score for P1 is reported in row 1 of Table 29 below. P1 demonstrates stable performance at T1 and T2, with a small deterioration at T3, and an improvement at T4 but this was not significant (Friedman's Test: Q= 3.00, p=0.3916). In order to elucidate the origin of any changes in the CAT picture description scores – specifically if this reflected improved verb retrieval, verb types and tokens produced were counted. These are reported for P1 in Table 6. 5.viii rows 2 and 3 below. This indicates no improvement in the number of verb types or tokens produced by P1 as a result of treatment.

Table 29 P1's scores on the CAT spoken picture description task.

P1	**T1**	**T2**	**T3**	**T4**
1. Total CAT Score	18	19	15.5	27
2. P1 Verb types	3	2	4	3
3. P1 Verb tokens	3	3	4	3
4. Appropriate ICWs	17	16	14	18
5. Inappropriate ICWs	4	2	3	0
6. Syntactic Variety	1	1	1	3
7. Grammatical Well-formedness	3	3	2.5	5
8. Speed	1	1	1	1

The CAT Spoken Picture Description scores were further scrutinized to locate the origin of any improvement i.e., the participants' scores on the individual CAT Spoken Picture Description Scoring criteria were compared and are reported in Table 29 rows 4 - 8. Perusal of these scores indicates that the improvement in the overall CAT picture description score for P1 was due to increased production of appropriate ICWs combined with a decrease in inappropriate ICWs, together with an increase in syntactic variety and grammatical well formedness.

The PAS of P1's CAT Spoken Picture Description indicates that P1's production of verbs with one argument structures improved immediately post treatment. (Table 30 below. Appendix E contains a transcription of P1's CAT Picture Description at all four time points to illustrate the process of coding, and for all six participants).

Table 30 The PAS of P1's CAT Spoken Picture Description.

	UTS	Verb	Verb + 1 argument	Verb + 2 arguments	Verb + 3 arguments	Verb + complex phrase
CAT Picture Description T1	8	0	5	2	0	0
CAT Picture Description T2	4	0	4	4	1	0
CAT Picture Description T3	5	0	8	0	0	0
CAT Picture Description T4	4	0	2	3	1	0

6.1.vi. P1's discourse production: The Autobiographical Memory Interview (AMI) (Kopelman et al., 1989).

Table 31 demonstrates no significant change in P1's AMIs but a potential assessment burden effect appears to be present as P1 produced fewer utterances at each successive time point (T1, 37 utterances, T2, 29 utterances and T4 24 utterances. The AMI was not administered at T3 as the assessment session was cancelled due to P1's hospitalisation).

Table 31 The PAS of P1's AMIs (Kopelman et al., 1989).

	T1	T2	T3	T4
UTS	14	13	n/a	12
Verb	0	1	n/a	4
Verb + 1 argument	3	5	n/a	0
Verb + 2 arguments	12	10	n/a	5
Verb + 3 arguments	5	0	n/a	3
Verb + complex phrase	3	0	n/a	0
Total Utterances	37	29	n/a	24

P1's AMIs were also scored using three of the criteria of the CAT Spoken Description Task to facilitate comparison with this task (Appropriate and Inappropriate Information Carrying Words could not be scored given that the AMI is a personal narrative). The scores for P1 are reported in Table 32 below and show no changes.

Table 32 CAT Spoken Picture Description scores for P1's AMIs (Kopelman et al., 1989).

P1	**T1**	**T2**	**T3**	**T4**
Syntactic Variety	3	2	n/a	2
Grammatical Well-formedness	3	3	n/a	3
Speed	1	1	n/a	1

6.1.vii. Impact of the SPT on P1's Functional Communication.

CETI scores for P1 are reported in Table 33 below.

P1 showed no significant change in his CETI scores on the Friedman test (Q =4.98, p=0.0827). However, at T3 his CETI rating is >5.2 SE below that at T1. This probably reflects P1's lack of engagement with the CETI at all testing points: he disliked completing it (the CETI at T2 is not reported as he did not complete all of it), and wanted this dislike to be recorded as part of his feedback in his exit interview. Thus, his scores on the CETI are unlikely to accurately reflect his functional communication.

Table 33 The impact of the SPT on functional communication for P1 (CETI scores (Lomas et al., 1989)*. (Significance level Friedman test: *=p.05; **p<.01; **p<.001; italics = significant using SE>5.2 change).*

CETI Score	T 1	T 2	T 3	T 4
P1	76.25	n/a	*65.438*	73.063

6.2 The impact of the SPT treatment on the production of verbs in isolation, in sentences, in discourse and on functional communication for P2.

6.2.i Summary of P2's Results.

P2 had a moderate non-fluent aphasia with moderately impaired noun and verb production. She had good comprehension of spoken and written nouns, and of spoken verbs. Her comprehension of spoken sentences was moderately impaired, with a more severe impairment of written sentence comprehension. She had very good ability to repeat and read aloud single words (see Chapter 5.1.ii for a fuller summary). P2 had eight face-to-face treatment sessions over 11 weeks and self-delivered a total of 59 hours of treatment during this period.

In interpreting the results of treatment for P2 it should be borne in mind that P2 continued to practise treated verbs using the SPT between T3 and T4, and therefore performance at T4 cannot be regarded as demonstrating maintenance of treatment effects. P2's production of treated PR verbs *in isolation* showed significant improvement at T3 and at T4, whilst production of untreated PR verbs improved significantly at T4 only (within level generalisation). There was no generalisation to untreated personally irrelevant verbs or to noun production (on the OANB). Production of PR treated verbs *in sentences* improved significantly post treatment as did the production of their related agents and objects. There was no significant change in the production of untreated PR verbs in sentences, or of personally irrelevant verbs in sentences. In terms of PAS, there was a significant improvement in the production of verbs with two arguments post treatment which was restricted to treated PR verbs, and this was accompanied by a decrease (trend) in the production of treated PR verbs alone. There were no changes in the PAS of untreated personally irrelevant verbs (on the SCT). In terms of discourse production, there was an improvement in P2's score on the CAT Picture

Description Task post-treatment which was not significant. The improvement in her score was driven by increased production of appropriate ICWs combined with an increase in syntactic variety and grammatical well-formedness. There was also an improvement in the production of verbs with one argument and with two arguments. There were no changes in P2's narrative production. In terms of functional communication P2 rated her communication more positively at T3 but this perceived improvement fell away at T4, whilst her SO rated her communication more positively at both T3 and T4.

6.2.ii. The effect of the SPT on P2's production of treated and untreated PR verbs in isolation.

The impact of treatment on the production of treated, untreated and the total set of verbs in isolation for P2 is reported in Table 34 below.

Table 34 P2's production of verbs in isolation pre and post treatment. *(Significant change (Wilcoxon Two Sample Test): *=p<.05; ** =p<.01; ***p=<.001).*

		T1	T2	T3	T4
P2	**Treated**	7	7	**20*** **	**19*** **
	Untreated	7	7	11	**16** **
	Total	14	14	**31*** **	**35** **

P2 showed a stable performance regarding verb production in isolation prior to treatment. Therefore, post-treatment comparisons are with T2. P2 showed a significant improvement in *treated* verb production at both T3 (p = 0.0001) and T4 (p = 0.0002), and a significant improvement in *untreated* verbs (within level generalisation) at T4 only (p = 0.0020). Regarding generalisation to untreated verbs, this reflected improvement in both semantically related and unrelated verbs. Thus, the most common topic of P2's PR verbs was *feelings and senses,* representing 12 of her 40 verbs, with six each in the treated and untreated verb sets. Of the nine untreated verbs which improved post-treatment, four of these related to feelings and senses and five to other verb topics. P2 showed a significant improvement in production of the total set of verbs at both T3 (p = 0.0002) and T4 (p = 0.0000) which was largely as a result of the improvement in treated verbs.

6.2.iii. The effect of the SPT on P2's production of untreated (personally irrelevant) verbs and on noun retrieval.

P2's scores on the OANB verb and nouns subsets are reported in Table 35 below. P2 demonstrated no significant change in production of untreated, personally irrelevant verbs or of nouns post treatment.

Table 35 P2's scores on a subset of verbs and nouns from the OANB*. (P2: verbs n=49; nouns n= 50).*

		T1	T2	T3	T4
P2	**OANB Verbs**	34	37	37	40
	OANB Nouns	39	43	45	47

6.2.iv. P2's sentence production using treated and untreated PR verbs and untreated personally irrelevant verbs.

Production of untreated and treated PR verbs, agents and objects in *sentences* for P2 is reported in Table 36 below. (Results for T3 are not reported as the recording of this assessment failed). P2 demonstrated a stable performance at baseline in terms of the production of PR verbs, agents and objects, therefore comparison post treatment is with scores at T2. Significant improvement in the production of treated PR verbs in sentences was observed post treatment (T4: verbs: p=0.0032), together with significant improvements in the production of agents and objects (T4: agents: p=0.0107; objects: p=0.0039). There were no significant improvements in the production of untreated PR verbs, agents or objects, but a significant improvement in the total set of agents, verbs and objects T4: verbs: p = 0.0032; agents: p=0.0481; objects: p=0.0021) which was driven by the improvement in the treated set. Change in the production of personally irrelevant verbs in sentences was measured via the SCT from the VAST. Performance was not stable prior to treatment with improvement occurring between T1 and T2. Nevertheless, T2 was used at the point of comparison with no change evident in the production of personally irrelevant verbs or their agents and objects.

Table 36 P2's production of verbs in sentences. *(Significant change compared to T2 (McNemar Test): *=p<.05; ** =p<.01; ***p=<.001).*

		T1	T2	T3	T4
PR Sentences Treated (n=20)					
	Agent	11	11	n/a	**19***
	Verb	11	7	n/a	**17****
	Object	6	4	n/a	**12****
PR Sentences Untreated (n=20)					
	Agent	12	15	n/a	15
	Verb	14	12	n/a	13
	Object	8	5	n/a	9
PR Sentences Total (n=40)					
	Agent	23	26	n/a	**34***
	Verb	25	19	n/a	**30****
	Object	14	9	n/a	**21****
Sentence Construction Test (n=20)					
	Agent	8	**20*****	n/a	18
	Verb	6	**16****	n/a	16
	Theme	4	8	n/a	4

Analysis of P2’s PAS using PR treated and untreated verbs is reported in Table 37 below. There was a significant improvement in P2’s production of two argument structures for treated verbs at T4 only (p = 0.0428) driving a significant improvement for the total set of verbs at T4 (p = 0.0317). (Production of verbs on their own was not analysed because of the zero vales at T4). The improvement in PAS reflected quicker production of sentences resulting in fewer instances of P2 being unable to produce any output within 10 seconds, and also enabling her to produce a sentence with a verb with two arguments (as opposed to an agent only or an agent with a verb) within that time limit.

Table 37 Analysis of P2's PAS using PR treated, untreated and the total set of verbs. *(UTS = undetermined thematic structure. Significant change compared to T2 (χ^2 Test): *=p<.05. Verbs alone and a verb + 3 arguments were not analysed due to the small number of instances).*

	No output <10 seconds	**UTS**	**Verb**	**Verb + 1 argument**	**Verb + 2 arguments**	**Verb + 3 arguments**	**Verb + complex phrase**
	Treated PR Verbs						
T1	3	3	3	5	6	0	0
T2	6	5	1	5	3	0	0
T3	n/a	n/a	n/a	n/a	n/a	n/a	n/a
T4	0	3	0	6	**10***	1	0
	Untreated PR Verbs						
T1	1	3	4	5	7	0	0
T2	2	6	3	4	5	0	0
T3	n/a	n/a	n/a	n/a	n/a	n/a	n/a
T4	1	5	0	6	8	0	0
	Total PR Verbs						
T1	4	6	7	10	13	0	0
T2	8	11	4	9	8	0	0
T3	n/a	n/a	n/a	n/a	n/a	n/a	n/a
T4	1	8	0	12	**18***	1	0

PAS using untreated personally irrelevant verbs was measured by the SCT. There was unstable performance at baseline which was significant for verbs with one argument. Post treatment performance was compared to T2 with no evidence of improved PAS structure as a result of treatment (see Table 38 below).

Table 38 Analysis of P2's PAS using personally irrelevant untreated verbs in the Sentence Construction Test (Bastiaanse et al., 2002). *(UTS = undetermined thematic structure. Significant change compared to T2 (χ^2 Test): *=p<.05. Verbs alone and a verb + 3 arguments were not analysed due to the small number of instances).*

	No output <10 seconds	UTS	Verb	Verb + 1 argument	Verb + 2 arguments	Verb + 3 arguments	Verb + complex phrase
SCT T1	12	2	0	2	3	1	0
SCT T2	0	3	0	**9***	8	0	0
SCT T3	n/a	n/a	n/a	n/a	n/a	n/a	n/a
SCT T4	2	4	0	9	5	0	0

6.2.v. P2's discourse production: picture description.

The total CAT Picture Description score for P2 is reported in row 1 of Table 39 below. P2 demonstrates stable performance at T1 and T2, with an increased score at T3 and T4 which was not significant (Friedman's Test: Q= 5.58, p=0.1337). In order to elucidate the origin of any changes in the CAT picture description scores – specifically if this reflected improved verb retrieval, the verb types and tokens produced were counted. These are reported for P2 in Table 6.2.vi rows 2 and 3 below. This indicates no improvement in the number of verb types produced by P2 as a result of treatment, with a small increase in the number of tokens (at T4). The CAT Spoken Picture Description scores were further scrutinised to locate the origin of any improvement. Perusal of these scores indicates that the improvement in the overall CAT picture description score for P2 was due to increased production of appropriate ICWs combined with an increase in syntactic variety and grammatical well formedness. To elaborate, post treatment, P2 made fewer grammatical errors in that she omitted less determiners, used correct prepositions and made fewer errors with verb inflections. P2's picture description also showed evidence of increased phrasal complexity. For example, at T2 she said "T*he man is sleeping*" whilst at T4 she said, "*man is snoozing and resting his weary legs on the the books*".

Another example is that at T1 she said, “*the the toddler he notice*” whilst at T4 she said, “*the toddler is trying to help no point out the danger with the the books*.”

Table 39 P2's scores on the CAT spoken picture description task.

P2	T1	T2	T3	T4
Total CAT Score	26.5	27	29	36
P2 Verb types	10	8	6	10
P2 Verb tokens	12	13	11	16
Appropriate ICWs	20	25	20	29
Inappropriate ICWs*	1	5	0	3
Syntactic Variety	2	2	3	4
Grammatical Well-formedness	4	4	5	5
Speed	1	1	1	1

The PAS of P2’s CAT Spoken Picture Description is reported in Table 40 below. This indicates that P2’s production of verbs together with an argument improved as a result of treatment: she produced an average of 10 utterances containing a verb together with an argument at T1/T2, and this increased to 19 utterances at T3, falling to 15 utterances at T4. More specifically verbs with one argument improved at T3 but returned to baseline at T4. Production of verbs with two arguments also improved slightly at T3 only, with tentative indications that production of verbs in longer and more complex phrases was emerging at T4.

Table 40 The PAS of P2’s CAT Spoken Picture Description.

	UTS	Verb	Verb + 1 argument	Verb + 2 arguments	Verb + 3 arguments	Verb + complex phrase
CAT Picture Description T1	1	1	4	4	0	0
CAT Picture Description T2	5	1	5	6	1	0
CAT Picture Description T3	2	0	11	8	0	0
CAT Picture Description T4	5	0	5	6	3	1

6.2.vi. P2's discourse production: The Autobiographical Memory Interview (AMI) (Kopelman et al., 1989).

Table 41 reports the PAS of P2's AMIs. This indicates little change in P2's discourse production with the exception of a reduction in utterances with an undetermined thematic structure (UTSs). An assessment burden effect is likely present as P2 produced fewer utterances at T3 and T4. Indeed, P2 commented during the AMI at T3 that it was "too much really" and the research student terminated the AMI at that point.

Table 41 The PAS of P2's AMIs (Kopelman et al., 1989).

	T1	T2	T3	T4
UTS	26	18	11	6
Verb	0	7	0	0
Verb + 1 argument	10	8	5	12
Verb + 2 arguments	18	19	17	21
Verb + 3 arguments	5	1	5	4
Verb + complex phrase	2	1	0	0
Total Utterances	61	54	38	43

P2's AMIs were also scored using three of the criteria of the CAT Spoken Description Task to facilitate comparison with this task. The scores for P2 are reported in Table 6.2.ix below and show an increase in scores for grammatical well-formedness reflecting more accurate use of verb inflections, prepositions and determiners.

Table 42 CAT Spoken Picture description scores for P2's AMIs (Kopelman et al., 1989).

	T1	T2	T3	T4
Syntactic Variety	4	4	3	4
Grammatical Well-formedness	1	1	3	3
Speed	1	1	1	1

6.2.vii. Impact of the SPT on P2's Functional Communication.

CETI scores for P2 and her SO are reported in Table 43 below. P2 showed an improvement in her CETI scores (>5.2 retest standard error of the mean) at baseline assessments (T1 and T2) as did the CETI scores of her SO. T2 was therefore used as the point of comparison for post treatment performance. There was no significant improvement in P2's CETI score between T2 and T4 according to Friedman's test (Q =2.89, p=0.2359) scores but the improvement was close to 5.2 retest standard error of the mean (5.0 at T3). The CETI rating of P2's SO improved significantly on the Friedman test (Q = 11.48, p=0.0032) post treatment, and this was also >5.2 retest standard error of the mean (at T3). The CETI ratings of both P2 and her SO decreased by >5.2 retest standard error of the mean at T4.

Table 43 The impact of the SPT on functional communication for P2 and her SO (CETI scores (Lomas et al., 1989)*. (Italicised scores show change >5.2 retest standard error of the mean).*

CETI Score	T 1	T 2	T 3	T 4
P2	38.25	*50.25*	*55.25*	*39.375*
P2 SO	29.063	*35.375*	*49.063*	*39.375*

6.3. The impact of the SPT treatment on the production of verbs in isolation, in sentences, in discourse and on functional communication for P3.

6.3.i. Summary of P3's Results.

P3 had a moderate – severe non-fluent aphasia together with a mild apraxia of speech. P3 had good understanding of spoken and written single words but had difficulty understanding both spoken and written sentences reversible sentences, and this was worse for written sentences. P3's ability to repeat words was affected by the number of syllables, but she had very good ability to read words aloud (see Chapter 5.1.iii for a fuller summary). P3 had eight face-to-face treatment sessions over nine weeks and self-managed a total of 18 hours of treatment during this period. P3 did not continue to practise between T3 and T4 so this represents a true maintenance period for her.

Following treatment P3's production of treated PR verbs in isolation showed improvement, which was significant at T3 and maintained at T4. There was no change in the production of untreated PR verbs. There was no generalization to untreated personally irrelevant verbs, but nouns improved significantly (on the OANB) immediately post treatment and at maintenance. Production of PR treated verbs in sentences improved significantly post treatment as did the production of the related agents, and this was maintained for both verbs and agents. There was a significant change in the production of untreated PR verbs and agents in sentences which was present at maintenance only. There was a significant improvement in the production of the total set of verbs and agents post-treatment, which was maintained, and a small but significant improvement in the production of the total set of objects evident at maintenance only. There was no change in the production of personally irrelevant verbs in sentences. In terms of PAS, there was a significant improvement in the production of treated PR verbs with one argument post treatment which was maintained. There was no improvement in PAS for untreated personally irrelevant verbs (on the SCT). In terms of discourse production, there was a significant improvement in P3's score on the CAT Picture Description Task post treatment. The improvement in her score was largely driven by increased production of appropriate ICWs combined with an increase in grammatical well-formedness. There was an increase in P3's production of verbs with two arguments in both picture description and personal narrative. In terms of functional communication both P3 and her SO rated her communication more positively post treatment and this was maintained.

6.3.ii. The effect of the SPT on P3's production of treated and untreated PR verbs in isolation.

The impact of treatment on the production of treated, untreated and the total set of verbs in isolation for P3 is reported in Table 44 below. P3 showed a stable performance regarding verb production in isolation prior to treatment for the treated and untreated sets of verbs. Therefore, post-treatment comparisons are with T2. For the total set of verbs, baseline performance was unstable therefore the point of comparison is the higher score at T1. P3 showed a statistically significant improvement in *treated* verb production at T3 (p = 0.0039) and at T4 (p = 0.0156). There was no improvement in *untreated* verbs at T3 or T4.

Table 44 P3's production of verbs in isolation pre and post treatment. *(Significant change (McNemar Test): *=p<.05; ** =p<.01).*

		T1	T2	T3	T4
P3	Treated	12	10	**18****	**16***
	Untreated	13	9	11	12
	Total	25	19*	29	28

6.3.iii. The effect of the SPT on P3's production of untreated (personally irrelevant) verbs and on noun retrieval.

P3's scores on the OANB verb and noun subsets are reported in Table 45 below. P3 demonstrated no significant change in production of untreated, personally irrelevant verbs post treatment, but there was a significant change in noun production at T3 (p = 0.0039) and T4 (p = 0.0195) even though none of the OANB nouns had been practised in treatment.

*Table 45 P3's scores on a subset of verbs and nouns from the OANB. (P3: verbs n=38; nouns n= 42). Significant change compared to T2 (Wilcoxon Two Sample Test: *=p<.05 **p=<0.01)*

		T1	T2	T3	T4
P3	**OANB Verbs**	28	33	36	37
	OANB Nouns	37	33	**41****	**40***

6.3.iv. P3's sentence production using treated and untreated PR verbs and untreated personally irrelevant verbs.

Production of untreated and treated PR verbs, agents and objects in *sentences* for P3 is reported in Table 46 below. Significant improvement in the production of treated PR verbs in sentences was observed post treatment at both T3 (p=0.0005) and T4 (p=0.0059), together with significant improvements in the production of treated agents (T3: p=0.0002; T4: p=0.0032). There was an improvement (trend) in the production of treated objects which had been at floor pre-treatment. There was a significant improvement in the production of *untreated* PR verbs and agents at T4 (maintenance) only (verbs: p=0.0107; agents: p=0.0010), with again a trend of improved production of objects.

Table 46 P3's production of verbs in sentences. *(Significant change compared to T2 (McNemar Test): *=p<.05; ** =p<.01; ***p=<.001).*

		T1	T2	T3	T4
PR Sentences Treated (n=20)					
Treated	Agent	5	4	**16*** **	**14** **
	Verb	9	5	**16*** **	**14** **
	Object	0	0	2	3
PR Sentences Untreated (n=20)					
Untreated	Agent	6	4	7	**14*** **
	Verb	7	6	8	**14* **
	Object	0	0	0	3
PR Sentences Total (n=40)					
Total	Agent	11	8	**23*** **	**28*** **
	Verb	16	11	**24** **	**28*** **
	Object	0	0	2	**6* **
Sentence Construction Test (n=18)					
	Agent	9	**16** **	n/a	13
	Verb	8	11	n/a	10
	Object	0	0	n/a	0

If P3's production of objects *after* the time limit of 10 seconds is considered, there was a more notable increase in production of objects: they were only produced on one occasion at T1 and two at T2 increasing to eight at T3 and nine at T4 (with the second argument almost exclusively being a theme as was targeted during treatment). There was no evidence of improved retrieval of untrained personally *irrelevant* verbs or agents in sentences (on the SCT). However, similarly to sentences using PR verbs, the production of objects after the time limit of 10 seconds improved post treatment: from two at T1 and zero at T2 to seven at T4 (T3 was not assessed). Examples of sentences which P3 produced exemplifying this improvement are reported in Table 47 below.

Table 47 P3's production of sentences including utterances >10 seconds (denoted by ||).

Target sentence PR verbs	P3's sentence: pre treatment	P3's sentence: post treatment
the players are arguing	|| they are arguing	The players are arguing || the players are arguing with their manager
the couple are buying a house	the couple ||is buy a house	the couple are buying a house
the children are watching TV	the children are listening|| no watching	they are watching a film
the hairdresser is cutting hair	the cutting || she is cutting her hair	the woman is having || er having her hair cut
the woman is listening to music	the woman is listening	the girl is listen to music
the woman is ordering food	the er the ||order the the she is ordering	she is ordering || a meal
the woman is paying for coffee	the woman is paying	the woman is paying for the drinks
the man is phoning	dialling	the man is phoning || his wife
the children are visiting grandad	|| visit um the children are visiting	the girls are visiting their || grandpa
Target sentence SCT verbs	**P3's sentence: pre treatment**	**P3's sentence: post treatment**
the man is drinking a glass of wine	the man is drinking	the man is drinking || a bottle of wine
the boy is hitting the girl	the man is || the man is /hin/ slapping her	the the || woman is being slapped um er by her husband
the man is smoking a pipe	the man is smoking	the man is blowing || smoke rings out of his pipe
the dog is biting the cat	the dog is chasing || the fast no the cat	the dog is || the cat is the cat is chasing the god the er the dog is chasing the cat
the man is painting the woman	the man is painting	the man is painting the woman
the clown is smiling	|| the clown is the clown is standing	|| the man is dressed as a clown
the child is scratching the man	the || er erm the er er /to/ toddler is scratching	the child is scratching his || er father

Analysis of P3's *PAS* using PR treated and untreated verbs is reported in Table 48 below. The significance of change in the production of verbs alone and verbs with two or more arguments was not analysed statistically due to the small numbers produced. However, production of trained PR verbs *alone* showed a trend of *reduced* production following treatment, accompanied by a significant *increase* in P3's production of *one argument* structures using trained verbs, with a trend of improved production of untrained verbs with one argument. This drove a significant improvement in production of this structure for the total set of verbs.

Table 48 Analysis of P3's PAS using PR treated, untreated and the total set of verbs. *(UTS = undetermined thematic structure. Significant change compared to T2 (χ^2 Test): *=p<.05; ** =p<.01. Verbs alone and a verb + 3 arguments were not analysed due to the small number of instances).*

	No output <10 seconds	UTS	Verb	Verb + 1 argument	Verb + 2 arguments	Verb + 3 arguments	Verb + complex phrase
	Treated PR Verbs						
T1	5	2	5	8	0	0	0
T2	12	2	2	4	0	0	0
T3	4	0	0	**14****	2	0	0
T4	4	2	0	**11***	3	0	0
	Untreated PR Verbs						
T1	10	4	1	5	0	0	0
T2	13	2	0	5	0	0	0
T3	6	5	2	7	0	0	0
T4	4	2	0	11	3	0	0
	Total PR Verbs						
T1	15	6	6	13	0	0	0
T2	25	4	2	9	0	0	0
T3	10	5	2	**21***	2	0	0
T4	8	4	0	**22****	6	0	0

PAS using untreated personally *irrelevant* verbs was measured by the SCT. There was no evidence of improved PAS structure as a result of treatment (see Table 49 below).

Table 49 Analysis of P3's PAS using personally irrelevant untreated verbs in the Sentence Construction Test (Bastiaanse et al., 2002). *(UTS = undetermined thematic structure. Verbs alone and a verb + 3 arguments were not analysed due to the small number of instances).*

	No output <10 seconds	UTS	Verb	Verb + 1 argument	Verb + 2 arguments	Verb + 3 arguments	Verb + complex phrase
SCT T1	6	5	0	8	0	0	0
SCT T2	2	3	0	12	0	0	0
SCT T3	n/a	n/a	n/a	n/a	n/a	n/a	n/a
SCT T4	4	3	0	10	1	0	0

6.3.v. P3’s discourse production: picture description.

The total CAT Picture Description score for P3 is reported in row 1 of Table 50 below. P3 demonstrates stable performance at T1 and T2, with an increased score at T3 and T4 which was significant (Friedman’s Test: Q= 12.22, p=0.0067). In order to elucidate the origin of any changes in the CAT picture description scores – specifically if this reflected improved verb retrieval, the verb types and tokens produced were counted. These are reported for P3 in Table 6. 3.viii rows 2 and 3 below. This indicates a slight improvement in the number of verb types and tokens produced at T4. The CAT Spoken Picture Description scores were further scrutinised to locate any other improvements and are reported in Table 6.7.viii rows 4 - 8. Improvement in the overall CAT picture description score for P3 was also related to increased production of appropriate ICWs combined with an increase in syntactic variety, grammatical well-formedness and speed.

Table 50 P3's scores on the CAT spoken picture description task.

P3	**T1**	**T2**	**T3**	**T4**
Total CAT Score	16	14	28	28.5
P3 Verb types	4	5	5	7
P3 Verb tokens	4	6	5	9
Appropriate ICWs	13	11	20	20
Inappropriate ICWs*	0	1	1	1
Syntactic Variety	1	1	2	2
Grammatical Well-formedness	2	3	6	6
Speed	0	0	1	1.5

The PAS of P3’s CAT Spoken Picture Description was also analysed and the results of this are reported in Table 51 below. This indicates that P3’s production of verbs together with an argument improved post treatment: at T1 and T2 she produced four utterances which included a verb with an argument, this rose to 10 at T3 and 12 at T4. Specifically, verbs with two arguments improved at T3 which was maintained at T4, with tentative indications that production of verbs with three arguments was emerging post-treatment.

Table 51 The PAS of P3's CAT Spoken Picture Description.

	UTS	Verb	Verb + 1 argument	Verb + 2 arguments	Verb + 3 arguments	Verb + complex phrase
CAT Picture Description T1	5	2	2	2	0	0
CAT Picture Description T2	1	2	3	1	0	0
CAT Picture Description T3	1	0	3	6	1	0
CAT Picture Description T4	0	0	4	7	1	0

6.3.vi. P3's discourse production: The Autobiographical Memory Interview (AMI) (Kopelman et al., 1989).

Table 52 reports the PAS of P3's AMIs. This indicates that there was an increase in P3's production of verbs with two arguments together with the emergence of verbs in complex sentences. (The AMI was not completed at T3 due to a medical emergency).

Table 52 The PAS of P3's AMIs (Kopelman et al., 1989).

	T1	T2	T3	T4
UTS	1	1	n/a	1
Verb	0	0	n/a	0
Verb + 1 argument	2	1	n/a	1
Verb + 2 arguments	6	10	n/a	13
Verb + 3 arguments	2	2	n/a	1
Verb + complex phrase	0	0	n/a	2
Total Utterances	11	14	n/a	18

P3's AMIs were also scored using three of the criteria of the CAT Spoken Description Task to facilitate comparison with this task. The scores for P3 are reported in Table 53 below and show small increases in scores for syntactic variety, grammatical well-formedness and speed.

Table 53 CAT Spoken Picture description scores for P3's AMIs (Kopelman et al., 1989).

P3	T1	T2	T3	T4
Syntactic Variety	2	2	n/a	3
Grammatical Well-formedness	4	4	n/a	5
Speed	0	0	n/a	0.5

6.3.vii. Impact of the SPT on P3's Functional Communication.

CETI scores for P3 and her SO are reported in Table 54 below. P3 showed a significant improvement in her CETI score post treatment (Friedman's Test, Q= 15.10, p=0.0005) and this change was also >5.2 retest standard error of the mean both immediately post-treatment and at maintenance. P3's SO also scored her significantly higher on the CETI post treatment (Friedman's Test, Q= 22.57, p=0.0000) and again this was >5.2 retest standard error of the mean both immediately post-treatment and at maintenance.

Table 54 The impact of the SPT on functional communication for P3 and her SO (CETI scores (Lomas et al., 1989). *(Italicised scores show change >5.2 retest standard error of the mean).*

CETI Score	T 1	T 2	T 3	T 4
P3	55.125	52.9375	*82.9375*	*76.5*
P3 SO	49.125	*n/a*	*61.25*	*69.688*

6.4. The impact of the SPT treatment on the production of verbs in isolation, in sentences, in discourse and on functional communication for P4.

6.4.i. Summary of P4's Results.

P4 had a moderate - severe fluent aphasia. He had significant impairments at multiple language levels: spoken and written word comprehension, spoken and written sentence comprehension, noun and verb naming, repetition, and writing. P4 had six face-to-face treatment sessions over 8 weeks (missing out on two sessions because he was visiting his terminally ill mother) and self-delivered a total of just under 29 hours of treatment during this period. P4 did not continue to practise between T3 and T4 so this represents a true maintenance period for him.

P4's production of treated PR verbs in isolation showed a non-significant trend of improvement immediately post-treatment only. There was no improvement in his production of untreated verbs or the total set of verbs. P4 demonstrated a trend of improvement in production of untreated, personally irrelevant nouns immediately post-treatment which was not statistically significant, and no improvement in verbs. In terms of sentence production, there was a trend of improvement in the production of treated agents both immediately post-treatment and at maintenance, and a significant improvement in the production of agents for the total set of verbs at maintenance. This pattern was mirrored in the production of untreated personally irrelevant agents with a significant change at maintenance. These were the only significant changes in the production of verbs, agents and objects in sentences. Regarding PAS, for the total set of PR verbs there was a trend of improvement in the number of verbs with one argument at maintenance, accompanied by an increase in the production of verbs with two arguments, neither of which was significant. By way of contrast PAS structure involving untreated personally irrelevant words showed a significant increase in the production of one argument structures at both T3 and T4 which was related to the significant increase in the production of agents. In terms of discourse production, in picture description P4's production of verbs with more than one argument showed improvement both immediately post-treatment and at maintenance. P4's production of a personal narrative demonstrated an increase in the relative proportion of utterances with a verb with one argument and with two arguments post treatment, together with a decrease in utterances with an undetermined thematic structure. P4 did not perceive his functional communication to have changed post treatment.

6.4.ii. The effect of the SPT on P4's production of treated and untreated PR verbs in isolation.

The impact of treatment on the production of PR treated, untreated and the total set of verbs in isolation for P4 is reported in Table 55 below. P4 showed significant improvement in treated verb production at T3 only (p= 0.0156).

Table 55 P4's production of verbs in isolation pre and post treatment. *(Significant change (McNemar Test): *=p<.05; ** =p<.01; ***p=<.001).*

		T1	T2	T3	T4
P4	Treated	6	4	**10***	6
	Untreated	7	3	2	5
	Total	13	7	12	11

6.4.iii. The effect of the SPT on P4's production of untreated (personally irrelevant) verbs and on noun retrieval.

P4's scores on the OANB verb and nouns subsets are reported in Table 56 below. P4 demonstrated a numerical improvement in production of untreated, personally irrelevant nouns at T3 which was not significant.

Table 56 P4's scores on a subset of verbs and nouns from the OANB. *(P4: verbs n=37; nouns n=42).*

		T1	T2	T3	T4
P1	OANB Verbs	11	6	11	11
	OANB Nouns	7	9	13	10

6.4.iv. P4's sentence production using treated and untreated PR verbs and untreated personally irrelevant verbs.

Production of untreated and treated PR verbs, agents and objects in *sentences* for P4 is reported in Table 57 below. (Production of untreated personally *irrelevant* verbs in sentences is reported via the VAST SCT in the table below). P4 demonstrated a stable performance at baseline in terms of the production of verbs, agents and objects. There was a (non-significant) trend of improvement in the production of treated and untreated agents at both T3 and T4 which resulted in a significant improvement in the production of agents for the total set of verbs at T4 (p=0.0064). This was the only significant change in the production of PR verbs, agents and objects in sentences. This pattern was mirrored in the production of untreated personally irrelevant verbs, agents and objects (in the SCT). There was unstable performance at baseline (poorer performance at T2) therefore, post treatment comparisons are with T1 (the higher baseline score). There were no significant changes in production of personally irrelevant verbs, agents, or objects.

Table 57 P4's production of verbs in sentences. *(Significant change compared to T2 (McNemar Test): *=p<.05; **=p<.01).*

Sentence Production	T1	T2	T3	T4
PR Sentences Treated (n=20)				
Agent	9	7	13	13
Verb	11	11	7	13
Object	6	7	7	8
PR Sentences Untreated (n=20)				
Agent	7	6	6	11
Verb	11	10	10	11
Object	3	6	9	7
PR Sentences Total (n=40)				
Agent	16	13	19	**24****
Verb	22	21	17	24
Object	9	13	16	15

VAST Sentence Construction Test (n=19)				
Agent	13	7	16	18
Verb	15	**10***	15	12
Object	9	**4***	4	9

Analysis of P4's PAS using PR treated and untreated PR verbs is reported in Table 58 below. For treated verbs there was a trend of decreased production of verbs with one argument at T4 which was not significant, accompanied by an increase in the production of verbs with two arguments which was also not significant. This improvement in treated verbs drove a similar pattern of improvement for the total set of verbs.

Table 58 Analysis of P4's PAS using PR treated, untreated and the total set of verbs. *(UTS = undetermined thematic structure).*

	No output <10 seconds	UTS	Verb	Verb + 1 argument	Verb + 2 arguments	Verb + 3 arguments	Verb + complex phrase
	Treated PR Verbs						
T1	1	8	2	6	3	0	0
T2	0	6	5	6	3	0	0
T3	1	8	1	5	5	0	0
T4	0	10	2	2	5	0	1
	Untreated PR Verbs						
T1	2	9	2	6	1	0	0
T2	0	10	5	5	0	0	0
T3	4	7	0	7	2	0	0
T4	0	7	2	7	3	0	0
	Total PR Verbs						
T1	3	17	2	12	4	0	0
T2	0	16	5	11	3	0	0
T3	5	15	0	12	7	0	0
T4	0	17	2	9	8	0	1

Whilst there was no significant quantitative change in PAS as a result of treatment, scrutiny of P4's sentence production pre-treatment strongly suggests that he had difficulty with event processing and that P4's ability to conceptualise events appears to have improved as a result of treatment with a knock-on effect of improved sentence production. Examples of sentences which demonstrate this change are given in Table 59 below.

Table 59 P4's production of sentences pre- and post-treatment.

Target Sentence: PR Verb Sentences	**P4's sentence: pre-treatment**	**P4's sentence: post-treatment**
the girls are visiting their grandad (in hospital)	man the shirt in his bed	the little girls at the hospital
the boys are watching TV	the two boys TV	The two were watch TV
the boy is watering the flowers	the flowers are watering the water	watering the garden
the couple have bought a house	the seller the car the house	a girl sell er is sold
the woman is driving the car	the woman how the car right or left	the girl is driving right then right and right
the people are talking	er /faid/ the one is the right	three four talking together
a ship hit by a torpedo	the hit by the side	he was bad hit
the man is selling bananas	bananas at a good price	sell a banana
1. Target Sentence: SCT Sentences	**2. P4's sentence: pre-treatment**	**3. P4's sentence: post-treatment**
the boy hits the girl	hits the man the baby the woman	the boy hits the girl on her head
the man is running	he's nice to running	the boy is running
the man is painting the woman	the /paincher/ of the white ||naked	the man painting the | | girls
the clown is smiling	the clown is the /krin/ (T: grin)	the clown do laughed
the woman is swimming	the woman no the swim across the distance	swimming fast
the baby is scratching the man	the baby's call the man going crazy	the little boy scratch the man

There was some evidence of improved PAS structure involving untreated personally irrelevant words (in the SCT - see Table 60 below), with the production of one argument structures increasing post- treatment, and this was significant at both T3 and T4 ($\chi^2 = 7.4831$, p=0.0062, $\chi^2 = 4.6060$, p=0.0318 respectively). Sentence production using personally irrelevant words also demonstrated an improvement in event processing: for example: target: the baby is scratching the man; P4's pre-treatment production: *the baby's call the man going crazy*; P4's post-treatment production: *the little boy scratch the man*.

*Table 60 Analysis of P4's PAS using personally irrelevant untreated verbs in the Sentence Construction Test (Bastiaanse et al., 2002). (UTS = undetermined thematic structure. Significant change compared to T2 (χ^2 Test): *=p<.05; ** =p<.01. Verbs alone were not analysed due to the small number of instances).*

	No output <10 seconds	UTS	Verb	Verb + 1 argument	Verb + 2 arguments	Verb + 3 arguments	Verb + complex phrase
SCT T1	0	7	0	5	5	3	0
SCT T2	3	11	0	2	3	0	0
SCT T3	1	3	1	**11****	2	1	0
SCT T4	0	7	0	**7***	5	0	0

6.4.v. P4's discourse production: picture description.

The total CAT Picture Description score for P4 is reported in row 1 of Table 61 below. P4's score was five points higher at T2 compared to T1. Post treatment his score was eight points higher at T3 and 6.5 points higher at T4 (compared to T2), but this was not significant (Friedman's Test: Q= 3.00, p=0.3916). The origin of the increase in P4's CAT Picture Description score post treatment was mostly due to increased production of appropriate ICWs together with a decrease in inappropriate ICWs, plus an increase in the score for syntactic variety (see rows 4 and 6 in Table 6.9.viii below).

Table 61 P4's scores on the CAT spoken picture description task.

P4	T1	T2	T3	T4
1. Total CAT Score	12.5	17.5	25.5	24
2. P1 Verb types	4	4	5	5
3. P1 Verb tokens	4	6	5	5
4. Appropriate ICWs	12	14	20	19
5. Inappropriate ICWs*	6	3	3	3
6. Syntactic Variety	1	2	4	3
7. Grammatical Well-formedness	5	3	3	3
8. Speed	1.5	1.5	1.5	1

The PAS of P4’s CAT Spoken Picture Description was also analysed using the approach adopted by Webster et al. (2007) and the results of this are reported in Table 62 below. This indicates that P4’s production of verbs together with an argument improved post-treatment: at T1 and T2 he produced four utterances including a verb together with an argument and this rose to nine at T3 and seven at T4. Specifically, production of verbs with two arguments improved both immediately and at maintenance with indications that production of verbs in more complex phrases was also increasing.

Table 62 The PAS of P4’s CAT Spoken Picture Description.

	UTS	Verb	Verb + 1 argument	Verb + 2 arguments	Verb + 3 arguments	Verb + complex phrase
CAT Picture Description T1	2	0	1	2	1	0
CAT Picture Description T2	4	1	3	1	0	1
CAT Picture Description T3	4	0	3	3	2	1
CAT Picture Description T4	3	0	0	4	1	2

6.4.vi. P4’s discourse production: The Autobiographical Memory Interview (AMI) (Kopelman et al., 1989).

Table 63 reports the PAS of P4’s AMIs. Immediately post treatment (at T3) there was a numerical improvement in the number of verbs produced with one argument and with two arguments. However, these results must be interpreted in the context of P4 producing approximately twice the number of utterances at T3 than at T1 and T2 (and T4 – see the bottom row of Table 6.9.x). Therefore, the relative proportion of utterances with a verb plus one argument, and plus two arguments, at each time point will be reported. For a verb with one argument this was 12% and 14% at T1 and T2 respectively, whilst immediately post-treatment at T3 it was 21% and at T4 23%. The proportion of utterances produced with a verb plus two arguments was 37% at

T1 and 36% at T2, whilst post-treatment the proportion was 45% at T3 returning to baseline (36%) at T4. There was also a decrease in the relative proportion of utterances produced with undetermined thematic structure (UTS). UTSs comprised 35% of utterances at T1 and 40% at T2, whilst at T3 this proportion decreased to 21% returning close to baseline at T4 (33%).

Table 63 The PAS of P4's AMIs (Kopelman et al., 1989).

	T1	T2	T3	T4
UTS	18	17	19	13
Verb	0	1	2	1
Verb + 1 argument	6	6	19	9
Verb + 2 arguments	19	15	41	14
Verb + 3 arguments	5	2	4	1
Verb + complex phrase	3	1	6	1
Total utterances	51	42	91	39

P4's AMIs were also scored using three of the criteria of the CAT Spoken Description Task to facilitate comparison with this task (Appropriate and Inappropriate Information Carrying Words could not be scored given that the AMI is a personal narrative). The scores for P4 are reported in Table 63 below and show minimal changes with an increase in grammatical-well-formedness at T3 only and a small increase in speed at T3 and T4.

Table 64 CAT Spoken Picture Description scores for P4's AMIs (Kopelman et al., 1989).

	T1	T2	T3	T4
Syntactic Variety	4	3	4	3
Grammatical Well-formedness	3	3	4	3
Speed	1	1	1.5	1.5

6.4.vii. Impact of the SPT on P4's Functional Communication.

CETI scores for P4 are reported in Table 65 below. P4 showed no significant change in his CETI ratings (Friedman test, Q=1.17, p=0.5558).

Table 65 The impact of the SPT on functional communication for P4 (CETI scores (Lomas et al., 1989). *(Significance level Friedman test: *=p.05; **p<.01; **p<.001; italics = significant using SE>5.2 change).*

CETI Score	T 1	T 2	T 3	T 4
P4	46.1875	45.375	48.9388	n/a

6.5. The impact of the SPT treatment on the production of verbs in isolation, in sentences, in discourse and on functional communication for P5.

6.5.i. Summary of P5's Results.

P5 had a mild anomic aphasia. She did however have difficulty with spoken and written comprehension whilst both her spoken and written picture descriptions demonstrated word finding difficulties and syntactic impairment. P5 had eight face-to-face treatment sessions over 17 weeks (due to participant illness) and self-managed a total of just under 7 hours of treatment during this period. (P5 carried on practicing her PR verbs using the SPT between T3 and T4 and so this is not a maintenance phase for her).

P5's production of PR treated verbs in isolation showed a significant improvement at T3 only, and this drove a significant improvement in the total set of verbs also at T3 only. There was no improvement in her production of untreated, personally irrelevant verbs or nouns at either T3 or T4. In terms of sentence production, there were no changes in the production of verbs and objects as a result of treatment. There was a significant improvement in production of untreated verb agents at T4 only. There was no improvement in the production of personally irrelevant verbs, agents or objects (on the SCT). P5's PAS using PR verbs showed a trend towards decreased production of untreated PR verbs with one argument, alongside an increase in the production of untreated PR verbs with three arguments at T4 only. This resulted in a similar pattern for the total set of PR verbs. PAS involving untreated personally irrelevant words showed no change at T3 or T4. In terms of discourse production, P5's picture description showed minimal changes as did her personal narrative. However, P5 perceived her functional communication to have improved significantly at T4.

6.5.ii. The effect of the SPT on P5's production of treated and untreated PR verbs in isolation.

The impact of treatment on the production of treated, untreated and the total set of verbs in isolation for P5 is reported in Table 66 below. P5 showed a stable performance regarding verb production in isolation prior to treatment. P5 showed significant improvement in treated verb production at T3 only (p=0.0313). There was also a non-significant trend of improvement in untreated verbs which resulted in a significant improvement in the total set of treated verbs at T3 only (p=0.0193).

Table 66 P5's production of verbs in isolation pre and post treatment. *(Significant change compared to T2 (McNemar Test): *=p<.05; **=p<.01).*

		T1	T2	T3	T4
P5	Treated	8	8	**13***	10
	Untreated	8	9	12	13
	Total	16	17	**25***	23

6.5.iii. The effect of the SPT on P5’s production of untreated (personally irrelevant) verbs and on noun retrieval.

P5’s scores on the OANB verb and nouns subsets are reported in Table 67 below. P5 showed unstable production of verbs pre-treatment, and no change in her production of untreated, personally irrelevant verbs and nouns as a result of treatment.

Table 67 P5's scores on a subset of verbs and nouns from the OANB. *(P5: verbs n=41; nouns n=41). (Significant change compared to T2 (McNemar Test): *=p<.05).*

		T1	T2	T3	T4
P5	**OANB Verbs**	31	37*	36	34
	OANB Nouns	38	37	36	38

6.5.iv. P5’s sentence production using treated and untreated PR verbs and untreated personally irrelevant verbs.

Production of untreated and treated PR verbs, agents and objects in sentences for P5 is reported in Table 68 below. (Production of untreated personally irrelevant verbs in sentences is reported via the VAST SCT in the table below). P5 demonstrated a stable performance at baseline in terms of the production of verbs and objects, but there was significant improvement in the production of treated verb agents between T1 and T2. There was improvement in the production of untreated agents at both T3 and T4 which was significant at T4 only (p=0.0313). There was no improvement in the production of personally irrelevant verbs, agents or objects (on the SCT).

Table 68 P5's production of verbs in sentences. *(Significant change compared to T2 (McNemar Test: *=p<.05).*

Sentence Production	T1	T2	T3	T4
PR Sentences Treated (n=20)				
Agent	13	**19***	18	19
Verb	10	14	15	13
Object	11	10	10	12
PR Sentences Untreated (n=20)				
Agent	13	14	18	**19***
Verb	14	11	12	14
Object	13	11	13	15
PR Sentences Total (n=40)				
Agent	26	*33*	36	38
Verb	24	25	27	27
Object	24	21	23	27
VAST Sentence Construction Test (n=19)				
Agent	17	18	17	18
Verb	15	17	17	16
Object	13	15	13	10

Analysis of P5's PAS using PR treated and untreated PR verbs is reported in Table 69 below. (Verbs alone and verbs with 1 argument (untreated verbs) were not analysed due to the small number of instances). There were no statistically significant changes in P5's PAS using PR verbs. However, at T4 there was a trend indicating decreased production of untreated verbs with one argument alongside an increase in the production of verbs with three arguments which resulted in a similar pattern for the total set of verbs. Examples of sentences with three arguments are given in Table 70 below.

Table 69 Analysis of P5's PAS using PR treated, untreated and the total set of verbs. *(UTS = undetermined thematic structure).*

	No output <10 seconds	UTS	Verb	Verb + 1 argument	Verb + 2 arguments	Verb + 3 arguments	Verb + complex phrase
	Treated PR Verbs						
T1	3	1	0	7	8	1	0
T2	1	1	0	6	12	0	0
T3	0	4	0	8	6	2	0
T4	1	3	0	4	10	2	0
	Untreated PR Verbs						
T1	2	1	2	3	11	1	0
T2	4	1	1	4	9	1	0
T3	0	2	0	6	10	2	0
T4	0	3	1	1	9	6	0
	Total PR Verbs						
T1	5	2	2	10	19	2	0
T2	5	2	1	10	21	1	0
T3	0	6	0	14	16	4	0
T4	1	6	1	5	19	8	0

Table 70 P5's production of verbs with three arguments post treatment.

Target sentence PR verbs	P5's sentence: T3	P5's sentence: T4
the chef is cooking	the girl is stirring the carrots	the man is frying something in the pan
the man is thinking	the man is thinking	the man has got something on his mind
the girls are visiting their grandpa	the girls come to visit him || in hospital	the girl is saying "Hi" to (I think it's) her grandad
the boy is planting a tree	the boy is planting the tree	the young boy is planting a tree in the garden
a woman is seasoning the sauce	she's putting pepper into || soup	the woman is putting pepper in the soup
someone is melting butter	she's melting butter in a pan	the woman is melting butter in the pan
the woman is blending	she's whisking something	the woman is mixing cornflour in a blender
the girls is helping the elderly lady	||she's helping this lady along	the young girl is helping the old woman to walk

In five of the eight sentences P5 produced post treatment, the third argument comprises a locative and this was a focus of treatment during the sentence phase of treatment (i.e., using the prompt word "Where?" to cue production of the locative). The effect of treatment appeared to be lexically based as the PAS of untrained personally irrelevant verbs did not change. That the treatment effect emerged at T4 may reflect the additional treatment that P5 self-managed between T3 and T4 which was greater than the treatment she self-managed during the treatment phase in terms of both amount and intensity as discussed earlier.

Analysis of P5's PAS using personally irrelevant untreated verbs is reported in Table 71 below. There was a non-significant trend in increased production of verbs with one argument post-treatment.

Table 71 Analysis of P5's PAS using personally irrelevant untreated verbs in the Sentence Construction Test (Bastiaanse et al., 2002). *(UTS = undetermined thematic structure).*

	No output <10 seconds	UTS	Verb	Verb + 1 argument	Verb + 2 arguments	Verb + 3 arguments	Verb + complex phrase
SCT T1	2	1	0	3	10	3	0
SCT T2	0	1	0	3	11	4	0
SCT T3	0	2	0	5	11	1	1
SCT T4	0	1	0	7	9	2	0

6.5.v. P5's discourse production: picture description.

The total CAT Picture Description score for P5 is reported in row 1 of Table 72 below. P5's score was five points higher at T3 compared to T1 and T2 but this was not significant (Friedman's Test: Q= 1.043, p=0.6997). The origin of the increase in P5's CAT Picture Description score post treatment was mostly due to decreased production of inappropriate ICWs together with an increase in the score for syntactic variety (at T3 only) and grammatical well-formedness (see rows 4 and 6 in Table 6.10.viii below). Perusal of P5's production of verbs during picture description indicated that the types and tokens of verbs produced did not increase post treatment.

Table 72 P5's scores on the CAT spoken picture description task.

P5	T1	T2	T3	T4
1. Total CAT Score	29	29	34	31
2. P1 Verb types	2	4	4	2
3. P1 Verb tokens	2	7	6	8
4. Appropriate ICWs	24	19	23	22
5. Inappropriate ICWs*	4	0	0	2
6. Syntactic Variety	2	3	4	3
7. Grammatical Well-formedness	4	5	5	5
8. Speed	3	2	2	2

The PAS of P5's CAT Spoken Picture Description reported in Table 73 below also did not change.

Table 73 The PAS of P5's CAT Spoken Picture Description.

	UTS	Verb	Verb + 1 argument	Verb + 2 arguments	Verb + 3 arguments	Verb + complex phrase
CAT Picture Description T1	1	0	0	7	1	0
CAT Picture Description T2	1	0	3	7	2	0
CAT Picture Description T3	4	0	0	5	1	0
CAT Picture Description T4	1	0	0	6	3	0

6.5.vi. P5's discourse production: the Autobiographical Memory Interview (AMI) (Kopelman et al., 1989).

Table 74 reports the PAS of P5's AMIs. The number of utterances with an UTS reduced at T3 matched by an increase in verbs with one argument, but there was little other evidence of change. There was however evidence of a potential assessment burden effect with fewer utterances produced at T4.

Table 74 The PAS of P5's AMIs (Kopelman et al., 1989).

	T1	T2	T3	T4
UTS	15	12	8	6
Verb	0	0	0	0
Verb + 1 argument	7	6	10	4
Verb + 2 arguments	36	43	44	30
Verb + 3 arguments	12	17	18	6
Verb + complex phrase	11	11	9	5
Total utterances	81	89	89	51

P5's AMIs were also scored using three of the criteria of the CAT Spoken Description Task to facilitate comparison with this task. The scores for P5 are reported in Table 75 and show no change.

Table 75 CAT Spoken Picture Description scores for P5's AMIs (Kopelman et al., 1989).

	T1	T2	T3	T4
Syntactic Variety	5	5	5	5
Grammatical Well-formedness	5	5	5	5
Speed	2	2	2	2

6.5.vii. Impact of the SPT on P5's Functional Communication.

CETI scores for P5 are reported in Table 76 below. P5 rated her communication significantly higher at T4 and this was significant according to Friedman's Test, (Q= 7.71, p=0.0211), as well as representing an increase >5.2 retest standard error of the mean.

Table 76 The impact of the SPT on functional communication for P5 (CETI scores (Lomas et al., 1989). *(Italics = significant using SE>5.2 change).*

CETI Score	T 1	T 2	T 3	T 4
P5	59.125	61.4375	59.3125	***74.6875***

6.6. The impact of the SPT treatment on the production of verbs in isolation, in sentences, in discourse and on functional communication for P6.

6.6.i. Summary of P6's Results.

P6 had a severe non-fluent aphasia. P6's spoken lexical retrieval was very poor, with verbs even more severely affected than nouns. He was, however, extremely phonemically cueable. P6 had excellent ability to repeat words but was completely unable to repeat nonwords. His ability to read words aloud was severely impaired but he had very good understanding of both spoken and written words and sentences, with some difficulty with reversible sentences indicating problems with mapping. P6 had eight face-to-face treatment sessions over 11 weeks and self-delivered a total of over 45 hours of treatment during this period. P6 continued to practise his PR verbs between T3 and T4 (for an additional 12 hours) meaning that this does not represent a true maintenance phase.

Following treatment P6's production of treated PR verbs *in isolation* showed a trend of improvement post-treatment which was not statistically significant. This did however represent an improvement of 125% at T3 and 50% at T4 (compared to T2). The improvement in treated verbs also drove an improvement in the production of the total set of verbs which was again not significant, but which represented an improvement of 86% at T3 and 57% at T4. P6 showed a significant improvement in untreated personally irrelevant verbs in isolation at T4 only, and no improvement in nouns. In terms of *sentence* production using PR verbs, the most notable (significant) improvement was in P6's ability to produce agents immediately post treatment (at T3). There was also significant improvement in his ability to produce PR verbs in sentences post treatment which was evident at T4 for the total set of verbs. There was no change in the production of personally *irrelevant* verbs in sentences. In term of *PAS*, prior to treatment P6 had been unable to produce any of his PR verbs together with an argument, but post treatment he was able to produce a verb with one argument for 11/40 of his PR verbs at T3 and 8/40 at T4. There was also a trend of improved production of untreated personally irrelevant verbs with one argument (on the SCT) at T4 only. There was no evidence of any improvement in P6's discourse production on the CAT Picture Description Task post treatment. (Narrative production was not assessed as this was too distressing). In terms of functional communication P6 rated his communication more positively immediately post treatment but his rating returned to baseline at T4.

6.6.ii. The effect of the SPT on P6's production of treated and untreated PR verbs in isolation.

The impact of treatment on the production of treated, untreated and the total set of verbs in isolation for P6 is reported in Table 77 below. P6 showed an improvement in verb production between baseline assessments which was not significant. Therefore, post-treatment comparisons are with T2. P6 showed a trend of improved production of treated verbs which was not statistically significant. It did however represent an improvement of 125% at T3 and 50% at T4. P6 also showed a trend of improvement in untreated verbs which represented an increase of 33% at T3 and 66% at T4. The improvement in treated verbs drove an improvement in the production of the total set of verbs which was again not statistically significant, but which represented an improvement of 86% at T3 and 57% at T4.

Table 77 P6's production of verbs in isolation pre and post treatment.

		T1	T2	T3	T4
P6	Treated	1	4	9	6
	Untreated	1	3	4	5
	Total	2	7	13	11

Finally, as a result of treatment, P6's ability to retrieve the written form of his PR verbs improved as did his ability to self-cue using the written form (see Table 78 below).

Table 78 Response to cueing and error type for P6's production of verbs in isolation.

	Correct	Successful Phonemic Cue	Unsuccessful Phonemic Cue	Delayed Response	Semantic Error	Verb Instrument	Phonological Error	Written Naming	Written Naming -> Self Cue
T1	2	29	1	3	5	1	1	0	0
T2	7	29	1	3	0	1	0	0	0
T3	13	10	0	2	1	1	0	8	7
T4	11	7	0	2	1	3	0	5	10

6.6.iii. The effect of the SPT on P6's production of untreated (personally irrelevant) verbs and on noun retrieval.

P6's scores on the OANB verb and nouns subsets are reported in Table 79 below. P6 demonstrated a statistically significant change in production of untreated, personally irrelevant verbs at T4 only (p=0.0195). Production of verbs at T4 represented an improvement of 78%.

Table 79 P6's scores on a subset of verbs and nouns from the OANB. *(P6: verbs n=42; nouns n= 42). (Significant change compared to T2. McNemar Test: *=p<.05; **=p<.01).*

		T1	T2	T3	T4
P6	**OANB Verbs**	9	9	11	**16***
	OANB Nouns	15	20	18	20

6.6.iv. P6's sentence production using treated and untreated PR verbs and untreated personally irrelevant verbs.

Production of untreated and treated PR verbs, agents and objects in *sentences* for P6 is reported in Table 80 below. P6's production of PR verbs in sentences at baseline was unstable with a statistically significant improvement from T1 to T2, but production of agents and objects was stable. (Comparison post treatment is with T2 for all). There was a trend of improved production of both treated and untreated PR verbs at T4 which was not statistically significant, but which drove a significant improvement in the total set of verbs at T4 (p=0.0193). Production of treated agents improved significantly at both T3 (p=0.0059) and T4 (p=0.0156), with tentative evidence that production of objects was emerging. In terms of *untreated* verbs, again there was a trend of improved production at T4 which was not statistically significant, but which nevertheless represented an improvement over baseline of 80%. Untreated agents improved significantly at T3 (p=0.0010) with no change evident in the production of objects. There was a significant improvement in the total set of agents at T3 (p=0.0000) and T4 (p=0.0010). Change in the production of personally irrelevant verbs in sentences was measured via the SCT from the VAST. Performance for agents was not stable prior to treatment with improvement occurring between T1 and T2. There was a trend of improved production of personally irrelevant verbs which was not significant.

Table 80 P6's production of verbs in sentences. *(Significant change compared to T2 (McNemar Test): *=p<.05; ** =p<.01; ***p=<.001).*

		T1	T2	T3	T4
Treated	Agent	1	1	**10****	**7***
	Verb	2	**8***	9	12
	Object	0	0	1	1
Untreated	Agent	0	0	**10*****	4
	Verb	3	5	5	9
	Object	0	0	0	0
Total	Agent	1	1	**20*****	**11*****
	Verb	5	**13***	14	**21***
	Object	0	0	1	1
Sentence Construction Test (n=18)	Agent	7	12	n/a	9
	Verb	0	2	n/a	5
	Object	0	2	n/a	0

Analysis of P6's *PAS* using treated and untreated PR verbs is reported in Table 81 below. (Production of verbs with an argument was not subjected to statistical analysis because of zero instances). Having produced no verbs together with an argument before treatment, P6 started to produce verbs with one argument following treatment, and this was true for both treated and untreated verbs (although the improvement in treated verbs was larger). It is also of note that before treatment P6 was unable to produce any output before 10 seconds for 30/40 of his PR verbs (average of T1 and T2), whereas after treatment this was the case for only 12.5 verbs (average of T3 and T4). P6's increased speed of sentence production meant that he was now able to produce PR verbs *together with an argument* within the 10 seconds threshold for the first time: at T3 he was able to produce 11/40 PR verbs together with an agent within 10 seconds, and at T4 9/40. P6 also demonstrates a pattern of emerging production of verbs with *two* arguments (agent + verb + theme) but only if responses made *after 10 seconds* are analysed (which accounts for the increase in his production of utterances of undetermined thematic structure (usually an agent + "is"). However, when these arguments produced after 10 seconds are included, P4 was able to produce a verb with two arguments on four occasions (*I drive the car; I like music; she is listening to music; she is roasting chicken*) but this ability only emerged at T4 suggesting that it was cumulative effect of the treatment P6 continued to self-deliver between T3 and T4.

Table 81 Analysis of P6's PAS using PR treated, untreated and the total set of verbs. *(UTS = undetermined thematic structure. Verbs alone and a verb + 3 arguments were not analysed due to the small number of instances).*

	No output <10 seconds	**UTS**	**Verb**	**Verb + 1 argument**	**Verb + 2 arguments**	**Verb + 3 arguments**	**Verb + complex phrase**
	Treated PR Verbs						
T1	17	1	2	0	0	0	0
T2	12	0	8	0	0	0	0
T3	5	5	3	7	0	0	0
T4	4	4	6	6	0	0	0
	Untreated PR Verbs						
T1	16	0	4	0	0	0	0
T2	15	0	5	0	0	0	0
T3	8	6	2	4	0	0	0
T4	8	2	8	2	0	0	0
	Total PR Verbs						
T1	33	1	6	0	0	0	0
T2	27	0	13	0	0	0	0
T3	13	11	5	11	0	0	0
T4	12	6	14	8	0	0	0

PAS using untreated personally *irrelevant verbs* was measured by the SCT (see Table 82 below). There was a non-significant trend in P6's production of personally irrelevant verbs with one argument as a result of treatment, mirroring his performance with PR verbs. Again, similar to the improvement with PR verbs, P6 was unable to produce any output before 10 seconds for 10/20 of the irrelevant verbs (average of T2 and T3), whereas after treatment this was the case for only 2 verbs. This also led to an increase in his production of utterances of undetermined thematic structure (again similar to his PR verbs usually an agent +/- "is").

Table 82 Analysis of P6's PAS using personally irrelevant untreated verbs in the Sentence Construction Test (Bastiaanse et al., 2002). *(UTS = undetermined thematic structure).*

	No output <10 seconds	**UTS**	**Verb**	**Verb + 1 argument**	**Verb + 2 arguments**	**Verb + 3 arguments**	**Verb + complex phrase**
SCT T1	13	7	0	0	0	0	0
SCT T2	7	11	0	2	0	0	0
SCT T3	n/a	n/a	n/a	n/a	n/a	n/a	n/a
SCT T4	2	13	0	5	0	0	0

6.6.v. P6's discourse production: picture description.

Table 83 below reports P6's scores on the CAT Picture Description task. P6's overall score improved because of an increase in the number of appropriate ICWs he produced. These were all nouns on each occasion of testing. The only verb P6 produced during picture description was "is" and it was unclear if this was the copula (e.g., guy is - no) or an attempt to produce the present progressive form of another verb (e.g., the boy is car).

Table 83 P6's scores on the CAT spoken picture description task.

P6	**T1**	**T2**	**T3**	**T4**
Total CAT Score	7	7	10	9
P6 Verb types	0	1	1	1
P6 Verb tokens	0	6	5	2
Appropriate ICWs	7	7	10	10
Inappropriate ICWs*	0	0	0	1
Syntactic Variety	0	0	0	0
Grammatical Well-formedness	0	0	0	0
Speed	0	0	0	0

This uncertainty is reflected in the PAS analysis of P6's CAT Spoken Picture Description in terms of the presence of utterances of undetermined thematic structure only (see Table 84 below).

Table 84 The PAS of P6's CAT Spoken Picture Description.

	UTS	Verb	Verb + 1 argument	Verb + 2 arguments	Verb + 3 arguments	Verb + complex phrase
CAT Picture Description T1	8	0	0	0	0	0
CAT Picture Description T2	9	0	0	0	0	0
CAT Picture Description T3	14	0	0	0	0	0
CAT Picture Description T4	13	0	0	0	0	0

6.6.vi. The impact of the SPT on P6's discourse production: the Autobiographical Memory Interview (AMI) (Kopelman et al., 1989).

P6's AMI's were not analysed as he was unable to produce any verbal response to the AMI questions (at T1, T2 and T3). The process was also distressing for him and so the interview was cut short by the research student at T1, T2 and T3, and not carried out at T4.

6.6.vii. The impact of the SPT on P6's Functional Communication.

P6's CETI scores are reported in Table 85 below. His scores were stable at baseline and showed a significant improvement post treatment (Friedman's Test, Q= 20.42, p=0.0000) which was also >5.2 retest standard error of the mean. His CETI score returned to baseline at T4.

Table 85 The impact of the SPT on functional communication for P6 (CETI scores (Lomas et al., 1989)). *(Italicised scores show change >5.2 retest standard error of the mean).*

CETI Score	T 1	T 2	T 3	T 4
P6	62.8125	60.3125	*78.125****	62.3125

6.7. Summary.

In summary, all six participants showed significant improvement in retrieval of PR verbs post-treatment. This was restricted to treated PR verbs for four participants. All six participants also showed significant improvements in lexical retrieval in sentences, and this reflected improved production of verbs and/or agents and /or objects, which led to improved PAS for five participants. The effect of the SPT on discourse was more limited, with three participants showing improvements. Four participants reported improvements in functional communication, and these were corroborated by two SOs.

The best responders to treatment (P1, P2, P3) all had non-fluent aphasia and reasonably intact lexical retrieval. P6 also had non-fluent aphasia but with a more severe impairment of lexical retrieval and of syntactic processing, His response to treatment, whilst also significant, was less widespread. P4 and P5 both had fluent aphasia and showed a more limited response to treatment. The fluent nature of their impairment may have accounted for this. However, also influential may have been the degree of their aphasia: P4 was the most severely impaired participant overall, whilst P5 was the most mildly impaired. Additionally, P5's treatment was much more distributed than intended, and she also self-delivered the least amount of treatment. Finally, it should be noted that three participants continued to self-deliver the SPT during the maintenance phase (P2, P5 and P6). This indicates that they found the SPT acceptable and perceived it to be beneficial.

Chapter 7. Discussion.

The contributions made by this PhD study to the existing literature will be summarized, including those of the two systematic scoping reviews carried out to inform the development of the novel SPT programme at the heart of the study. The SPT programme itself will then be discussed in relation to the current situation regarding self-delivered, computer-based aphasia treatments, and the unique characteristics of the SPT highlighted. Finally, feasibility results and those of preliminary efficacy testing are discussed.

7.1. Background.

This study explored the feasibility, acceptability, compliance and fidelity of a novel verb and sentence production treatment, together with some preliminary efficacy testing. The verb and sentence treatment program was low dose and clinician delivered, supplemented by self-managed computer-based treatment, and this mode of delivery has not been explored before in relation to sentence level treatments. This PhD study makes several additional unique contributions to the evidence base for verb and sentence production treatments in aphasia, and the desire to do this arose from multiple factors which inspired the PhD.

The first was an acute awareness of the limited publicly funded speech and language therapy for PwA leading them to increasingly rely on computer-based treatments which they could administer themselves (e.g., Cann, 2021; van de Sandt-Koenderman, 2011; Kurland, 2014). However, the evidence for the effectiveness of such treatment is limited (e.g., Lavoie, Macoir & Bier, 2017; Zheng, Lynch & Taylor, 2016) yet the prescription of it is, as Kurland et al. (2018) lament, becoming almost routine as a means of supplementing face-to-face treatment. In particular, evidence is lacking for computer-based treatments which aim to improve impairments *beyond* single word processing i.e., sentence level treatments. Specifically, to date, there are only two studies that investigated computer-delivery of *sentence* treatments (Furnas & Edmonds, 2014; Thompson, Choy, Holland & Cole, 2010) and in neither of these was treatment *self-delivered*. There are six studies which investigate the self-delivery of *verb* treatments via computer (Kurland et al., 2014; 2018; Mortley et al., 2004; Palmer et al., 2012; 2019; Routhier et al., 2016), but in only one of these studies was the treatment specifically designed for verbs (Routhier et al., 2016), with the others treating both nouns and verbs. This pointed to the need to i) develop a bespoke treatment programme to improve verb and sentence production which could be delivered via computer, ii) explore the feasibility of this mode of treatment delivery, and iii) carry out some preliminary efficacy testing of the computer delivered sentence treatment to begin to address the gap in the evidence base. Thus, the Sentence Production Treatment (SPT) programme was developed for this PhD research and preliminary efficacy testing was conducted using pre-post studies of individual cases (n=6).

The SPT investigated in this study is a complex intervention (as defined by the Medical Research Council (MRC) (2019): https://mrc.ukri.org/documents/pdf/complex-interventions-guidance/) because it has several interacting components. As recommended by the MRC, the development of the SPT was therefore informed by (two) systematic scoping reviews to ensure that it was based on the best available evidence, and each of these reviews added to the existing knowledge base. The first review investigated treatments which aimed to improve the retrieval of verbs in isolation (Hickin et al., 2020: Chapter 1), and the second investigated treatments which aimed to improve sentence production by treating verbs in the context of a sentence, targeting a verb together with its arguments (Hickin et al., 2022: Chapter 2). The reviews included studies which used a variety of designs including group studies, case series and single case studies, and studies which investigated treatments using a *single* technique (e.g., phonological cueing of verbs) or which used a *combination* of treatments. The broad nature of the scoping reviews conducted (i.e., that they included a variety of research designs and treatments) makes them unique as previous reviews have been restricted to particular techniques (e.g. Semantic Feature Analysis (SFA) (Efstratiadou et al., 2018; Maddy et al., 2014) or one type of design (e.g. de Aguiar et al., (2016) who conducted a meta-analysis of verb treatments restricted to single case studies). Because the SPT programme developed for this PhD research was designed to be delivered by computer, studies which investigated verb or sentence treatment delivered via computer were also included in the reviews (although, as pointed out above, there were few of them). Again, this makes the reviews carried out for this PhD research unique as no previous reviews of verb or sentence treatments have included computer-based treatments.

A further gap in the evidence base concerned the lack of a review of the *fidelity* of verb and sentence treatments. The reviews carried out evaluated the fidelity of both verb and sentence treatments which has not been reported on before. The reviews found that the fidelity of verb and sentence treatments has been infrequently assessed and thus they provide fresh evidence supporting the need for the fidelity of verb and sentence treatments to be addressed.

Existing verb and sentence treatment research has used a variety of outcome measures to assess the efficacy of treatment, and this represents a further challenge in interpreting the evidence in this field, as does the inconsistent measurement of the effect of treatment at different levels of language (Conroy et al., 2006; Webster & Whitworth, 2012). The reviews carried out charted the levels at which each reviewed study assessed the impact of treatment, namely the impact of treatment on a) production of trained verbs in isolation and in sentences, b) production of untrained verbs in isolation and in sentences (within and across level generalisation), c) functional communication and d) discourse, as well the ways in which this was assessed. This updates the review of verb treatments carried out by Webster and Whitworth almost a decade ago (2012). The level of the evidence of verb and sentence treatments is also reported, as are the potential “active ingredients” of verb and sentence treatments (so that these could be included in the SPT), and, again, this is novel.

Conducting reviews of both verb and sentence treatments also enabled comparisons to be made between the two fields yielding findings of interest in relation to guiding future research. Specifically, sentence treatments appeared to be more effective in improving sentence and discourse production than verb treatments. However, sentence treatments were usually more complex than verb treatments (i.e., they had more potential active ingredients, often including those given in verb treatments), and they tended to be given in larger doses. Further confounding factors were that whilst sentence and discourse production were frequently assessed in studies of sentence treatment, they were infrequently assessed in verb treatment studies meaning that a positive effect of verb treatment on sentence and discourse production may have been missed. Finally, the wide range of outcome measures used in both types of study further confounded interpretation of the evidence. Therefore, future research into verb and sentence treatments should: ensure that they are given in comparable doses; try to elucidate what are the active ingredients of treatments (e.g. is it the syntactic components of sentence treatments that are most active, or do these need to be combined with the ingredients of verb treatments?); routinely assess key levels of language production across both types of treatment; and agree a core set of outcome measures to facilitate comparison of studies/interpretation of the evidence.

In terms of the treatment which the reviews informed, the SPT developed for this PhD study is the first sentence level treatment developed for self-delivery by computer to incorporate a *variety* of evidence-based verb *and* sentence treatments (identified from the reviews). Thus, for instance, Phase 1 of treatment has verb exercises based on the studies by, for example, Conroy et al. (2009a, b & c) and on SFA adapted for verbs in a series of studies by Wambaugh and colleagues (e.g., Wambaugh et al., 2014). Phase 2 comprises sentence level exercises based on mapping treatments (e.g., Byng et al., 1994; Nickels et al., 1991) and VNeST (e.g., Edmonds et al., 2014). VNeST has been adapted (from face-to-face treatment) for delivery by computer (Furnas & Edmonds, 2014), however this adaptation was not designed for self-delivery, whilst mapping treatments have never to date, to my knowledge, been adapted for delivery by computer. Phase 3 of the SPT which was designed to facilitate generalisation of treatment to real life communication drew upon discourse level treatments delivered face-to-face (e.g., Carragher, Sage & Conroy, 2014) and a discourse level treatment adapted for computer delivery: script training (Cherney, Kaye & Van Vuuren, 2014).

Finally, a search of the Aphasia Software Finder (https://www.aphasiasoftwarefinder.org/advanced-software-search) identified 14 apps which targeted spoken sentence production, and only seven of these were deemed to be theory based (according to the checklist provided by the Aphasia Software Finder). This highlights the urgent need for the design of *theory-based* sentence level treatments for self-delivery by computer (including mapping treatments, PAS treatments and VNeST), and the SPT is a prototype of such a computer-based, self-delivered treatment. It should also be noted that the computer-based aphasia treatments with the best evidence base are currently StepByStep (e.g., Mortley et al., 2004; Palmer et al., 2012; 2019) and Constant Therapy (e.g., Braley et al., 2021; Godlove et al., 2019) because they have RCTs demonstrating their effectiveness. Both StepByStep and Constant Therapy predominantly treat words in a *single word* context and concentrate on nouns. For example, Constant Therapy has only two exercises devoted to *sentence* planning (Des Roches et al., 2015, Table 2C), and these were excluded from the analysis of the effectiveness of exercises by the authors because there was not enough data suggesting that they were not frequently practiced (Table 3A)). Regarding verb treatment, less than 3% of the PR words chosen by the 100 participants in the Big CACTUS study of StepByStep were verbs (Palmer et al., 2017). Again, this suggests that more attention be paid to verb and sentence production in computer-based aphasia treatments as is the case for the SPT. Indeed, as noted in the reviews of verb and sentence treatments reported in Chapters 1 and 2, targeting verbs may be more effective in improving sentence production than targeting nouns because of the pivotal role of verbs in sentence production. As reported by Palmer et al. (2019), although treatment using StepByStep improved retrieval of the PR words chosen by participants, this did not generalise to functional communication or to everyday conversation. They note that treatment may need adapting to enable generalisation, and treating more verbs, including in a sentence context could be one such adaptation. Generalisation may also require specific treatment exercises (e.g., Newton et al., 2017) and an evidence-based generalisation phase of treatment is, unique to self-delivered computer-based treatments, included in the SPT.

7.2 The feasibility of a) recruiting and retaining suitable participants to the SPT, b) self-management of the SPT using a computer, and c) selecting a set of PR verbs.

7.2.i. Recruitment and retention of suitable participants to the SPT.

In total, 12 participants were screened for suitability for inclusion in the study. Of these, eight participants (66%) were suitable for the SPT, with six (75%) consenting to participate in the study and all of these participants were successfully retained for the duration of the study. Bearing in mind the very small number of participants, the results of this study suggest that it is feasible to recruit and retain participants to the SPT and that a future, larger scale study is achievable. Candidacy in terms of the characteristics of PwA that make them most likely to respond well to the SPT – is discussed in section 7.7 below.

In terms of retention, all (six) candidates were retained for the duration of the study indicating that the self-delivery of the SPT was not too onerous for them. However, the researcher also implemented strategies to improve participant retention, namely visiting participants in their homes when they were unable to attend the clinic, accommodating their schedules in arranging treatment sessions (e.g., to allow P4 to visit his terminally ill mother) and keeping in regular communication with them. These strategies were amongst those found to be effective in improving retention by Spell, Richardson, Basilakos, Stark, Teklehaimanot et al. (2020) in a treatment study involving 150 participants for a total of 42 weeks (the Predicting Outcomes of Language Rehabilitation - POLAR – study) and it is intended that similar strategies are implemented in any future study.

7.2.ii. The feasibility of self-management of the SPT.

The *technical feasibility* of the SPT will be discussed first. Because no funding was available to purchase devices for potential participants to the study, an initial concern was that PwA may be excluded from the study because they did not own a computer. This did not turn out to be the case. However, recruitment material did state that treatment exercises would be on participant's own computers, so it is possible that potential participants without a computer did not come forward. Therefore, any future implementation of the study would include funding to loan a device to a PwA who did not own a computer (as was the case in, for example, Kurland et al., 2018 and Palmer et al., 2019, although Harrison et al. (2020) note that participants who used their own device tended to practice more).

A second important finding was that the SPT worked on a variety of devices (namely tablets, laptops, desktops, and a smart phone), and on Windows based and Apple devices. Menger et al. (2016) discuss the many factors which make PwA at risk of digital exclusion and include financial barriers as one of these. It was thus important that the SPT worked on multiple platforms so that additional devices did not need to be purchased[23]. The capacity for computer-based aphasia

[23] NB although P5's husband did purchase a new computer, this was in addition to an iPad they already had.

treatment packages to work on multiple platforms was also a high priority for SLTs, as reported by Swales et al. (2016). Finally, the SPT exercises are PowerPoint slides. PowerPoint was deliberately used because this software is commonly available on a range of devices and thus it was not necessary to purchase software to use the SPT[24].

Another factor that could affect the feasibility of self-managed computer-based aphasia treatment was the *ability of participants* to use the SPT independently: five of the six participants achieved independent use of the SPT. This is similar to the figure reported by Kurland et al. (2018) in relation to their iPad-based verb treatment (they report that 18/21 participants (86%) achieved independent use) and De Cock, Batens, Feiken, Hemelsoet, Oostra and De Herdt (2021) who found that 79% of participants achieved independent use of a tablet-based intervention given at the acute stage of treatment. The level of independent use likely reflects the aphasia-friendly design of the SPT which was informed by the small but developing field of research into aphasia-friendly HCIs. Future implementations of the SPT will seek to refine the aphasia-friendliness of the SPT by involving PwA in a co-design process, including not just the design of the HCI but also the content of the SPT (see later discussion). That one participant did not attain independent use of the SPT will be addressed in future implementations of the SPT by including a more nuanced assessment of *current* use of technology (which appeared to be the decisive factor for this candidate). For example, Hickin, Woolf, Caute and Robey (2011) developed a questionnaire to investigate how independent PwA were in using technology. The Functional Assessment of Computer use in Aphasia (FACUA) was developed from structured observations of two PwA accessing their computers to send emails and the difficulties they had in doing so. PwA were asked to rate their independence at the various

[24] NB The SPT can also be used in KeyNote the Apple equivalent to Powerpoint.

stages of sending emails using a visual analogue scale. They also completed a brief questionnaire about the frequency and purposes for which they used the Internet (e.g., banking, shopping, social networking). An adapted form of the FACUA questionnaire could be used to assess how independently participants were in carrying out tasks relevant to using the SPT (such as dealing with Pop Ups relating to software updates), as well as establishing how frequently and for what purposes they used technology which would also be indicative of competence levels. Additionally, specific training sessions will be arranged for those participants identified as needing this (rather than training participants during assessment sessions as was the case in this study).

A second issue in relation to the independent use of the SPT relates to the process of updating exercises. New exercises were uploaded in two different ways: by email and by downloading them from a memory stick. Whilst all but P5 were able to *carry out* the SPT exercises independently, only three participants were entirely independent in *updating* exercises. Kurland et al. (2018) also report different levels of independence in their self-managed iPad-based verb treatment: 18/21 participants were independent in carrying out treatment exercises, but only 14/21 participants could access weekly remote supervision independently. To optimise the level of independence of participants in updating exercises, future implementations would include training in uploading new exercises as well as using the SPT exercises themselves.

7.2.iii. The feasibility of selecting a set of personally relevant (PR) verbs for treatment.

The feasibility of selecting a set of PR verbs is discussed in relation to the utility of the various methods used to help with this process, the time taken to choose PR verbs, and range of topics to which selected verbs belonged.

All six participants successfully selected a set of 40 PR verbs for treatment, although the degree to which they proactively selected their verbs varied considerably: P2 choose her verbs entirely independently with the help of her partner, whilst P1 relied on the researcher to choose his, with the other four participants using a combination of self-selection together with verbs suggested by the researcher. The various methods used to assist the selection process also varied in their utility. (The process is described in detail in Chapter 4.4.i). Key to the selection of PR verbs was - getting to know each participant's interests and hobbies; who were the important people with whom they communicated; and what situations were the most difficult communication wise. The People, Situations and Topics prompt cards from the CAPPA (Whitworth et al., 1997) were useful in this regard because they facilitated broad discussion of participants' communication needs in that the prompt cards helped them to consider communication outside of the home and immediate family and friends (such as at clubs attended or at health appointments). The CETI (Lomas et al., 1989) also helped to expose contexts where communication was difficult, and, in particular, that participants without exception rated "discussing things in depth" as difficult, supporting the decision to encourage participants to consider not only what they *had* to say but what they *wanted* to say. This is in line with the findings from Wallace et al.'s (2017) international study of the treatment outcomes desired by PwA, including that they wanted to have complex conversations involving giving explanations and expressing their emotions and opinions. Thus, participants were encouraged to consider choosing more abstract verbs (such as "hope" and "feel"), and this was also in line with the aim of addressing the likely bias towards targeting highly concrete words in PR vocabularies as a result of the pictographic resources commonly used in treatment (as discussed by Renvall et al. (2013a) – see Chapter 3.2.ii for a fuller discussion). A goal action planning (GAP) process (Scobbie et al., 2013; 2011) was also integral to the selection of PR verbs because it enabled the selection of verbs in line with these goals.

In summary, the process of selecting PR verbs combined several different techniques which were used during baseline assessment sessions, and where possible the selection process was integrated with baseline assessment tasks (e.g., using the results of the CETI to inform PR verb selection, and discussing OANB verbs which were difficult as potential PR verbs). This meant that the process of PR verb selection was not overly time-consuming, with final selection of verbs taking approximately half an hour across two assessment sessions. Once PR verbs had been determined, the researcher spent an average of two – three hours searching Google images for appropriate photographs to represent PR verbs, and for P2, P5 and P6, uploading personal photographs they had provided to represent PR verbs. This process became quicker for later participants because some PR verbs were common across participants (e.g., eat, pay, talk) and therefore the same images could be used.

The range of topics to which PR verbs belonged will now be discussed. The most popular topics to which PR verbs chosen by the six participants belonged were *entertainment and hobbies*, *food and drink* and *nature and gardening.* This finding is in line with Palmer et al. (2017) (in relation to PR nouns) and suggests that verbs as well as nouns relating to these topics are valid suggestions for clinicians to put forward when helping PwA to choose PR words for treatment. Additionally, the selection process used to select PR verbs in this study appeared to result in the selection of more abstract verbs as well as material ones, as was hoped. This is attested to by the frequency with which verbs relating to *feelings and senses* and *communication mediums and modes* were selected: 11.7% and 10% of PR verbs respectively (the fourth and fifth most popular topics – see Table 5.12). This compares to Palmer et al. (2017) where *nouns* relating to feelings and senses were chosen 3% of the time, and nouns relating to communication mediums and modes 2% of the time. The importance of targeting more abstract verbs in treatment is attested to by P2 who explained that she felt like she had "lost her voice at the dining table" and wanted to be able to express her opinions more. She chose verbs such as “aim, hate, hope” and “think” in relation to this. If we are to restore the voices of people with aphasia in this sense – and PwA express this desire strongly (e.g., Wallace et al, 2017) – it is imperative that they are given the opportunity to choose abstract verbs to work on in treatment as well as more concrete ones.

In terms of the *process* of selecting participants’ PR verbs, only P2’s partner was proactive in helping her to select her verbs. This meant that the research student was very actively involved in working together with the other five participants to identify suitable PR verbs, and this leads to the possibility that this may have biased/influenced their choice of verbs. However, it is reassuring to note that the verb topics chosen by P2 together with her partner overlapped considerably with those chosen by the other five participants (together with the research student), the only topic unique to P2 being money (with two verbs selected: “save” and “withdraw” – see Table 5.13). That the involvement of the research student did not unduly restrict the selection of PR verbs by participants is also supported by the range of verbs chosen: the six participants chose 120 different verbs (out of a possible 240 total) meaning that 50% of verbs were unique to a participant. Thus, whilst the possibility that the involvement of the researcher in choosing PR words for five of the six participants biased their choice of words remains, this finding would suggest that participants still chose a diverse range of verbs.

In summary, it is a unique feature of this study that it targeted PR *verbs* exclusively, using self-delivered computer-based treatment. Whilst Palmer et al. (2019) also targeted all PR words in their RCT of self-delivered computer-based treatment, these comprised mostly nouns. Indeed, it is interesting to note that Palmer et al’s 100 participants only chose verbs in their set of 100 treated PR words on 2.3% of occasions (Palmer et al., 2017). The reasons for the paucity of verbs is not clear but may lie in Palmer et al’s use of picture prompts to assist participants in choosing their target words. It is not reported if these prompts depicted any verbs, and, as discussed above, the use of picture prompts runs the risk of biasing selection towards highly imageable and more concrete words (i.e., nouns), because verbs are less imageable and less concrete. That verbs should be targeted more frequently in aphasia treatment is underlined by their frequency in spoken language, as highlighted by Renvall et al. (2013a & b). Thus, it is vital that clinicians – and researchers – ensure that PwA are given equal opportunity to select verbs as well as nouns for treatment.

Finally, the lack of generalisation of verb treatments in particular (to untrained verbs for only 15% of participants (e.g., Hickin et al., 2020)) emphasises the importance of verbs targeted in treatment being *personally relevant*, and the results of this study suggest that choosing PR verbs is feasible. Because an eclectic approach was used to select the PR verbs it is not possible to identify which parts of the process were most important to the successful selection of PR verbs, and indeed it may be that the broad range of strategies used was what made the process feasible. It should also be noted that whilst the selection process was eclectic it was not time-consuming, generally taking about 30 minutes of a baseline assessment session. However, in a future study, adopting a more rigorous process for noting the time taken to discuss and record PR verbs, including any discarded, might provide more concrete information to guide further research. Finally, it is possible that the PR nature of the verbs selected for treatment may have contributed to the acceptability of and compliance with treatment and this will be discussed in the following two sections.

7.3. The acceptability of the SPT to the participants with aphasia and their significant others.

7.3.i Summary.

The six participants and the two SOs who completed exit interviews were very positive about the SPT, with very few disadvantages identified. This is in line with the findings of Kearns et al. (2019) who reviewed 27 studies that included self-report of the acceptability of computer-based aphasia treatment. The acceptability of the SPT likely reflects the influence of several aspects of the SPT programme. Firstly, its aphasia-friendly design meant it was largely very accessible and hence presumably acceptable. Second, the targeting of PR verbs in the program likely increased the saliency of the stimuli and therefore the SPT itself, helping participants to stay interested and motivated. This finding is in line with those of the Big CACTUS study where both participants with aphasia and SLTs implementing the treatment regarded personalisation of treatment as important to motivation (Harrison et al., 2021, Burke et al., 2021 respectively). Third, the SPT was related to participants' goals: Kearns et al. (2019) found this to be very important to the acceptability of computer-based treatment. Fourth, the SPT included weekly contact with the research student. Harrison et al. (ibid) found a positive correlation between the amount of SLT/SLTA support given to participants and the amount of self-managed treatment, and both participants with aphasia and carers regarded this as an important factor in making the self-managed computer-based treatment acceptable.

Harrison et al. also found a relationship between time post-stroke and degree of compliance with treatment, theorizing that the underlying reasons for this were likely a combination of reduced access to face-to-face treatment at the chronic stage increasing the acceptability of self-managed computer-based treatment, whilst reduced commitments generally (e.g., with other health related appointments) may have given participants more time to practice. This suggests that PwA at the chronic stage are likely to find the SPT most acceptable. When interviewing SLTs who implemented treatment in the Big CACTUS study (Burke et al., 2021), therapists also reported that they regarded PwA who had some proficiency with technology as most likely to find computer-based self-managed treatment acceptable, as well as those who had insight into their condition and were self-motivated (e.g., as opposed to a carer being more motivated). This suggests that PwA who have some experience with technology, have a good understanding of the impact of their aphasia and who are self-motivated may find the SPT most acceptable.

7.3.ii Detailed Discussion.

The acceptability of the SPT to participants and their SOs was investigated via the use of a rating scale questionnaire for the participants (based on that used by Palmer et al., 2013), together with open-ended questions about the perceived benefits and disadvantages of self-delivered computer treatment (ibid). Two SOs (of P2 and P3) were also interviewed about the acceptability of the SPT and its perceived benefits and disadvantages using a topic guide again based on Palmer et al. (2013).

Overall, both participants and carers were very positive about the acceptability of the SPT. Very few disadvantages were mentioned, and these mainly related to the *perceived* superiority of face-to-face treatment. The only other disadvantages reported were technical issues (by P1 and P5 accurately reflecting their experience), whilst P4 wanted treatment to be even more intensive.

In terms of perceived advantages of the SPT, three participants reported improvements in confidence. All participants felt that treatment had worked, with four participants reporting *a lot of change* in their language as a result of treatment, and three reporting *some change*. All participants also reported that they had used their PR verbs in daily communication, with four reporting this to be the case *every day* and two on *most days.* The acceptability of the self-delivered computer-based treatment investigated in this study is in line with previous findings (Amaya et al., 2018; Kearns et al., 2019; Palmer et al., 2013; Wade et al., 2003). Themes common to the previous studies and the current study are perceived improvements in confidence and in language as a result of computer-based aphasia treatment. Participants perceived there to be very few disadvantages of the SPT. It was not possible to explore the reasons for participants' positive views of the treatment in detail due to time and ethical constraints (that is ethical approval for the PhD study required that time taken for outcome measurement be kept to the minimum possible). However, their positive responses to questions about the difficulty level of self-delivering the treatment and how much help they needed indicate that this contributed to the overall positive perception of the SPT. It should be noted at this point that all six questions in the rating sale questionnaire were phrased neutrally (e.g., How did you find doing therapy on a computer? How much help did you need with the computer therapy? – see Appendix F) avoiding the positive bias identified in some of the acceptability questionnaires used in the studies reviewed by Kearns et al. (2019). However, it should also be noted that it was the research student who administered the exit questionnaire and conducted the exit interview, and this may have resulted in a positive bias. A future investigation of the SPT would incorporate the use of an independent assessor to conduct exit interviews to avoid this.

Having noted this proviso, the ease of use of the SPT reported by the participants was likely influenced by its aphasia-friendly design. As discussed in Chapter 3 (sections 3.5.ii and 3.5.iii), the literature pertaining to the *design* of *accessible, aphasia-friendly* HCIs, and to the features which influence the *acceptability* of computer-based treatment to PwA, their SOs and SLTs was reviewed so that it could inform the design of the SPT. Thus, the SPT incorporated the aphasia-friendly features of *multimodality, large buttons, stable interface*, *simple navigation* and *visual simplicity (*as recommended by Brandenberg et al., 2013). The SPT also minimised interaction between the PwA and the SPT and allowed the PwA to control the pace (aphasia-friendly features recommended by Wilson et al. (n.d.). These features appear to have been effective in making the SPT accessible to the majority of the six participants. Indeed, the reason that only three participants were fully independent in *uploading* the SPT exercises may have been that this required the use of aspects of the HCI which could *not* be made aphasia-friendly (e.g., downloading files from emails, or copying files across using Windows). The accessibility of the SPT exercises themselves presumably contributed to the acceptability of the SPT. Acceptability was also likely influenced by the use of PR stimuli a feature requested by SLTs (Swales et al., 2016)), and by tailoring the SPT to address goals jointly formulated with participants, the latter being identified as important to the acceptability of computer-based treatment by Kearns et al. (2019). Regarding P5, who had the most difficulty accessing the SPT, future implementations of the treatment would incorporate a more structured approach to training participants. This would include checking that they are able to independently self-deliver a trial treatment session between baseline assessment sessions, rather than observing them self-deliver treatment during a face-to-face baseline assessment session (as was the case with this study), plus specific training in downloading SPT exercises.

7.4 Compliance with the SPT.

As discussed in Chapter 3, the factors which influence compliance with self-delivery of computer-based aphasia treatment are poorly understood, at least in part because so little research has investigated this issue. However, in a recently published paper, Harrison et al. (2020) shed some much-needed light on this neglected area. They investigated the factors which influenced adherence to computer-based treatment for the 85 participants in the intervention arm of their Big CACTUS RCT. Important findings included that there was a significant positive correlation between amount of practice and: years post stroke; length of access to computer treatment; and the amount of therapist/volunteer support. Some interesting (non-significant) trends were also noted. These included that PwA using their *own* devices tended to practice more, as did those using *more portable* devices (e.g., a tablet compared to a desktop), and that PwA aged *56-65* tended to practice *more* than those *>76*. Whilst Harrison et al. did not assess the relationship between *ability to use technology* and compliance per se, they note its important influence, and consequently that of the *availability of support* from an informal carer for those participants who could not use the computer-based treatment independently. *Personalisation* of vocabulary, *alignment* between personal *goals* and the *intended outcome* of treatment, and a sense of *self-efficacy* in being able to self-deliver treatment were noted as important motivational factors. *Periods of illness* and a *lack of a PwA clearly understanding* the impact of their aphasia on their communication were noted as barriers to practice.

The findings of this small-scale feasibility study regarding the factors influencing compliance with the self-delivered, computer-based SPT are in line with those of Harrison et al. First, in terms of actual compliance with the SPT, four of the six participants (66%) were compliant with the total amount of self-delivered treatment requested (16 hours), and three participants delivered it to the requested intensity of at least two hours a week. The compliance rate is similar to that reported in Palmer et al. (2012) (66.7%) but lower than that in Kurland et al. (2018) (83%). The factors which appeared to interact to reduce the compliance of two participants (P1 and P5) were in line with those identified by Harrison et al. First, P1 and P5 had more limited *competence with technology* than the other four participants, with P5 in particular standing out as the only participant who was not a frequent user of technology when she started treatment (see Table 5.14.i)[25]. This meant that P1 was not immediately independent in self-delivering treatment (though he rapidly became so), whilst P5 only learnt to use the SPT independently after the end of the treatment phase (i.e., during the maintenance phase). Also of likely significance is that whilst P1 and P5 *needed help* to use the SPT, as did P4, only P4 had access to someone who could help (his wife). When P1 and P5 ran into difficulty with the SPT there was no one to assist them to overcome this[26]. Finally of likely relevance is the *health* of P1 and P5: both were hospitalised during the treatment phase (twice in the case of

[25] P5 was also by far the oldest participant in the study at >80 years.
[26] Both P1 and P5 did contact the research student by phone when they were having technical difficulties (as requested) but it was not possible to solve their problems remotely.

P5) and this also reduced their ability to comply with treatment. The failure of P1 and P5 to comply with the amount of treatment requested reflected the interaction of three factors namely - poorer competence with technology, lack of support to solve technical issues, and poor health.

In terms of the impact of *aphasia severity* on compliance with treatment, there is a concern that more severe aphasia could limit a PwA's ability to comply with self-delivered treatment. However, Harrison et al. (2020) and Kurland et al. (2018) found no relationship between aphasia severity and compliance, whilst Des Roches et al. (2015) found a negative correlation with more severely impaired participants practicing more often. In this study there was no apparent relationship between compliance and aphasia severity since the participant who had the mildest aphasia practised the least (P5), whilst the two participants with the highest levels of compliance had moderate aphasia (P2 and P6). However, the small number of participants in this study together with the interaction between several likely influential factors (e.g., age and health as well as aphasia severity) means no clear conclusions can be drawn.

In summary, it is highly likely that several different factors *interact* to influence compliance with self-delivered computer-based treatment, and these probably include years post-stroke, physical health, support from a therapist and from volunteers/assistants, competence with technology, the availability of informal technical support, personalisation of treatment and the perceived relevance of treatment to a PwA's goals. To improve the compliance of PwA with self-delivered computer-based treatment, it will be necessary to systematically chart *all* of the factors which potentially influence compliance, and the study by Harrison et al. (2020) is an important step forward in this regard. Models used within the technology industry to represent the factors which influence the take up of new technology are also likely to be useful, as discussed by Kearns et al. (2019). For example, the degree to which a person perceives that using technology will allow them to *improve their skills* is seen to be the most important predictor of the take up of that technology.

This speaks to the need to ensure that PwA understand very clearly how any self-delivered computer-based treatment will help them to *improve their language,* and vindicates the use of a goal action planning process in this feasibility study, to ensure alignment between a PwA's goals and those set in relation to the SPT. Other factors identified by the technology industry as influencing compliance are: ease of use (including both how easy it is to *learn to use* technology and *remember* this – underlining the importance of the evidence-based, aphasia-friendly design of the SPT PowerPoint exercise slides), social factors (i.e. how important *significant others* think it is for a PwA to use the technology – carers were involved in initial discussions about the SPT with the consent of the PwA) and facilitating factors (i.e. the availability of *support* for technological problems). All of these factors are now recognised to likely determine compliance with computer-based aphasia treatment, with relevant research consequently emerging. For example, there are studies relating to the design of aphasia-friendly HCIs (e.g., Roper, Davey, Wilson, Neate, Marshall & Grellmann, 2018), the views of carers on computer-based treatments (e.g., Harrison et al., 2020), and the feasibility of supporting self-delivered computer-based treatment remotely (e.g., Braley et al., 2021). A final recommendation is that, whenever possible, the amount of treatment self-delivered (i.e., compliance) is monitored remotely by computer-based treatment which is programmed to do so (e.g., Kurland et al., 2018).

Lastly, it is interesting to note that three of the six participants (50%) continued to use the SPT during the maintenance phase and this included P5 who finally taught herself to use the SPT independently during this phase. Palmer et al. (2019) report that 61% of their participants (n=57) continued to use the StepByStep treatment program beyond maintenance assessment. As noted in Chapter 3 this speaks to the determination of PwA to overcome any difficulties in accessing technology so that they can continue to proactively self-manage their treatment.

7.5 The fidelity of the Sentence Production Treatment.

A number of reviews have identified that aphasia treatment studies have neglected to assess treatment fidelity (e.g., Brogan, Ciccone & Godecke, 2019; Dipper, Franklin, de Aguiar, Baumgaertner, Brady, et al., 2021; Hinckley and Douglas, 2013). This has resulted in the recommendation that future research should attend to treatment fidelity in terms of treatment delivery, receipt and enactment and the SPT will now be discussed in relation to these.

In terms of treatment delivery and enactment, a treatment manual was written for the SPT (available as supplementary material) and this was used to develop fidelity checklists for treatment delivered face-to-face (Appendix B). Face-to-face treatment was delivered to an average of 98% adherence to the treatment protocol described in the fidelity checklists (range 92-100%). This is similar to the levels of adherence to the VNeST treatment protocol reported by Edmonds and colleagues in a series of studies (e.g., Edmonds et al., 2009; Edmonds & Babb 2011; Edmonds et al., 2014) who reported 98%-99% adherence. However, it should be noted that the research student rated her own adherence to the treatment protocol, whereas ideally this would be carried out by someone independent of the study, and this would be part of the design of any future research study investigating the SPT.

In terms of treatment *receipt*, the total amount of face-to-face treatment delivered was close to the target of 8 sessions per participant at 97% i.e., a total of 46.5 sessions out of the intended total of 48 sessions were delivered (with 6.5 sessions being delivered to P4 instead of eight sessions due to the illness of P4's mother). (The amount of treatment self-delivered (i.e., compliance) is discussed in the preceding section and will not be discussed further here). In terms of treatment receipt, participants' views of the SPT are reported in section 7.3 above and again will not be discussed further other than to note that the SPT is currently the only sentence treatment – computer-based or face-to-face – which has solicited participants' views.

In summary, the study found that the SPT treatment was feasible in terms of recruitment and retention of participants, and most participants (5/6) were able to self-manage treatment independently. This is the first evidence that *self-delivery* of *sentence* level treatment via *computer-based* exercises is feasible: Hickin et al. (2022) conducted a systematic scoping review of 33 studies of sentence treatments and found none was self-delivered via computer. The SPT was found to be acceptable to participants and their significant others with very few disadvantages identified. Given that sentence treatments are acknowledged to be complex (see e.g., Conroy et al., 2006; Conroy et al., 2009c) this may have meant that they were not amenable to computer-based self-deliver. This PhD study provides preliminary evidence that this is not the case and that this mode of delivering sentence treatments is worthy of further and larger scale exploration. Four out of the six participants complied with the expected amount of additional self-delivered treatment, three of these exceeding this amount significantly. This level of compliance is similar to that for noun and verb treatments self-delivered by computer (e.g., Palmer et al., 2012; Kurland et al., 2018) and again represents the first, preliminary evidence that sentence level treatments, despite their complexity, can be successfully self-delivered. Conclusions regarding factors which influenced compliance pointed to the importance of being a current user of technology and of having someone available to assist with technical issues. Finally, the clinician delivered component of treatment was delivered to a high standard of fidelity (98%).

Future studies of the SPT would investigate the fidelity of the clinician-delivered component of the SPT using an independent assessor. A future study would also investigate the fidelity of the *self-delivered* component of the SPT to establish how often and how successfully participants were using the cues provided in the SPT exercises, and to consequently adapt and refine these. To expand, exploration of the fidelity of self-delivered, computer-based aphasia treatment is an emerging field of research. In a pioneering study, Ball, De Riesthal and Steele (2018) explored the degree to which four participants complied with recommended treatment procedures (treatment fidelity) during self-delivered computer-based anomia treatment, including whether adherence influenced accuracy of performance. Sessions were video recorded to monitor participants' interaction with the computer-based treatment. Treatment fidelity was maintained in only 45-61% of sessions, and interestingly the number of successful naming attempts was higher in naming attempts where fidelity wasn't maintained than when it was (77% - 93% versus 47% - 61%). The authors speculate that participants identified whether a particular cue was necessary for successful naming, and consequently modified the treatment protocol to maximise their success levels. Des Roches, Mitko and Kiran (2017) also carried out a study pertinent to the fidelity of self-delivered computer-based aphasia treatment. They report a detailed analysis of how participants in their 2015 study (Des Roches et al., 2015) interacted with Constant Therapy. Specifically, they investigated the therapeutic activity of "client acts" (as defined by Baker, 2012a) by establishing how participants' used cues in Constant Therapy exercises. They found that participants who used more cues tended to use these less successfully, and that this was negatively correlated with their aphasia severity (i.e., participants with more severe aphasia used more cues less successfully and vice versa). Thus, in terms of Baker's model, cue use by those with more severe aphasia meant that potentially a less effective dose of treatment was being self-delivered than by those with milder aphasia. Although both more and less severe participants showed reduced cue use and increased accuracy of use over time (i.e., both groups began to self-deliver their treatment more effectively), this disparity remained. The studies by both Des Roches et al. and Ball et al. speak to the need to monitor cue use and modify cues provided as treatment progresses. In terms of Baker's model, this would ensure that an effective dose of treatment is being consistently delivered. Thus, in future investigations of the feasibility and efficacy of the SPT the use of cues during treatment would be monitored, ideally with an inbuilt facility such as that incorporated into StepByStep (e.g., Palmer et al., 2019) and Constant Therapy (e.g., Des Roches et al., 2017). This would allow the efficacy of cues to be investigated, and consequent refinement of the SPT treatment protocol.

7.6 Preliminary efficacy testing of the effect of the SPT on the production of i) trained and untrained verbs, ii) untrained nouns, iii) sentence production using trained and untrained verbs, iv) verb and sentence production in discourse and v) in functional communication as perceived by i) the participants with aphasia themselves and ii) their significant others.

The effects of the SPT on each level of language will be described in turn for each participant, with potential influential factors discussed.

7.6.i The effect of the SPT treatment on the language of P1.

P1 had a moderate non-fluent aphasia with moderately impaired noun and verb production. He had eight face-to-face treatment sessions over 11 weeks and self-managed a total of 11 hours of treatment during this period.

P1 responded well to the SPT, showing statistically significant improvements in the production of personally relevant verbs in isolation and in sentences, and this applied to both trained and untrained verbs (within level generalization). This improvement reflected increased access to the verb form leading to reduced production of a verb argument instead of the verb. P1 also showed statistically significant improvements in predicate argument structure using both trained and untrained verbs, and this reflected faster sentence production. This likely reflects the generalised improvement in his verb retrieval (discussed above), which consequently facilitates the (faster) retrieval of the arguments of the sentence via spreading activation (Collins & Loftus, 1975). That treatment produced changes in both trained and untrained verb and sentence production may be due to P1 having moderately (rather than severely) impaired verb and noun naming, as the (limited) evidence suggests that this is likely to increase responsiveness to sentence treatment (e.g., Schwartz et al., 1994; Edmonds et al., 2015). However, P1's production of untrained personally irrelevant verbs remained unchanged both in isolation and in sentences. This may reflect the lower salience of the personally irrelevant verbs as there is (again limited) evidence suggesting that highly salient stimuli may respond better to treatment (e.g., Kleim & Jones, 2008).

In terms of discourse, P1 demonstrated a trend in improved PAS (in CAT picture description) but this was not statistically significant, and there were no changes in narrative (AMI) production. The lack of change in narrative production may be explained by a fatigue effect, as P1 produced fewer utterances on each occasion of doing the AMI. In other words, how a sample of discourse was captured may not have been sensitive to capturing any changes. It is also of note that the Generalisation Phase of treatment during which discourse tasks were practiced constituted only two weeks of the eight weeks of the SPT programme for P1. This amount of treatment may not have been enough to produce a change in discourse. Finally, P1 showed no significant change in his CETI scores, and this may reflect the lack of change in his discourse production. It may also reflect P1's lack of engagement with the CETI, his dislike of which he wished to be recorded at his post-treatment interview.

7.6.ii The effect of the SPT treatment on the language of P2.

Turning now to P2, she also had a moderate non-fluent aphasia with moderately impaired noun and verb production. P2 had eight face-to-face treatment sessions over 11 weeks and self-managed a total of 59 hours during this period. P2 also continued to practice for an unspecified amount of time between T3 and T4 and therefore this does not constitute a maintenance period for her.

In terms of her response to the SPT, P2 showed a statistically significant improvement in the production of both trained and untrained PR verbs in isolation. However, the effect on untrained verbs only appeared at T4 suggesting that the extra treatment P2 self-managed between T3 and T4 was operative here. In relation to P2's improved production of the total set of verbs in isolation, this improved by 150% (i.e., relative to initial performance). P2's good repetition skills likely contributed to the success of treatment. This is in line with the findings of de Aguiar et al. (2016) who found that participants who were able to repeat words with >49% accuracy were more likely to improve as a result of verb treatment. However, given that participants needed to be able to repeat at least 75% of words accurately to be included in the research project, this does not, by itself, explain why P2's verb production improved so much. The most likely additional contributing factor is the amount of self-managed treatment P2 administered during the treatment phase (59 hours) together with an unspecified amount of practice between T3 and T4.

In terms of sentence production P2 showed a statistically significant improvement in the production of verbs, agents, and objects which was restricted to trained verbs. This drove an improvement in PAS: the production of PR verbs with two arguments increased, but this was also restricted to trained verbs. This improvement reflected quicker production of sentences, resulting in less instances of P2 being unable to produce any output within 10 seconds and enabled her to produce a sentence with a verb with two arguments (as opposed to an agent only or an agent with a verb). P2 showed no improvement in sentence production using either untrained PR or personally irrelevant verbs. This is hard to explain since untrained PR verb production in isolation improved, yet this did not drive an improvement in sentence production (unlike for P1). A possible reason for this is that more treatment was required to produce improvement in untrained sentence production (hinted at by the improvement in untrained verb production in isolation which only occurred after the additional treatment P2 self-managed between T3 and T4).

In terms of discourse production, P2 showed some improvement in PAS: in picture description her production of utterances containing a verb together with an argument increased, and the production of verbs with complex phrases was emerging. P2's CAT Picture Description scores for syntactic complexity increased reflected her emerging use of more complex verb phrases, and her improved scores for grammatical well-formedness reflected more accurate use of verb inflections, prepositions and determiners. In terms of the latter, the SPT did not aim to improve the production of verb morphology or sentence grammar, and this was made explicit to participants: they were asked to concentrate on the who – verb – what elements of sentences and guided to “not to worry about the little words.” However, P2 (who, as a teacher, had taught grammar) was determined throughout her participation in the research to produce as grammatically correct sentences as she could. She was acutely aware of when she made a grammatical error during face-to-face sessions and would ask for feedback. As part of her self-management of treatment P2 practiced PR verb sentences by writing them out, and these sentences also demonstrate attention to and self-correction of verb morphology. This is the likely reason for her improved grammatical scores following treatment. In terms of her functional communication, both P2 and her SO rated her communication better on the CETI post-treatment but at T3 only, even though P2 maintained the improvements she had made in her language production at T4. This may reflect the lack of face-to-face contact with the research student between T3 and T4: P2 enjoyed the face-to-face treatment sessions a lot (with her partner commenting that he could hear gales of laughter drifting down the garden during these sessions).

7.6.iii The effect of the SPT treatment on the language of P3.

P3’s results will now be discussed. P3 had a more severe nonfluent aphasia. Her expressive language in everyday communication was very effortful, and she also had a mild apraxia of speech. P3 had eight face-to-face treatment sessions over nine weeks and self-managed a total of 18 hours of treatment during this period.

P3 demonstrated a statistically significant improvement in the production of PR verbs in isolation immediately post-treatment which was restricted to trained verbs. The lack of within level generalisation for P3 may have resulted from the nature of her naming impairment. P3's commonest naming error was delayed production (>10 seconds) or the production of a phonological paraphasia (likely caused by her mild dyspraxia) and she was phonemically cueable. This suggests that she had difficulty accessing phonology from an intact semantic system (i.e., that her anomia was post-semantic) thus any effect of treatment would likely be specific to trained items as it would strengthen mapping between semantics and phonology on a lexical basis (as suggested by Howard (2000) in attempting to explain the lack of within level generalization of anomia treatment).

Contrary to the lack of generalisation in verb retrieval, P3 showed a significant change in untrained noun production (on the OANB). Her improvement reflected both quicker naming and a reduction in phonological paraphasias, and this is hard to explain in the light of P3's post semantic anomia. It may, however, reflect increased awareness of her phonological errors and/or an improvement in her ability to edit them prior to production as a result of repeated practice and self-monitoring whilst doing the SPT (i.e., improvement in her dyspraxia). P3 had particular difficulty with consonant clusters involving /s/ and those involving the juxtaposition of alveolar and velar consonants (e.g., stool became school, stamp -> /sl sk spa/, picnic -> /piklic/ or /pikenik/) and this type of cluster occurred much more frequently in the OANB nouns than either the OANB verbs or her PR verbs. This meant that 42% of her errors on OANB noun naming were phonological compared to only 15% of her errors on OANB (matched) verb naming. Thus, her improved ability to correct these phonological errors had a disproportionate effect on her noun naming as opposed to her verb naming. That P3's naming of both verbs and nouns improved as a result of treatment is perhaps surprising given that her repetition was impaired and therefore her ability to benefit from the sound cues provided as part of the SPT was likely reduced. She did, however, have very good ability to read aloud, and this enabled her to make good use of the written cues in the SPT and thus benefit from treatment.

In terms of sentence production using PR verbs, P3's improvement reflected an overall increase in the speed at which she was able to produce sentences. In terms of lexical retrieval in sentences, P3's production of trained PR verbs and agents improved significantly, resulting in a significant improvement in the production of verbs with one argument immediately post-treatment, and this was maintained. Untrained verbs and agents also improved significantly but only at maintenance suggesting that effects of treatment were slower to "trickle through" to untrained items. The production of objects had been at floor for P3 prior to treatment but started to emerge immediately post treatment for trained verbs. This resulted in a trend of improved production of verbs with two arguments, with again a treatment effect being slower to emerge for untrained objects. If P3's production of arguments after the time limit of 10 seconds is considered there was a more notable increase in production of objects: they were only produced on one occasion at T1 and two at T2 increasing to eight at T3 and nine at T4 (with the second argument almost exclusively being a theme as was targeted during treatment).

The improvement in the production of objects likely reflected improved syntactic processing for P3. Specifically, P3 was the only participant who did not show a syntactic bootstrapping effect in PR verb production. That is, she was the only participant who successfully retrieved more verbs in isolation than in sentences. This suggests that the root of P3's difficulties with sentence production did not lie in lexical retrieval *per se*, but in allocating successfully retrieved items to the correct thematic role in the sentence, or, as Marshall (1995) would view it, P3 had a procedural rather than a lexical mapping deficit. This theory is supported by P3 having mildly impaired retrieval of both nouns and verbs (on the OANB) suggesting that problems with lexical retrieval were indeed not the main cause of her sentence production difficulties. That P3 did have problems with mapping is supported by her performance in sentence comprehension where she had difficulty understanding reversible sentences. The mechanism by which SPT seems to have worked for P3 therefore, is to enable her to map syntactic structures onto thematic roles more quickly and this meant that treatment generalized to production of sentences using untrained verbs. However, as Mitchum et al. (2000) argue, evidence of improved mapping in production of structures where arguments have moved (such as *passive* sentences) would constitute stronger evidence of this. That improvements in sentence production using personally irrelevant verbs were less than for PR verbs suggests that the saliency of the PR verbs increased the potency of the SPT.

In terms of *discourse*, the most notable changes in PAS for P3 were in production of verbs together with two arguments, which increased in both picture description and narrative production. This contrasts to the pattern of improvement in *sentence* production where verbs with one argument improved. The reason for this relates to the speed at which P3 produced sentences during assessment of sentence production (as discussed above). Thus, P3's production of verbs with two arguments in sentences did improve but because of the time limit this was not credited whereas, because there was no time limit in assessment of discourse, her production of verbs with two arguments was credited. Finally of note in P3's discourse is the emergence of the production of verbs in complex phrases in her personal narrative production only (e.g., I had a friend who live in Spain). That such phrases were not used in picture description probably reflects the nature of the task, i.e., it does not require the use of complex language. This speaks to the need to use suitable measures (e.g., that are of interest and motivating) to stimulate the use of complex language as noted by Hickin, Mehta and Dipper (2015). Lastly, both P3 and her SO rated her functional communication as significantly better post-treatment and at maintenance suggesting that for P3 treatment had an impact on communication in real life.

7.6.iv The effect of the SPT treatment on the language of P4.

P4 had a moderate – severe fluent aphasia. He had moderately impaired spoken expression but a severe impairment on many of the CAT subtests. P4 had six face-to-face treatment sessions over 8 weeks, and self-managed a total of just under 29 hours of treatment during this period.

The impact of treatment on P4's production of words in isolation was minimal: trained PR verbs showed significant improvement but only immediately post treatment, and there was no improvement in the production of untrained PR verbs. Untrained personally irrelevant verbs did not improve, whilst nouns showed a trend of improvement immediately post-treatment only. P4's relatively poor response to treatment in terms of his lexical retrieval may be due to the severity of his impairment. According to the criteria Palmer et al. (2019) used in their RCT of self-managed, computer delivered treatment, P4 would be categorized as having both a severe naming impairment (he scored 16/48 on CAT Object Naming). He was also outside of normal limits on (CAT) spoken sentence comprehension (21/32). Palmer et al. note that of their 97 participants, those whose naming impairment was more severe and whose comprehension was outside of normal limits responded less well to treatment. P4 was also very poor at repeating words. Conroy et al. (2009a, b & c) suggest that good repetition is an important factor in determining response to treatment, and, at 75%, P4 is on the threshold of their acceptance criteria. In terms of the trend of improved untrained noun production, P4 had the most impaired noun retrieval and was the only participant who was more impaired on noun than verb production. So, perhaps there was more potential for nouns to improve than verbs. Also of potential relevance is that P4's difficulties with noun retrieval were very evident in his attempts at sentence production, and so for him, practicing nouns during the SPT may have been just as active an ingredient as practicing verbs (perhaps even more so) and this may account for the improvement in his noun retrieval. However, given that the

OANB nouns were untrained why should treatment have an effect? This may be explained by the nature of P4's anomia. P4 made both semantic and phonological errors in his naming attempts (e.g. nest -> bird; nun -> church woman; brain -> bread; box -> /bok/) with his commonest error being no response. This suggests that he had multiple deficits (in the semantic system and in access to/within the phonological output lexicon) underlying his anomia. Treatment may therefore have improved production of untrained nouns via a generic effect on semantics and/or phonological processing.

In terms of the effect of treatment on P4's lexical retrieval in sentences, the only effect of treatment was on his production of agents. It has been suggested that an appreciation of agency is an important aspect of both event processing and of verb meaning (Cairns et al., 2007). Prior to treatment, P4's attempts at sentence production strongly suggested that P4 had difficulty identifying the agent of an event, and indeed that he had difficulty with event processing per se. Cairns et al. (2007) suggest that sentence treatments "might help (participants) to maintain a useful focus over events, for instance by 'anchoring' their attention while the object of focus is fitted to available language" (p.231). The SPT slides were animated to always focus the participant on the verb first, followed by the agent and then the theme (and finally an adjunct if this was being trained) and so it is certainly possible that the SPT worked in this manner. It was not anticipated that the SPT would work like this since the focus on treated verbs, agents and themes was informed by the sentence treatment literature (particularly mapping treatment studies and those of VNeST) not the small treatment literature relating to event processing (e.g., Cairns et al., ibid). However, as has been noted many times before (e.g., Byng et al., 1994; Howard, 2000; Marshall, 1995) psycho-linguistically motivated treatments for aphasia do not always work in the theorised manner.

In terms of PAS in sentences, there was a small trend of improvement in the production of PR verbs with two arguments after treatment (accompanied by a trend for reduced production of PR verbs with one argument). In terms of PAS relating to personally *irrelevant* verbs, there was a statistically significant improvement in the production of verbs with one argument. The improvement in the production of verbs together with their arguments likely reflects P4's better ability to conceptualise the events depicted in the pictures he was describing as discussed above. Why PR verbs and personally irrelevant verbs responded differently to treatment is not clear. It may however, reflect the greater *salience* of the PR verbs in that this may have made both the verbs and the events they depicted more interesting/motivating to describe thus adding to the "power" of treatment. Although the importance of saliency in improving the effectiveness of treatment is being increasingly acknowledged (e.g., Dignam et al., 2016; Kliem & Jones, 2008; Raymer et al., 2008), the actual mechanisms by which saliency works are not clear and thus the nature of saliency (i.e., what the requisite properties of saliency are) is also not clear. As noted by Raymer et al. (2008) the definition of saliency needs to be refined. Nevertheless, it seems likely that an essential property of salient stimuli is that they are perceived as meaningful (i.e., personally relevant) so that they evoke emotions for participants which promote neuroplastic change. Kleim and Jones (2008) allude to the evidence that emotions "modulate the strength of memory consolidation" (p.231) presumably because stronger emotions invoke more robust neuroplastic changes. Salient stimuli are also likely to induce greater motivation and attention and consequently better engagement in treatment which may also facilitate neuroplastic change (ibid).

Turning to P4's discourse production, there was evidence that treatment had some impact at this level of communication too. In picture description there was a small increase in the production of verbs together with two arguments (mirroring sentence production), with tentative indications of verbs being produced in more complex structures. In personal narrative, production of verbs together with an argument also increased, accompanied by a decrease in utterances with undetermined thematic structure, however this was evident immediately post treatment only. The results in relation to P4's production of personal narrative must be interpreted in the context of him producing almost twice the number of utterances immediately post-treatment (at T3) compared to the other three occasions of sampling discourse (hence the relative proportion of utterances was reported). This likely reason for this is that, uniquely at T3, P4 talked about what happened when he had his stroke, narrating how his wife had taken the decision to come back to the UK from France against his wishes. His narrative clearly provoked some powerful emotions (e.g., he said "But she forced me to come here. Actually, that was it. She she made her mind up. And all girls and boys (their children) said no. But she bought here.") Perhaps salience was again at play here, in that the powerful emotions which accompanied his stroke narrative drove P4's to produce more language. This speaks to the need to make not just treatment *stimuli* salient but also assessment (and crucially treatment) *tasks* salient as well (Raymer et al, 2008; Hickin et al., 2015). Despite the changes in P4's discourse and sentence production, he did not perceive any changes in his functional communication.

7.6.v The effect of the SPT treatment on the language of P5.

Turning to P5, she had a mild anomic aphasia. P5 had the most distributed treatment of the six participants: eight face-to-face treatment sessions over 17 weeks (due to participant illness). She also self-managed the least amount of treatment: a total of just under 7 hours of treatment during this period as she was unable to use the SPT independently. (P5 ultimately taught herself to use the SPT and carried on practicing her PR verbs using the SPT between T3 and T4 and so this is not a maintenance phase for her).

Considering P5's production of *PR* verbs in isolation first, trained PR verb production improved significantly immediately post-treatment only, whilst production of untrained PR and personally *irrelevant* verbs (and nouns) remained unchanged. It is possible that in relation to the latter a ceiling effect occurred as P5 named 88% of OANB verbs and nouns successfully at P2 (having improved since T1). In terms of P5's *lexical retrieval* in sentences using PR verbs, the only change resulting from treatment was a statistically significant improvement in the production of agents of untrained verbs. The lack of change in the production of agents of trained verbs may also reflect a ceiling effect (which was not present for untrained verbs because trained and untrained verbs were matched based on performance when naming verbs *in isolation* not in sentences). A ceiling effect may also have been in operation in relation to the lack of change in *agents* of untrained *irrelevant* verbs in isolation (in the SCT).

The impact of the SPT on P5's discourse production was minimal in terms of both picture description and personal narrative production. A possible explanation of this in relation to the AMI is that P5 showed a fatigue effect at T4 (producing just over half the number of utterances of previous AMIs). A more likely explanation, however, is that of a ceiling effect (as demonstrated in relation to lexical retrieval in isolation and in sentences for P5). Adding weight to this argument is that P5 was already producing a variety of sentence structures in her AMIs pre-treatment, giving little room for change. Indeed, P5 performed better in less constrained tasks such as narrative production (compared to verb in isolation and sentence production) because she appeared to be experiencing what Cahana- Amitay et al. (2011) have termed "linguistic anxiety." Cahana-Amitay et al. define linguistic anxiety as follows:

> "Specifically, a person with linguistic anxiety is one in whom the deliberate, effortful production of language involves anticipation of an error, with the imminence of linguistic failure serving as the threat" (Cahana-Amitay et al., 2011, p.593)

Cahana-Amitay et al. (2011) propose that PwA may show higher levels of anxiety in constrained language tasks (such as naming) than in more naturalistic tasks such as picture description or narrative production. This is because the risk of exposing their language impairment – and therefore of failure – is higher in constrained tasks, and this is indeed the pattern that P5 showed. At times P5 became tense and expressed nervousness during constrained assessment tasks, but never during CAT picture description or the AMI which, she stated, she enjoyed.

The limited impact of treatment for P5 on verb, sentence and discourse production is thus likely due to a combination of reasons. Her treatment was the most distributed and she also self-managed the smallest amount of treatment because she was unable to overcome technological barriers to using the SPT. P5's underlying linguistic deficits were also the mildest leading to some ceiling effects, and linguistic anxiety may have contributed significantly to P5's aphasia. Thus, the SPT likely did not target the source of her difficulties with sentence production. It may be that treatment at the level of discourse (such as that delivered by for example Whitworth (2010)) may have been more appropriate for P5. Regarding her linguistic anxiety, it may be that group treatment such as that provided by Community Aphasia Groups (e.g., Attard, Loupis, Togher & Rose, 2018) would have been more effective as it which would have given P5 the opportunity to both demonstrate and develop confidence in her linguistic skills in a supportive environment (as suggested by Cahana-Amitay et al. (2011).

Finally, despite the limited impact of treatment on her language production per se, P5 perceived that her functional communication had improved. This may also reflect the role of linguistic anxiety in her aphasia in that she reported in the exit interview that she felt she had benefitted from the treatment. This may have led to increased confidence in her language abilities, with a subsequent reduction in her anxiety. However, this hypothesis cannot be confirmed by any other evidence.

7.6.vi The effect of the SPT treatment on the language of P6.

The final participant, P6, had a moderate - severe nonfluent aphasia in that he had very little spontaneous verbal output but good comprehension. P6's spoken lexical retrieval was moderate - severely impaired, with verbs more impaired than nouns. P6 had eight face-to-face treatment sessions over 11 weeks and self-managed a total of over 45 hours of treatment during this period. P6 continued to practice his PR verbs between T3 and T4 (for an additional 12 hours) meaning that this does not represent a true maintenance phase.

P6 demonstrated a trend in improved production of both trained and untrained PR verbs which was greater for trained verbs. Although this change was not statistically significant, in terms of percentage improvement it represented a large change: 86% improvement of the total verbs at T3 and 57% at T4. P6's production of untrained personally irrelevant (OANB) verbs improved significantly, but at T4 only suggesting that the extra treatment P6 self-managed between T3 and T4 was operative here. That P6's lexical retrieval improved as a result of treatment despite the severity of his deficit is encouraging. Palmer et al. (2012) found that self-managed computer treatment did not benefit participants who scored <10% on the OANB, however P6 scored higher than this (42% for nouns and 21% for verbs). Palmer et al. (2019) in their larger RCT of the self-managed computer treatment found that benefits of the treatment were "slightly higher for participants with mild word finding difficulties and for those whose verbal comprehension was within normal limits" (p.830). P6 is regarded as being just within the category of having moderate word finding difficulties as he scored 18 on the CAT Object Naming Test with their criteria being 18 – 30. However, he scored 27/32 on spoken sentence comprehension and is thus regarded as within normal limits on this criterion. P6 was also extremely phonemically cueable and had excellent ability to repeat words meaning that he was able to benefit from the sound cues provided in the SPT. Treatment also appeared to teach P6 how to use his much better written naming to self-cue his spoken naming. The most likely mechanism for this is the juxtaposition of spoken and written cues in the SPT exercise slides, and his repeated use of these. The improvement in untrained personally irrelevant verbs may reflect increased P6's metalinguistic awareness of verbs because of treatment. This mechanism of improvement has been proposed by Carragher et al. (2015) who felt that earlier involvement in a verb treatment study may have increased metalinguistic awareness of their participants, who were subsequently able to benefit from sentence treatment. Marshall (1995) also proposes that some participants may benefit from treatment because it increases their psycholinguistic insight rather than reducing their impairment, and that this may explain why participants with similar impairments given similar treatments respond differently. Whilst it is not possible to reach a definite conclusion about why treatment showed some generalization to untrained verbs for P6, the improvement in his sentence production may also shed light on this.

It should first be noted that P6's production of PR verbs in sentences varied significantly at baseline. This is possibly related to the salience of P6's PR verbs as production of personally irrelevant verbs in the SCT did not improve at baseline. This explanation is supported by the amount of thought P6 had clearly devoted to his selection of PR verbs: for example, he pointed out to the research student that "play" would be a good verb to treat because it could be used not only in relation to the sports he enjoyed (as suggested by the student), but to the music and bands he loved as well. Despite the variation in sentence production at baseline, production of trained and untrained agents improved significantly as a result of treatment, and there was a trend of improvement for trained and untrained verb production in sentences, which was significant for the total set of verbs at T4 only. The latter likely reflects a cumulative effect of treatment, i.e., the severity of P6's deficit meant that the additional treatment P6 self-managed between T3 and T4 was required to produce a treatment effect. However, P6's production of trained and untrained *agents* improved significantly earlier (at T3), with improvement being significant for the total set of agents at both T3 and T4. The apparent earlier response of agents to treatment may reflect the primacy of agents in relation to the event processing required prior to sentence formulation. Marshall and Cairns (2005) demonstrated this effect in normal controls who showed a predisposition to name the agent first (followed by theme then instrument) when describing pictures in sentences, and this effect remained even when themes were also animate i.e., that this was an effect of agency rather than animacy. In terms of the effect of treatment on the production of *objects*, no objects were produced by P6 before treatment, but these were beginning to emerge post treatment (see discussion below). In terms of personally *irrelevant* verbs, agents, and objects there was little evidence of improved production: verb production increased post treatment, but this appeared to be at the cost of production of agents and objects which decreased. As with P3, the saliency of the PR verbs appears to have increased the potency of the SPT in improving sentence production.

Turning now to changes in P6's production of PAS in sentences, the most notable change is that prior to treatment P6 was not able to produce any PR verbs together with an argument within 10 seconds whereas after treatment he was. Additionally, if responses made *after 10 seconds* are analysed, he was also able to produce verbs with *two* arguments (agent + verb + theme) for the first time, doing so on four occasions (I drive the car; I like music; she is listening to music; she is roasting chicken) but this ability only emerged at T4. As mentioned earlier in relation to verb retrieval in sentences, this likely reflects a cumulative effect of treatment i.e., the extra treatment self-administered between T3 and T4 was required to enable P6 to produce a second argument. The smaller effect of treatment on objects may be due to the complete absence of objects in P6's output pre-treatment compared to agents which were evident though very minimally (on 2/80 occasions). The later response of objects to treatment may also be a consequence of the primacy of agents over objects in event processing as discussed earlier.

The possible mechanism of treatment in relation to P6's improved sentence production is hinted at by the few (four) attempts P6 made at producing PAS pre-treatment. These demonstrate both verb omissions ("guy is wine; man is pipe") suggesting that treatment may have worked by improving P6's ability to retrieve verbs. P6's other two errors are likely mapping errors: "driving the driver; eat the boy" suggesting that treatment may also have increased his knowledge of PAS i.e., how to map thematic roles onto syntactic structures. Indeed, it is possible that treatment exploited P6's covert knowledge of thematic role structure (as evidenced by his very good performance on sentence comprehension tasks), bringing it to conscious awareness. This is suggested by Cairns et al. (2007) as a potential mechanism of treatment. However, the limited improvement in PAS using personally irrelevant somewhat undermines the proposed explanation of improved awareness of PAS/knowledge of mapping. Once again, it may be the salient nature of P6's PR verbs that was at play here.

P6's discourse production was only assessed via picture description because the severity of P6's expressive difficulties made assessment of narrative production too distressing. Although P6's CAT picture description score improved slightly post treatment, this was entirely due to an increase in appropriate ICWs. There was no evidence of an impact of treatment on discourse production. The severity of P6's expressive impairment likely required more (and more intensive) treatment to produce an impact on discourse as is hinted at by improvement in sentence production occurring predominantly only at T4 (after additional self-managed treatment). Despite the lack of an impact on discourse, P6 rated his functional communication better post treatment but only at T3. This likely reflected P6's metalinguistic awareness of the improvements that he had made during the treatment itself. That is, that he had only been able to produce a verb together with an argument very occasionally before treatment but was able to do this both more often and more quickly after treatment, and that his ability to produce a sentence with two arguments was also emerging. He was rightly proud of the hard work (i.e., the amount of self-managed treatment) he had put in to make this progress.

7.7 Summary and discussion of preliminary efficacy testing.

Overall indicative key findings are now summarised. The SPT resulted in a significant improvement in PR *verb production in isolation* for five of the six participants, which was restricted to treated verbs for three participants. P6 showed a non-significant trend of improved PR verb production with the likely influential factor being the severity of his verb retrieval deficit, and it is possible that more treatment may have produced a significant improvement. *Sentence production* using PR verbs improved significantly for all six participants, but this was largely restricted to treated verbs, and changes were minimal for P4 and P5 who were the two participants with fluent aphasia. However, the SPT improved lexical retrieval in sentences, most commonly for agents including untreated agents (for five participants) followed by verbs and then objects, and this may reflect the primacy of agents in the early stages of sentence production (i.e., event processing: Cairns et al., 2007). There was also evidence of improved PAS (i.e., syntactic processing) for the four non-fluent participants. The impact of treatment on verb and sentence production using personally *irrelevant* verbs was minimal. This indicates that personal relevance (i.e., salience) *may* make words more responsive to aphasia treatments. However, much larger scale research is required to confirm this preliminary indicative finding.

The lack of studies investigating self-delivered, computer-based verb – and particularly sentence – treatments is likely based on the belief that verb and sentence treatments are too complex to be successfully delivered in this manner. This study therefore provides preliminary evidence that this is not the case and indicates that verb and sentence treatments self-delivered by computer are worthy of more investigation. In terms of the impact of the SPT on the production of verbs in isolation, trained PR verbs improved for all participants, and untrained PR verbs for two participants (with a trend of improvement for a third participant). This is also worthy of investigation since generalisation of self-delivered computer-based verb treatment has only been reported for one of 27 participants in which it has been investigated (Routhier et al., 2016 - see Table 3).

The preliminary efficacy of the SPT on *sentence production* using trained verbs was comparable to that of face-to-face treatments. In terms of improvement in untrained sentence production the SPT was somewhat less effective than face-to-face treatments and this may have been due to a smaller dose of sentence treatment being given in the SPT which meant that it was not powerful enough to generalise to untrained sentence production. Indeed, future implementations of the SPT would aim to tailor the treatment more specifically to individual level of impairment with the aim of ensuring that the appropriate dose of treatment at each phase was given. The tailoring would be specified *a priori* stating the "what, why, when and how" as recommended in the TIDieR checklist (Hoffman et al., 2014). Thus for example, P3 was the only participant not to show a syntactic bootstrapping effect in production of PR verbs in sentences, with her performance being worse than verb retrieval in isolation. This pattern of improvement indicated that her difficulties with sentence production were not due to poor lexical retrieval but were most likely syntactic in nature. Thus, for her it would likely have been best to spend most time in the sentence phase of the SPT. Indeed, this is what she did (Verb Phase: two sessions; Sentence Phase: three sessions; Generalisation Phase: 2.5 sessions - see Table 5.16) but it may have been better for her to have spent even more time in this phase. In a future implementation of the SPT note will be taken of any discrepancy in verb retrieval in isolation versus in sentences, and this information used to tailor the amount of treatment spent in the Verb and Sentence Phases of treatment respectively.

The impact of the SPT on *discourse* was the most limited, with only one participant showing a statistically significant improvement. However, four participants showed a trend of improvement, and it should be noted that discourse measures were not subjected to inferential statistical analysis (other than in relation to the CAT picture description score) because of the unconstrained nature of the tasks. The reasons for the more limited impact of the SPT on discourse are unclear but may relate to the self-delivery of treatment via computer giving fewer opportunities to practise discourse. The Generalisation Phase of the SPT was specifically designed to address this potential weakness but on average participants spent least time in this phase of treatment. This may have diluted the dose of treatment given to discourse production and rendered it less effective. The means used to gather a sample of discourse (the AMI) may also have been influential in that it likely produced a fatigue effect for five of the six participants. Thus, a future study of the SPT would use a different measure of discourse (e.g., The Scenario Test (van de Meulen et al., 2010)). In summary, it is not possible to reach any firm conclusions regarding the reasons for the more limited impact of the SPT on discourse and this would be a focus of future research.

Finally, the impact of the SPT on *functional communication* was promising. Four of the six participants perceived the SPT to have improved their functional communication and this was corroborated by the two SOs who completed the CETI. The perceived improvement in functional communication was backed up by improvements in language measures for all four participants, although for P5 the perceived improvement likely reflected an improvement in her *confidence* when communicating (as discussed earlier). This study therefore provides preliminary evidence that self-delivered computer-based treatment can have an impact on functional communication. This has only been formally investigated in one other study of self-delivered computer-based treatment. Palmer et al. (2019) found no impact on functional communication as measured by the COAST (Long et al., 2008) and highlight that developing self-delivered computer-treatments which have a functional impact should be an important priority of future research. The findings of this study represent a possible way forward here in that it may be the PR nature of the trained verbs that was key to producing a functional impact. However, Palmer et al. also targeted PR words which were in this case mostly nouns. It is therefore possible that the targeting of PR *verbs* was key, and/or that they were practised in PR tasks and/or that trained verbs included more abstract verbs. This latter possibility is hinted at by two of the four CETI questions which were perceived to have improved by at least three of the participants relating to communication activities which required the use of just such verbs - namely discussing emotions (question 4) and discussing or describing something in depth (question 16).

In terms of candidacy for the SPT, all findings are tentative given the small number of participants (n = 6). There are, however, indications that the SPT is more suitable for participants with non-fluent aphasia as the two participants with fluent aphasia (P4 and P5) showed least benefit from the SPT. There are also indications that the SPT is most suitable for people with relatively intact lexical retrieval and syntactic processing (P1 P2, and P3) – a finding in line with the research investigating VNeST (e.g., Edmonds et al., 2014), whilst people with a more severe expressive impairment might require more treatment than was provided in this study to show maximum benefit (P6). However, the SPT may have the potential to improve deficits at different levels of language processing namely, event processing (P4), lexical retrieval (e.g., P1 and P2) and syntactic processing (e.g., P3 and P6).

To clearly establish candidacy for the SPT, it will need to be trialed with a larger number of participants with non-fluent aphasia, and moderately impaired lexical retrieval and syntactic processing. Such a trial would also include a detailed survey of participants' strengths and deficits to establish the key skills which contribute to the success of treatment. Thus, for example, repetition has been identified as important in lexical retrieval treatment (Conroy et al., 2009a, b & c) and therefore all participants recruited to the SPT study had to have reasonable repetition skills. However, of the two participants with the worst repetition (P3 and P4), P4 responded poorly to treatment and P3 responded well. As already noted, P4's overall severity likely affected his ability to respond to treatment. P3 on the other hand was moderately impaired and in particular had good reading aloud skills. This enabled her to benefit from the written cues provided in the SPT, presumably compensating for her reduced ability to use sound cues. By way of contrast, P6 had severely impaired reading aloud meaning written cues were of limited use for him, but he had very good repetition and was also very phonemically cueable. This meant he was able use the sound cues provided by the SPT very well, presumably compensating for his reduced use of written cues. Thus, detailed assessment of participants' strengths and deficits is also likely to inform good candidacy for the SPT.

7.8 Limitations.

This study had several limitations, and these will now be discussed. Firstly, the study used pre-post studies of individual cases (n=6) and as such is not regarded as using single case methodology according to SCRIBE guidelines (Tate et al., 2016). This is because the lack of assessment during the intervention phase in pre-post studies weakens internal validity as it does not control for variables such as spontaneous recovery (ibid p.144). The study could have used a withdrawal design or a multiple baseline design (both of which are regarded as singe case methodology) although it is possible that there may have been ethical concerns about each of these designs because they would lengthen the amount of time participants were involved in the study (the actual study was shortened following feedback from the ethics committee). Whatever single case design is used, it should be noted that the results from single cases cannot be generalised because it is unclear why treatment was effective/ineffective (e.g., was it because of the treatment itself or something to do with the individual participant?) and this is another limitation.

An alternative design for this PhD study would have been to do a case series, where outcome measures and the treatment regime are specified in advance and applied to a series of single cases recruited to pre-determined criteria (Howard, Morris and Buerk, 2016). Each participant in a case series can be analysed as a single case, and the series of cases can be combined to yield a combined result which is more generalisable than that from a single case. However, in a case series the treatment regime must be rigidly applied. Given that this was a feasibility study exploring a prototype, complex intervention, the SPT was applied more flexibly to accommodate participants' capacity to practice and rate of progress through exercises for example. Immediate future studies will therefore continue to use single cases, to explore candidacy for the SPT and the potential active ingredients of treatment. Immediate studies will also include collaboration with PwA in a process of co-design (e.g., Kearns et al, 2019) to refine both the HCI of the SPT and its content to maximise its acceptability. Thus, feedback from participants in this feasibility study (during face-to-face visits) indicated that the exercises included in the Verb Phase of treatment were very acceptable. In particular, the first two exercises which used sound cues, letter cues and sentence closure resulted in high levels of success which was seen as important in building confidence in using the SPT early on. The Wh- question exercises and the exercise influenced by VNeST (Verb Networks B) in the Sentence Phase of treatment were on the other hand, more challenging for participants and likely require refinement. Such a refinement process is recommended by the MRC guidance on developing complex interventions (www.mrc.ac.uk/complexinterventionsguidance) with a series of studies potentially required to progressively refine the design of treatment prior to full-scale evaluation. Collaboration with PwA would also include usability testing to maximise the accessibility of the SPT (e.g., Roper et al., 2018) and potential collaboration with technologists to develop the platform on which the SPT is delivered (ibid). In this regard, it should be emphasized here that the SPT is a prototype treatment, and this includes technologically. Thus, it is not yet programmed to monitor participant performance, nor does it use, for example, avatar clinicians to deliver treatment (which is being pioneered in self-delivery of, for example, Aphasia Script Training). It is hoped that the SPT will be developed in future versions to include such technological advances. The SPT's rudimentary technology may, however, have had the advantage of reducing technological difficulties by minimizing demands on the processing capacity of whatever device was owned by participants. Harrison et al. (2020) report that the main cause of technological difficulties experienced by participants in the Big CACTUS study was the newly implemented voice recognition facility of StepByStep which was not supported by some of the (older) devices used to deliver treatment. Thus future, more technologically advanced, implementations of the SPT need to attend to minimizing such barriers to delivering the computer-based treatment.

In future studies, there are several other limitations of this feasibility study which need to be addressed. These include the fidelity of monitoring the amount of treatment self-delivered. Because participants recorded this manually, this was not always accurate. Ideally, in a future study the platform used to deliver the SPT would record the amount of SPT self-delivered remotely as recommended by Kurland et al. (2018), and this would also be an aim of collaboration with technologists. Having noted that participants did not always record the amount of self-delivered treatment accurately, this varied very significantly between participants in both amount and intensity. Although the time commitment was explained very clearly to participants at the recruitment stage (using aphasia-friendly PowerPoint slides and principles of supported conversation), this would be reiterated to participants and schedules discussed in more detail with participants with aim of improving compliance. The SPT was also not delivered to the intended amount or intensity on a face-to-face basis. This was mostly due to participants cancelling session due to personal reasons (such as ill health or holidays). However, on one occasion the research student was unable to deliver face-to-face treatment sessions because of issues with public transport, and on several occasions the time taken to travel to participants was extended for the same reason. A possible solution to this is to carry out face-to-face treatment sessions remotely and this would have the additional economic advantage of reducing the amount of clinician time required to supervise the SPT. (Remote supervision has been found to be effective in maintaining compliance with self-delivered computer-based treatment (Braley et al., 2021; Kurland et al., ibid) and to be acceptable (Amaya et al., 2018)). Remote supervision is also of increased relevance because of the COVID 19 pandemic and future studies of the SPT could explore this mode of supporting participants.

A more structured approach to training participants to use the SPT would also be taken in a future research study to improve compliance with self-delivered treatment, namely specific training session/s on uploading the SPT exercises and using them. The importance of SLT training and support for participants for compliance with treatment is noted in the Big CACTUS study (Burke et al., 2021; Harrison et al., 2020). Attending to the training of participants would also be part of the refinement process recommended for complex interventions by the MRC. Further limitations of the study included three of the participants continuing to use the SPT during the maintenance phase of the study which consequently was not a maintenance phase for these participants. (However, this continued use was informative because: i) it indicated that these participants found the SPT acceptable and beneficial, and ii) that more training in using the SPT may benefit less technologically competent PwA since P5 taught herself how to use the SPT during this period). Future studies would withdraw the treatment from participants by removing it from devices, thus enforcing the maintenance phase.

Finally, a future study would amend some of the background assessments and outcome measures used in the feasibility study. For example, the Verb Comprehension Test from the VAST (Bastiaanse et al., 2002) was used as a background assessment to detect any deficit in verb semantics. However, all participants performed almost at ceiling and the test would not therefore be used in a future study to reduce the assessment burden. In terms of outcome measures, the AMI used to measure narrative production resulted in fatigue effects for participants and could not be used with P6 as he found it too distressing because of the severity of his expressive deficit. Thus, (as discussed earlier), The Scenario Test (van der Meulen et al., 2010) would be used as an alternative outcome measure because it allows the use of non-verbal methods of communication as well as verbal and is likely to take less time to administer due to its structured nature.

7.9 Overall conclusions.

This PhD feasibility study set out to explore the feasibility, acceptability, compliance and fidelity of a novel verb and sentence production treatment, with some preliminary efficacy information arising from pre-post studies of individual cases. The treatment program was low dose and clinician delivered, supplemented by self-managed computer-based treatment.

The study began with two systematic scoping reviews to inform the development of the SPT. Both of these have been published (Hickin et al., 2020 & 2022). The reviews contribute to the research base by reporting on the fidelity and the level of evidence for verb and sentence production treatments respectively, together with the impact of treatments on key levels of language.

The feasibility component of the study indicated that it was feasible to recruit suitable participants for the SPT, with initial indications being that individuals with nonfluent aphasia and moderately impaired lexical retrieval and syntactic processing may benefit the most from treatment. The SPT also appears to be technologically feasible, with five of the six participants achieving independent use of the SPT, and four out of six complying with the requested amount of self-delivered treatment. Future studies of the SPT will seek to identify participants with the necessary skill set to self-deliver the SPT/identify those who need additional training by a more nuanced assessment of competence with technology at the recruitment stage. The SPT targeted PR verbs and it was found to be feasible to select a set of 40 PR verbs for participants. The use of PR stimuli may also have been important in maintaining motivation to use the self-delivered computer-based treatment and this finding is in line with previous studies.

Preliminary efficacy testing of the SPT indicated that it was effective in improving PR verb retrieval in isolation, but that this was largely restricted to treated items. Sentence production using treated PR verbs also improved as a result of the SPT, with some limited generalization to untreated PR verbs. The impact of the SPT on the production of personally irrelevant verbs in isolation and in sentences was very minimal, hinting that the salience of the PR verbs targeted in treatment may be an important (active) ingredient of treatment. The impact of the SPT on discourse production was also minimal and this may have been because the SPT did not target discourse specifically and/or the outcome measures were insensitive to any changes. Despite the lack of changes in discourse production, four participants perceived their functional communication to have improved and this was corroborated by two significant others. In summary, whilst the findings of this study represent only preliminary, low level evidence, they are important because they indicate that the complex intervention of sentence treatment may successfully be self-delivered by PwA using a computer. Given the increasing reliance of PwA on self-delivered treatment, particularly since the COVID 19 pandemic, and the need to develop computer-based treatments which go beyond single word processing and which have a functional impact, the findings of the study represent a small but potentially significant step forward.

References.

Adrian J.A., Gonzalez. M., Buiza, J.J. & Sage, K. (2011). Extending the use of spanish computer-assisted anomia rehabilitation program (CARP-2) in people with aphasia. *Journal of Communication Disorders,* 44, 666–677.

Allen, L., Mehta, S., McClure, J.A. & Teasell, R. (2012) Therapeutic Interventions for Aphasia Initiated More than Six Months Post Stroke: A Review of the Evidence. *Topics in Stroke Rehabilitation, 19*:6, 523-535. https://doi.org/10.1310/tsr1906-523.

Amaya, A., Woolf, C., Devane, N., Galliers, J., Talbot, R., Wilson, S. & Marshall, J. (2018) Receiving aphasia intervention in a virtual environment: the participants' perspective, *Aphasiology*, 32, 538-558. DOI: https://doi.org/10.1080/02687038.2018.1431831

Aphasia Software Finder. (n.d.) Available at: https://www.aphasiasoftwarefinder.org/talking?tid=28. [Accessed: 1 September, 2021].

Attard, M.C., Loupis, Y., Togher L. & Rose, M.L. (2018) The efficacy of an inter-disciplinary community aphasia group for living well with aphasia. *Aphasiology, 32*, 105-138, DOI: 10.1080/02687038.2017.1381877

Baker, E. (2012a). Optimal intervention intensity, *International Journal of Speech-Language Pathology, 14*:5, 401-409, DOI: 10.3109/17549507.2012.700323

Baker, E. (2012b). Optimal intervention intensity in speech-language pathology: Discoveries, challenges, and unchartered territories. *International Journal of Speech-Language Pathology, 14*(5), 478-485. 10.3109/17549507.2012.717967

Ball, A.L., de Riesthal, M. & Steele, R.D. (2018). Exploring treatment fidelity in persons with aphasia autonomously practicing with computerized therapy materials. *American Journal of Speech-Language Pathology, 27*, 454–463

Basic English Key Words List (undated) Available at: https://www.thoughtco.com/basic-english-key-words-1212307 [Accessed: 14 February 2019].

Bastiaanse, R., Hurkmans, J. & Links, P.(2006) The training of verb production in Broca's aphasia: A multiple-baseline across-behaviours study, *Aphasiology, 20*:02-04, 298-311, DOI: 10.1080/02687030500474922

Bastiaanse, R., Edwards, S. & Rispens, J. (2002). *The Verb and Sentence Test*. Bury St Edmunds: Thames Valley Test Company Limited.

Bazzini, A., Zonca, G., Craca, A., Cafforio, E., Cellamare, F., Guarnaschelli, G., Felicetti, G. & Luzzatti, C. (2012). Rehabilitation of argument structure deficits in aphasia. *Aphasiology, 26*, 1440-1460, DOI:10.1080/02687038.2012.690023

Behn, N., Moss, B., McVicker, S., Roper, A., Northcott, S., Marshall, J., . . . Hilari, K. (2021). Supporting wellbeing through PeeR-befriending (SUPERB) feasibility trial: Fidelity of peer-befriending for people with aphasia. BMJ Open, 11(8) doi:10.1136/bmjopen-2020-047994.

Biber, D. and Conrad, S. (undated). Corpus Linguistics and Grammar Teaching. Available at: https://www.scribd.com/document/297216707/Pl-Biber-Conrad-Monograph5-Lo [Accessed: 9 November 2017].

Biber, D., S. Johansson, G. Leech, S. Conrad, & E. Finegan. (1999). Longman Grammar of Spoken and Written English. London: Longman.

Biran, M., & Fisher, S. (2015). Structured treatment can improve predicate argument structure impairment. *Aphasiology, 29*(1), 29-56. doi:10.1080/02687038.2014.961122

Blankestijn-Wilmsen, J., Damen, I., Voorbraak-Timmerman, V., Hurkmans, J., Koning, B. D., Pross, A., & Jonkers, R. (2017). The effect of static versus dynamic depictions of actions in verb and sentence production in aphasia. Aphasiology, 31(10), 1166-1182. doi:10.1080/02687038.2016.1258537

Blomert, L., Kean, M. L., Koster, C., & Schokker, J. (1994). Amsterdam—Nijmegen everyday language test: Construction, reliability and validity. Aphasiology, 8(4), 381-407. 10.1080/02687039408248666

Bonifazi, S., Tomaiuolo, F., Altoè, G., Ceravolo, M. G., Provinciali, L. & Marangolo, P. (2013) Action observation as a useful approach for enhancing recovery of verb production: new evidence from aphasia. European Journal of Physical and Rehabilitation Medicine, 49(4), 473-481.

Boo, M., & Rose, M. L. (2011). The efficacy of repetition, semantic, and gesture treatments for verb retrieval and use in broca's aphasia. *Aphasiology, 25*(2), 154-175. doi:10.1080/02687031003743789

Bowen, D.J., Kreuter, M., Spring, B., Cofta-Woerpel, L., Linnan, L et al. (2009) How We Design Feasibility Studies. *Am J Prev Med, 36*(5): 452–457. doi:10.1016/j.amepre.2009.02.002.

Boyle, M. (2017). Semantic treatments for word and sentence production deficits in aphasia. *Seminars in Speech & Language, 38*(1), 52-61. 10.1055/s-0036-1597256

Brady, M.C., Kelly, H., Godwin, J., Enderby, P. and Campbell, P. (2016) 'Speech and language therapy for aphasia following stroke', Cochrane Database of Systematic Reviews, 6(Art. No.: CD000425). DOI: 10.1002/14651858.CD000425.pub4.

Braley, M., Sims Pierce, J., Saxena, S., De Oliveira, E., Taraboanta, L., Anantha, V., Lakhan, S.E. and Kiran, S. (2021). A Virtual, Randomized, Control Trial of a Digital Therapeutic for Speech, Language, and Cognitive Intervention in Post-stroke Persons with Aphasia. *Frontiers in Neurology, 12*, Article 626780.

Brandenburg, C., Worrall, L., Rodriguez, A. D. & Copland, D. (2013). Mobile computing technology and aphasia: An integrated review of accessibility and potential uses. *Aphasiology, 27*, 4, 444–461.

Brennan, A., Worrall, L., & McKenna, K. (2005) The relationship between specific features of aphasia-friendly written material and comprehension of written material for people with aphasia: An exploratory study. *Aphasiology, 19*, 693-711, DOI: 10.1080/02687030444000958

Brogan, E., Ciccone, N., & Godecke, E. (2019). Treatment fidelity in aphasia randomised controlled trials. *Aphasiology, 33*(7), 759-779. doi:10.1080/02687038.2019.1576442

Bryant, L., Ferguson, A., & Spencer, E. (2016). Linguistic analysis of discourse in aphasia: A review of the literature. Clinical Linguistics & Phonetics, 30, 489-518.

Burke, J., Palmer, R., & Harrison, M. (2021). What are the factors that may influence the implementation of self-managed computer therapy for people with long term aphasia following stroke? A qualitative study of speech and language therapists' experiences in the big CACTUS trial. *Disability and Rehabilitation,* doi:10.1080/09638288.2020.1871519

Byng, S., Nickels, L. and Black, M. (1994). Replicating therapy for mapping deficits in agrammatism: Remapping the deficit?, *Aphasiology, 8*, 315-341, DOI: 10.1080/02687039408248663

Byng, S. (1988) Sentence processing deficits: Theory and therapy, Cognitive Neuropsychology, 5, 629-676, DOI:http://dx.doi.org/10.1080/02643298808253277

Cahana-Amitay, D., Albert , M.L., Pyun, S-B., Westwood, A., Jenkins, T. Wolford, S. et al. (2011) Language as a stressor in aphasia, Aphasiology, 25, 593-614 DOI:http://dx.doi.org/10.1080/02687038.2010.541469

Cairns, D., Marshall, J., Cairns, P. & Dipper, L. (2007). Event processing through naming: Investigating event focus in two people with aphasia. Language and Cognitive Processes, 22:2, 201-233, DOI: 10.1080/01690960500489644

Cambridge English Dictionaries (n.d) Available at: https://dictionary.cambridge.org/dictionary/english/personal [Accessed: 19 September

2020].

Cann, K. (2021) A rich picture. Bulletin of The Royal College of Speech and Language Therapists. Summer, p.40-41.

Carragher, M., Sage, K. and Conroy, P. (2015). Outcomes of treatment targeting syntax production in people with Broca's-type aphasia: evidence from psycholinguistic assessment tasks and everyday conversation. *International Journal of Language & Communication Disorders,* 50, 322-336.

Carragher, M., Sage, K. & Conroy, P. (2014): Preliminary analysis from a novel treatment targeting the exchange of new information within storytelling for people with nonfluent aphasia and their partners, Aphasiology, DOI: 10.1080/02687038.2014.988110

Carragher, M., Sage K, and Conroy, P. (2013). The effects of verb retrieval therapy for people with non-fluent aphasia: Evidence from assessment tasks and conversation. *Neuropsychological Rehabilitation*, 23, 846-887. doi:10.1080/09602011.2013.832335

CELEX (undated) Available at: https://catalog.ldc.upenn.edu/LDC96L14 [Accessed: 14 February 2019].

Cherney, L. R., Lee, J. B., Kim, K. A. & Van Vuuren, S. (2021). Web-based oral reading for language in aphasia (web ORLA ®): A pilot randomized control trial. *Clinical Rehabilitation, 35*(7), 976-987. doi:10.1177/0269215520988475

Cherney, L. R., Kaye, R. C., Lee, J. B., & van Vuuren, S. (2015). Impact of personal relevance on acquisition and generalization of script training for aphasia: A preliminary analysis. American Journal of Speech-Language Pathology, 24(4), S913-S922. doi:10.1044/2015_AJSLP-14-0162

Cherney, L. R. (2012). Aphasia treatment: Intensity, dose parameters, and script training. International Journal of Speech-Language Pathology, 14, 424-431. doi:10.3109/17549507.2012.686629

Cherney, L. R., Halper, A.S., Holland, A.L, & Cole, R. (2008). Computerized Script Training for Aphasia: Preliminary Results. *American Journal of Speech-Language Pathology, 17*, 19 –34.

Clarke, D.J., Burton, L.-J., Tyson, S.F., et al. (2018) Why do stroke survivors not receive recommended amounts of active therapy? Findings from the ReAcT study, a mixed-methods case-study evaluation in eight stroke units. *Clinical Rehabilitation, 32* (8). pp. 1119-1132. https://doi.org/10.1177/0269215518765329

Collins, A.M. & Loftus, E.M. (1975). A Spreading-Activation Theory of Semantic Processing *Psychological Review*, 82(6), 407-428.

Conlon, E. L., Braun, E. J., Babbitt, E. M., & Cherney, L. R. (2020). Treatment fidelity procedures for an aphasia intervention within a randomized controlled trial: Design, feasibility, and results. *American Journal of Speech-Language Pathology, 29*, 412-424. doi:10.1044/2019_AJSLP-CAC48-18-0227

Conroy, P. & Scowcroft, J., (2012). Decreasing cues for a dynamic list of noun and verb naming targets: A case-series aphasia therapy study. Neuropsychological Rehabilitation, 22(2), 295-318.

Conroy, P., Sage, K., & Lambon-Ralph, M. (2009a). A comparison of word versus sentence cues as therapy for verb naming in aphasia. *Aphasiology,* 23 (4) 462 – 482.

Conroy, P., Sage, K., & Lambon-Ralph, M. (2009b). The effects of decreasing and increasing cue therapy on improving naming speed and accuracy for verbs and nouns in aphasia, *Aphasiology,* 23 (6), 707–730.

Conroy, P., Sage, K. & Ralph, M.A.L. (2009c). Errorless and errorful therapy for verb and noun naming in aphasia. *Aphasiology,* 23 (11), 1311-1337.

Conroy, P., Sage, K. & Ralph, M.A.L. (2006). Towards theory-driven therapies for aphasic verb impairments: A review of current theory and practice. *Aphasiology, 20*:12, 1159-1185, DOI: 10.1080/02687030600792009

Constant Therapy (undated) Available at: https://constanttherapyhealth.com/constant-therapy/ [Accessed 25th March 2022].

Cruice, M., Pritchard, M. & Dipper, L. (2014) Verb use in aphasic and non-aphasic personal discourse: What is normal? *Journal of Neurolinguistics 28,* 31–47. doi:10.1016/j.jneuroling.2013.12.002.

Davidson, B., Worrall, L. & Hickson, L. (2003) Identifying the communication activities of older people with aphasia: evidence from naturalistic observation. *Aphasiology, 17*, 243-264.

Davis, G.A. and Tan, L.L. (1987) Stimulation of sentence production in a case with agrammatism. Journal of Communication Disorders, 20, 447-457.

De Aguiar, V., Bastiaanse, R. & Miceli, G. (2016). Improving production of treated and untreated verbs in aphasia: a meta-analysis. *Frontiers in Human Neuroscience, 10,* 468, doi: 10.3389/fnhum.2016.00468

De Cock, E., Batens, K., Feiken, J., Hemelsoet, D., Oostra, K., & De Herdt, V. (2021). The feasibility, usability and acceptability of a tablet-based aphasia therapy in the acute phase following stroke. *Journal of Communication Disorders, 89*, N.PAG-N.PAG. doi:10.1016/j.jcomdis.2020.106070

Des Roches, C.A., Mitko, A. & Kiran, S. (2017). Relationship between Self-Administered Cues and Rehabilitation Outcomes in Individuals with Aphasia: Understanding Individual Responsiveness to a Technology-Based Rehabilitation Program. *Frontiers in Human*

Neuroscience, 11, 7, doi: 10.3389/fnhum.2017.00007

Des Roches, C.A., Balachandran, I., Ascenso, E.M.,Tripodis,Y., & Kiran, S. (2015). Effectiveness of an impairment-based individualized rehabilitation program using an iPad-based software platform. Frontiers of Human Neuroscience, 8, 1015, doi: 10.3389/fnhum.2014.01015

Dietz, & Boyle, M. (2017). Discourse measurement in aphasia research: have we reached the tipping point? *Aphasiology*, https://doi.org/10.1080/02687038.2017.1398803

Dignam, J. K., Rodriguez, A. D., & Copland, D. A. (2016). Evidence for intensive aphasia therapy: Consideration of theories from neuroscience and cognitive psychology. PM & R: Journal of Injury, Function & Rehabilitation, 8(3), 254-267. https://doi.org/10.1016/j.pmrj.2015.06.010

Dijkers, M. (2015). What is a scoping review. KT Update, 4. An e-newsletter from the Center on Knowledge Translation for Disability and Rehabilitation Research. [http://ktdrr.org/products/update/v4n1]

Dipper, L. T., Franklin, S., de Aguiar, V., Baumgaertner, A., Brady, M., Best, W., . . . Visch-Brink, E. G. (2021). An umbrella review of aphasia intervention descriPtion in research: The AsPIRE project. Aphasiology, doi:10.1080/02687038.2020.1852001

Dipper, L., Marshall, J., Boyle, M., Botting, N., Hersh, D., Pritchard, M. … Cruice, M. (2020). Treatment for improving discourse in aphasia: a systematic review and synthesis of the evidence base. Aphasiology. doi:10.1080/02687038.2020.1765305.

Douglas, J. M., O'Flaherty, C. A., & Snow, P. C. (2000). Measuring perception of communicative ability: The development and evaluation of the La Trobe communication questionnaire. *Aphasiology, 14*(3), 251–268.

Druks, J. & Masterson, J. (2000) An Object Action Naming Battery. Hove: Psychology Press.

Ebbels, S.H. van der Lely, H. K.. & Dockrell, J.E. (2007) Intervention for Verb Argument Structure in Children With Persistent SLI: A Randomized Control Trial. *Journal of Speech, Language, and*

Hearing Research, *50*, 1330 –1349.

Edmonds, L. A. (2016). A review of verb network strengthening treatment. *Topics in Language Disorders, 36*(2), 123-135. doi:10.1097/TLD.0000000000000088.

Edmonds, L. A., Obermeyer, J., & Kernan, B. (2015). Investigation of pretreatment sentence production impairments in individuals with aphasia: Towards understanding the linguistic variables that impact generalisation in verb network strengthening treatment. *Aphasiology, 29*(11), 1312-1344. doi:10.1080/02687038.2014.975180

Edmonds, L. A. (2014). Tutorial for verb network strengthening treatment (VNeST): Detailed description of the treatment protocol with corresponding theoretical rationale. *Perspectives on Neurophysiology & Neurogenic Speech & Language Disorders, 24*(3), 78-88.

Edmonds, L. A., Mammino, K., & Ojeda, J. (2014). Effect of verb network strengthening treatment (VNeST) in persons with aphasia: Extension and replication of previous findings. *American Journal of Speech-Language Pathology*, 23(2), S312-29. doi:10.1044/2014_AJSLP-13-0098

Edmonds, L. A., & Babb, M. (2011). Effect of verb network strengthening treatment in moderate-to-severe aphasia. *American Journal of Speech-Language Pathology*, 20(2), 131-145. doi:10.1044/1058-0360(2011/10-0036)

Edmonds, L.A., Nadeau, S.E. & Kiran, S. (2009) Effect of Verb Network Strengthening Treatment (VNeST) on lexical retrieval of content words in sentences in persons with aphasia , *Aphasiology, 23*:3, 402-424, DOI: 10.1080/02687030802291339

Edwards, S. & Tucker, K. (2006). Verb retrieval in fluent aphasia: A clinical study. *Aphasiology*, 20, 644-675, DOI: 10.1080/02687030600631827

Efstratiadou, E.A., Papathanasiou, I., Holland, R., Archonti, A. & Hilari, K. (2018) A Systematic Review of Semantic Feature Analysis Therapy Studies for Aphasia. Journal of Speech, Language, and Hearing Research, 61, 1261–1278.

Faroqi-Shah. Y & Graham, L.E. (2011). Treatment of semantic verb classes in aphasia: acquisition and generalization effects. *Clinical Linguistics & Phonetics, 25*(5): 399–418.

Ferretti, T. R., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language, 44*, 516–547.

Fink, R.B., Schwartz, M.F. & Myers, J.L. (1998). Investigations of the Sentence-Query Approach to Mapping Therapy. *Brain and Language. 65,* 203-207.

Floel, A. (2019). Computerised speech and language therapy in post-stroke aphasia. The Lancet Neurology, 18, 806-807.

Frattali, C.M., Thompson, C.K., Holland, A.L., Wohl, C.B. and Ferketic, M.M. (1995) *The American Speech-Language Hearing Association Functional Assessment of Communication Skills for Adults (ASHA FACS).* ASHA

Furnas, D. W., & Edmonds, L. A. (2014). The effect of computerised Verb Network Strengthening Treatment on lexical retrieval in aphasia. *Aphasiology,* 28(4), 401-420. doi: 10.1080/02687038.2013.869304

Garrett, M. (1988). Processing in language production. In F. Newmeyer (Ed.), Linguistics: *The Cambridge Survey, Vol 3.* Cambridge: Cambridge University Press.

Gleason, J. B., Goodglass H., Green, E., Ackerman, N., Hyde, M. (1975). The retrieval of syntax in Broca's aphasia. *Brain and Language*, 24,451-471.

Godlove, J., Anantha, V., Advani, M., Des Roches, C. and Kiran, S. (2019) Comparison of Therapy Practice at Home and in the Clinic: A Retrospective Analysis of the Constant Therapy Platform Data Set. *Frontiers in Neurology, 10*, Article 140. doi: 10.3389/fneur.2019.00140

Goral, M. and Kempler, D. (2009) Training verb production in communicative context: Evidence from a person with chronic non-fluent aphasia. *Aphasiology,* 23 (12), 1383–1397.

Gravier, M.L., Dickey, M.W., Hula, W.D, Evans, W.S., Owens, R.L., Winans-Mitrik, R.L. & Doyle, P.J. (2018) What Matters in Semantic Feature Analysis: Practice-Related Predictors of Treatment Response in Aphasia. *American Journal of Speech-Language Pathology*, 27, 438–453. https://doi.org/10.1044/2017_AJSLP-16-0196

Haley, K. L., Womack, J., Helm-Estabrooks, N., Lovette, B., & Goff, R. (2013). Supporting autonomy for people with aphasia: Use of the Life Interests and Values (LIV) cards. Topics in Stroke Rehabilitation, 20(1), 22–35. https://doi.org/10.1044/2018_AJSLP-18-0163

Harrison, M., Palmer, R. & Cooper, C. (2020). Factors associated with adherence to self-managed aphasia therapy practice on a computer—a mixed methods study alongside a randomized controlled trial. *Frontiers in Neurology*. doi: 10.3389/fneur.2020.582328

Helm-Estabrooks, N. & Ramsberger, G. (1986) Treatment of agrammaticism in long term Broca's aphasia. *British Journal of Disorders of Communication, 21*, 39-45.

Helm-Estabrooks, N. (1981). Helm Elicited Language Program for Syntax Stimulation. Austin, Texas: Exceptional Resources Inc.

Hersh, D., Worrall, L., Howe, T., Sherratt, S., Davidson, B., (2012), SMARTER goal setting in aphasia rehabilitation. *Aphasiology*, 26(2), 220-233, Abingdon, Oxon, DOI: 10.1080/02687038.2011.640392.

Hickin, J., Cruice, M. & Dipper, L. (2022). A systematically conducted scoping review of the evidence and fidelity of treatments for verb and sentence deficits in aphasia: sentence treatments. *American Journal of Speech Language Pathology, 31,* 1: 431-462. https://doi.org/10.1044/2021_AJSLP-21-00120.

Hickin, J., Cruice, M. & Dipper, L. (2020). A systematically conducted scoping review of the evidence and fidelity of treatments for verb deficits in aphasia: verb in isolation treatments. *American Journal of Speech Language Pathology*, 29, 530-559. https://doi.org/10.1044/2019_AJSLP-CAC48-18-0234

Hickin, J., Mehta, B. & Dipper, L (2015) To the sentence and beyond: a single case therapy report for mild aphasia. *Aphasiology,* 29 (9), 1038-1061. doi:10.1080/02687038.2015.1010474

Hickin, J., Caute, A. and Woolf, C. (2012) *Information Technology in Aphasia Therapy 1: New Developments*. Symposium Presentation at the British Aphasiology Society Biennial International Therapy Symposium. September, City University London.

Hickin, J., Woolf, C. and Caute, A., and Robey, J., (2011, September) A pilot outcome measure for the use of email by people with aphasia. Poster presentation at the British Aphasiology Society Biennial International Conference, Reading.

Hinckley, J. J. & Douglas, N. F. (2013). Treatment fidelity: Its importance and reported frequency in aphasia treatment studies. *American Journal of Speech-Language Pathology*, 22, 279–284.

Hoffman, T.C., Paul P Glasziou, P.P., Boutron, I., Milne, R., Perera, R., Moher, D., Altman, D.G., . . . Michie, S. (2014). Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *British Medical Journal*, 348:g1687 doi: 10.1136/bmj.g1687

Hoover, E.L., Caplan, D., Waters, G. & Budson, A. (2015) Effects of impairment-based individual and socially oriented group therapies on verb production in aphasia, *Aphasiology,* 29:7, 781-798, DOI: 10.1080/02687038.2014.989953

Hopper, T., & Holland, A. L. (2005). Aphasia and learning in adults: Key concepts and clinical considerations. *Topics in Geriatric Rehabilitation*, 21(4), 315-322. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=rzh&AN=106428077&site=ehost-live

Howard, D., Best, W. & Nickels, L. (2015) Optimising the design of intervention studies: critiques and ways forward, *Aphasiology,* 29:5, 526-562, DOI: 10.1080/02687038.2014.985884

Howard D. (2000). Cognitive neuropsychology and aphasia therapy: the case of word retrieval. In: Papathanasiou, I, ed. *Acquired Neurogenic Communication Disorders: A Clinical Perspective*. London: Whurr.

Jones, E. V. (1986). Building the foundations for sentence production in a non-fluent aphasic. *International Journal of Language & Communication Disorders, 21*, 63-82. doi:10.3109/13682828609018544

Kaderavek, J.N. and Justice L.M. (2010). Fidelity: An Essential Component of Evidence-Based Practice in Speech-Language Pathology. *American Journal of Speech-Language Pathology,19*, 369–379.

Kearns, Á., Kelly, H., & Pitt, I. (2019). Self-reported feedback in ICT-delivered aphasia rehabilitation: A literature review. Disability and Rehabilitation, 1-15. doi:10.1080/09638288.2019.1655803

Kelly, H., Kennedy, F., Britton, H., McGuire, G. & Law, J. (2016) Narrowing the "digital divide"—facilitating access to computer technology to enhance the lives of those with aphasia: a feasibility study. *Aphasiology, 30*, 2-3, 133-163. https://doi.org/10.1080/02687038.2015.1077926

Kempler, D., & Goral, M. (2011). A comparison of drill- and communication-based treatment for aphasia. *Aphasiology*, 25(11), 1327-1346. 10.1080/02687038.2011.599364

Kim, M., & Thompson, C. K. (2000). Patterns of comprehension and production of nouns and verbs in agrammatism: implications for lexical organization. *Brain and Language, 74,* 1-25.

Kleim, J.A. & Jones, T.A. (2008). Principles of Experience-Dependent Neural Plasticity: Implications for Rehabilitation After Brain Damage. *Journal of Speech, Language, and Hearing Research,* 51, S225–S239.

Knoph, M. I. N., Lind, M., & Simonsen, H. G. (2015). Semantic feature analysis targeting verbs in a quadrilingual speaker with aphasia. *Aphasiology, 29*(12), 1473-1496. doi:10.1080/02687038.2015.1049583

Kolk, H. (1995). A time-based approach to agrammatic production. *Brain and Language, 50*, 282–303.

Kopelman, M.D., Wilson, B.A. & Baddeley, A.D. (1989). The autobiographical memory interview: A new assessment of autobiographical and personal semantic memory in amnesic patients, *Journal of Clinical and Experimental Neuropsychology, 11*:5, 724-744, DOI: 10.1080/01688638908400928

Kristensson, J. Behrns, I. & Saldert, C. (2015) Effects on communication from intensive treatment with semantic feature analysis in aphasia. *Aphasiology*, 29:4, 466-487, DOI: 10.1080/02687038.2014.973359

Kurland, J., Liu, A. and Stokes, P. (2018) Effects of a tablet-based home practice program with telepractice on treatment outcomes in Chronic aphasia. Journal of Speech, Language, and Hearing Research, 61, 1140–1156. https://doi.org/10.1044/2018_JSLHR-L-17-0277

Kurland, J. (2014). iRehab in aphasia treatment. *Seminars in Speech & Language*, 35(1), 3-4.

Kurland, J., Wilkins, A. R., & Stokes, P. (2014). iPractice: Piloting the effectiveness of a tablet-based home practice program in aphasia treatment. *Seminars in Speech & Language, 35*(1), 51-63. doi:10.1055/s-0033-1362991.

Lavoie, M., Macoir, J., & Bier, N. (2017). Effectiveness of technologies in the treatment of post-stroke anomia: A systematic review. Journal of Communication Disorders, 65, 43–53. http://dx.doi.org/10.1016/j.jcomdis.2017.01.001

Lavoie, M.. Routhier, S., Légaré, A. & Macoir, J. (2016) Treatment of verb anomia in aphasia: efficacy of self-administered therapy using a smart tablet. *Neurocase, 22*, 109–118, http://dx.doi.org/10.1080/13554794.2015.1051055

Le Dorze, G., Jacob, A. & Coderre, L. (1991). Aphasia rehabilitation with a case of agrammatism: A partial replication. *Aphasiology, 5*, 63-85. DOI:10.1080/02687039108248520

Lesser, R. & Algar, L. (1995) Towards combining the cognitive neuropsychological and the pragmatic in aphasia therapy. *Neuropsychological Rehabilitation,* 5:1-2, 67-92, DOI: 10.1080/09602019508520176

Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge MA: MIT Press.

Lincoln. N. B., Mcguirk, E., Mulley, G. P., Lendrem, W . , Jones. A. C . Mitchell, J. R. A. (1984). Effectiveness of speech therapy for aphasic stroke patients; a randomised controlled trial. *Lancet,* 1, 1197-1200.

Linebarger, M. C., Romania, J. F., Fink, R. B., Bartlett, M. R., & Schwartz, M. F. (2008). Building on residual speech: A portable processing prosthesis for aphasia. Journal of Rehabilitation Research and Development, 45(9), 1401-1414. doi:10.1682/JRRD.2007.10.0171

Links, P., Hurkmans, J., & Bastiaanse, R. (2010). Training verb and sentence production in agrammatic broca's aphasia. *Aphasiology, 24*(11), 1303-1325. doi:10.1080/02687030903437666

Lomas, J., Pickard, L., Bester, S., Elbard, H., Finlayson, A. & Zoghaib, C. (1989) The Communicative Effectiveness Index: development and psychometric evaluation of a functional communication measure for adult aphasia. *Journal of Speech and Hearing Disorders*, 54, 113-124.

Long, A. F., Hesketh, A., Paszek, G., Booth, M., & Bowen, A. (2008). Development of a reliable self-report outcome measure for pragmatic trials of communication therapy following stroke: The Communication Outcome after Stroke (COAST) scale. *Clinical Rehabilitation*, 22, 1083–1094. doi:10.1177/0269215508090091

Luzzatti, C., Raggi, R., Zonca, G., Pistarini, C., Contardi, A., & Pinna, G. (2001). On the nature of the selective impairment of verb and noun retrieval. *Cortex*, 37, 724–726.

Maddy, K.M., Capilouto, J.G. & McComas, ,K.L. (2014) The effectiveness of semantic feature analysis: An evidence-based systematic review. *Annals of Physical and Rehabilitation Medicine,* 57 (2014) 254–267. Doi: http://dx.doi.org/10.1016/j.rehab.2014.03.002

Marangolo P, Cipollari S, Fiori V, Razzano C. & Caltagirone C. (2012) Walking but Not Barking Improves Verb Recovery: Implications for Action Observation Treatment in Aphasia Rehabilitation. *PLoS ONE 7(*6): e38610. https://doi.org/10.1371/journal.pone.0038610

Marangolo P, Bonifazi S, Tomaiuolo F, Craighero L, Coccia M, et al. (2010) Improving Language without words: first evidence from aphasia. Neuropsychologia 48: 3824–33.

Marshall, J., Booth, T., Devane, N., Galliers, J., Greenwood, H., Hilari, K. … Woolf, C. (2016). Evaluating the benefits of aphasia intervention delivered in virtual reality: Results of a quasi-randomised study. *PLoS ONE, 11(*8). doi:10.1371/journal.pone.0160381.

Marshall, J., Roper, A., Galliers, J., Wilson, S., Cocks, N., Muscroft, S. & Pring, T. (2013) Computer delivery of gesture therapy for people with severe Aphasia. *Aphasiology,* 27, (9), 1128–1146. http://dx.doi.org/10.1080/02687038.2013.786803

Marshall, J. & Cairns, D. (2005) Therapy for sentence processing problems in aphasia: Working on thinking for speaking. *Aphasiology*, 19, 1009-1020. DOI:10.1080/02687030544000218

Marshall, J., Pring, T. & Chiat, S. (1998) Verb Retrieval and Sentence Production in Aphasia. *Brain and Language* 63, 159–183.

Marshall, J., Chiat, S. & Pring, T. (1997). An impairment in processing verbs' thematic roles: A therapy study. *Aphasiology, 11*, 855 — 876. DOI: 10.1080/02687039708250461

Marshall, J. (1995). The mapping hypothesis and aphasia therapy, *Aphasiology*, 9(6), 517-539 doi: http://dx.doi.org/10.1080/02687039508248712

Mason, C., Nickels, L., McDonald, B., Moses, M., Makin, K., & Taylor, C. (2011). Treatment of word retrieval impairments in aphasia: Evaluation of a self-administered home programme using personally chosen words. Aphasiology, 25(2), 245-268. doi:10.1080/02687038.2010.489258

Maul, K. K., Conner, P. S., Kempler, D., Radvanski, C., & Goral, M. (2014). Using informative verbal exchanges to promote verb retrieval in nonfluent aphasia. *American Journal of Speech-Language Pathology, 23*(3), 407-420. doi:10.1044/2014_AJSLP-13-0004

McCall, D., Virata, T., Linebarger, M. C., & Berndt, R. S. (2009). Integrating technology and targeted treatment to improve narrative production in aphasia: A case study. Aphasiology, 23(4), 438-461. doi:10.1080/02687030701818000

McCann C. & Doleman, J. (2011). Verb retrieval in nonfluent aphasia: A replication of Edwards & Tucker, 2006. *Journal of Neurolinguistics* 24, 237–248.

McKelvey, M. L., Hux, K., Dietz, A., & Beukelman, D. R. (2010). Impact of personal relevance and contextualization on word-picture matching by people with aphasia. American Journal of Speech-Language Pathology, 19(1), 22-33. doi:10.1044/1058-0360(2009/08-0021)

McNeil , M.R., Doyle, P.J. , Spencer, K. Jackson Goda, V.,Flores, D. & Small, S.L. (1998) Effects of training multiple form classes on acquisition,generalization and maintenance of word retrieval in a single subject, *Aphasiology, 12*, 7-8, 575-585. DOI: 10.1080/02687039808249559

McRae, K., Hare, M., Elman, J. L., & Ferretti, T. (2005). A basis for generating expectancies for verbs from nouns. *Memory and Cognition, 33*(7), 1174–84.

Medical Research Council (2006) Developing and evaluating complex interventions:new guidance. Available at: https://webarchive.nationalarchives.gov.uk/ukgwa/20220207162925/http://mrc.ukri.org/documents/pdf/complex-interventions-guidance/

Menger, F., Morris, J., & Salis, C. (2016) Aphasia in an Internet age: wider perspectives on digital inclusion, Aphasiology, 30:2-3, 112-132, DOI:10.1080/02687038.2015.1109050

Mitchum, C. C., Greenwald, M. L., & Berndt, R. S. (2000). Cognitive treatments of sentence processing disorders: What have we learned? *Neuropsychological Rehabilitation, 10*(3), 311-336. doi:10.1080/096020100389174

Mitchum, C.C., Greenwald, M.L., & Berndt, R.S. (1997). Production-specific thematic mapping impairment: A treatment study. *Brain and Language, 60*, 121–123.

Mortley, J., Wade, J. & Enderby, P. (2004). Superhighway to promoting a client-therapist partnership? Using the Internet to deliver word-retrieval computer therapy, monitored remotely with minimal speech and language therapy input. *Aphasiology*, 18, (3), 193-211.

MRC Psycholinguistic Database (n.d) Available at: https://websites.psychology.uwa.edu.au/school/MRCDatabase/uwa_mrc.htm [Accessed: 21 December 2021].

Murad, M.H., Sultan, S., Haffar S. and Bazerbachi, F. (2018). Methodological quality and synthesis of case series and case reports *BMJ, Evidence-Based Medicine*, 23:60–63. doi:10.1136/bmjebm-2017-110853

Munsell, M., De Oliveira, E., Saxena, S., Godlove, J. and Kiran, S. (2020). Closing the Digital Divide in Speech, Language, and Cognitive Therapy: Cohort Study of the Factors Associated With

Technology Usage for Rehabilitation. Journal of Medical Internet Research, 22(2): e16286. doi: 10.2196/16286:

Newton, C., Kirby, P., & Bruce, C. (2017). Getting into shape: The effect of shape coding on the spoken language production of two men with chronic aphasia. *Aphasiology,* 1-23. doi:10.1080/02687038.2017.1306639

Nichol, L., Hill, A. J., Wallace, S. J., Pitt, R., Baker, C., & Rodriguez, A. D. (2019). Self-management of aphasia: A scoping review. *Aphasiology*, 1-40. doi:10.1080/02687038.2019.1575065

Nicholas, L. E., & Brookshire, R. H. (1993). A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. *Journal of Speech Language and Hearing Research, 36,* 338–350. doi:10.1044/jshr.3602.338

Nickels, L. (2002). Therapy for naming disorders: Revisiting, revising,and reviewing. Aphasiology, 16, 935-979. DOI:10.1080/02687030244000563

Nickels, L., Byng, S. & Black, M. (1991). Sentence processing deficits: A replication of therapy. *International Journal of Language & Communication Disorders*, 26, (2), 175-199.

Oxford Centre for Evidence-Based Medicine 2011 Levels of Evidence. (2011). Available at: http://www.cebm.net/wp-content/uploads/2014/06/CEBM-Levels-of-Evidence-2.1.pdf [Accessed: 23 November 2017]

Oxford English Living Dictionaries (2019) Available at: https://en.oxforddictionaries.com/definition/functional [Accessed: 12 February 2019].§

Palmer, R., Dimairo, M., Cooper, C., Enderby, P., Brady, M., Bowen, A. et al. (2019) Self-managed, computerised speech and language therapy for patients with chronic aphasia post-stroke compared with usual care or attention control (Big CACTUS): a multicentre, single-blinded, randomised controlled trial. *Lancet Neurology, 18*, 821–33

Palmer R, Witts H, Chater, T. (2018) What speech and language therapy do community dwelling stroke survivors with aphasia receive in the UK? *PLoS ONE 13*(7): e0200096. https://doi.org/10.1371/journal.pone.0200096

Palmer, R., Hughes, H., & Chater, T. (2017). What do people with aphasia want to be able to say? A content analysis of words identified as personally relevant by people with aphasia. *PLoS ONE 12*(3): e0174065. https://doi.org/10.1371/journal.pone.0174065

Palmer, R., Enderby, P. & Paterson, G. (2013). Using computers to enable self-management of aphasia therapy exercises for word finding: the patient and carer perspective. *International Journal of Language & Communication Disorders, 48* (5), 508-521.

Palmer, R., Enderby, P., Cooper, C., Latimer, N., Julious, S., Paterson, G., . . . Dixon, S. (2012). Computer therapy compared with usual care for people with long-standing aphasia poststroke: A pilot randomized controlled trial. *Stroke (00392499), 43*(7), 1904-1911. doi:10.1161/STROKEAHA.112.650671

Park, Y., Goral, M., Verkuilen, J. & Kempler, D. (2013). Effects of noun–verb conceptual/phonological relatedness on verb production changes in Broca's aphasia. *Aphasiology, 27*, 7, 811-827.

Perkins, L., Crisp, J., & Walshaw, D. (1999). Exploring conversation analysis as an assessment tool for aphasia: The issue of reliability.13(4-5), 259-281. doi:10.1080/026870399402091

Peters, Micah D.J., Godfrey, Christina M., Khalil, Hanan, McInerney, Patricia, Parker, Deborah and Soares, Cassia Baldini (2015) Guidance for conducting systematic scoping reviews. *International Journal of Evidence-Based Healthcare,* 13 3: 141-146. doi: 10.1097/XEB.0000000000000050

Pitt, R., Theodoros, D., Hill, A.J. & Russell, T. (2019). The development and feasibility of an online aphasia group intervention and networking program – TeleGAIN, *International Journal of Speech-Language Pathology*, 21:1, 23-36. https://doi.org/10.1080/17549507.2017.1369567

Quique, Y.M., Evans, W.S. & Dickey, M.W. (2018) Acquisition and Generalization Responses in Aphasia Naming Treatment: A Meta-Analysis of Semantic Feature Analysis Outcomes. American Journal of Speech-Language Pathology, 28, 1-17. https://doi.org/10.1044/2018_AJSLP-17-0155

Raymer, A. M., Beeson, P., Holland, A., Kendall, D., Maher, L. M., Martin, N., . . . Gonzalez Rothi, L. J. (2008). Translational research in aphasia: From neuroscience to neurorehabilitation. *Journal of Speech, Language, and Hearing Research (JSLHR), 51*(1), S259-S275. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=eoah&AN=22354157&site=ehost-live

Raymer, A.M. , Ciampitti , M., Holliway, B., Singletary, F., Blonder, L.X., Ketterson, T., Anderson , S., Lehnen, J., Heilman, K.M. & Gonzalez Rothi, L.J. (2007) Semantic-phonologic treatment for noun and verb retrieval impairments in aphasia, *Neuropsychological Rehabilitation, 17*:2, 244-270, DOI:10.1080/09602010600814661

Raymer, A.M. Singletary, F., Rodriguez, A.,Ciampitti, M., Heilman, K.M. & Gonzalez Rothi, L.J. (2006) Effects of gesture+verbal treatment for noun and verb retrieval in aphasia. Journal of the International Neuropsychological Society (2006), 12, 867–882. Doi: 10.10170S1355617706061042

Raymer, A., & Kohen, F. (2006). Word-retrieval treatment in aphasia: Effects of sentence context. *Journal of Rehabilitation Research & Development, 43*(3), 367-377.

Raymer, A., Rodriguez, A., Ciampitti, M., Singletary, F. Fuller, R., Heilman, K. & Gonzalez Rothi, L. (2005). Effects of Pantomime Training for Noun and Verb Retrieval in Aphasia *Neurorehabilitation and Neural Repair 19*(4), 378.

Raymer, A.M. & Ellsworth, T.A. (2002). Response to contrasting verb retrieval treatments: A case study. *Aphasiology*, 16, 1031-1045, DOI: 10.1080/026870401430000609

Renvall, K., Nickels, L. & Davidson, B. (2013a) Functionally relevant items in the treatment of aphasia (part I): Challenges for current practice, *Aphasiology, 27*(6), 636-650, DOI: 10.1080/02687038.2013.786804

Renvall, K., Nickels, L. & Davidson, B. (2013b) Functionally relevant items in the treatment of aphasia (part II): Further perspectives and specific tools,. *Aphasiology, 27*(6), 651-677, DOI: 10.1080/02687038.2013.796507

Ritchie, J. & Spencer, L. (1994). Qualitative data analysis for applied policy research. In A. Bryman & R. G. Burgess (Eds), *Analyzing Qualitative Data* (pp. 173-194). London: Routledge.

Robey, R. R., & Schultz, M. C. (1998). A model for conducting clinical outcome research: An adaptation of the standard protocol for use in aphasiology. *Aphasiology,* 12, 787–810.

Rochon, E., Laird, L., Bose, A. & Scofield, J. (2005) Mapping therapy for sentence production impairments in nonfluent aphasia. *Neuropsychological Rehabilitation, 15*:1, 1-36, DOI: 10.1080/09602010343000327

Rochon, E. & Reichman, S. (2003). A Modular Treatment for Sentence Processing Impairments in Aphasia: Sentence Production. *Journal of Speech-Language Pathology and Audiology,* 27, 202 -210.

Rodriguez, A. D., Raymer, A. M., & Gonzalez Rothi, L. J. (2006). Effects of gesture+verbal and semantic-phonologic treatments for verb retrieval in aphasia. Aphasiology, 20(2-4), 286-297. 10.1080/02687030500474898

Roper, A., Davey, I., Wilson, S., Neate, T., Marshall, J. & Grellmann, B. (2018). Usability Testing – An Aphasia Perspective. In: ASSETS '18 Proceedings of the 20th International ACM SIG ACCESS Conference on Computers and Accessibility. (pp. 102-106). New York, USA: ACM. ISBN 978-1-4503-5650-3

Rose, M.L., Raymer, A.M.,] Lanyon, L.E. &] Attard, M.C. (2013) A systematic review of gesture treatments for post-stroke aphasia, *Aphasiology, 27*:9, 1090-1127, DOI: 10.1080/02687038.2013.805726

Rose, T., Worrall, L., Hickson, L., & Hoffmann, T. (2012) Guiding principles for printed education materials: Design preferences of people with aphasia, *International Journal of Speech-Language Pathology*, 14:1, 11-23, DOI:10.3109/17549507.2011.631583

Rose, T., Worrall, L., Hickson, L., & Hoffmann, T. (2011). Aphasia friendly written health information: Content and design characteristics. *International Journal of Speech-Language Pathology*, 13, 335 – 347.

Rose, M., & Sussmilch, G. (2008). The effects of semantic and gesture treatments on verb retrieval and verb use in aphasia. *Aphasiology, 22*(7-8), 691-706. doi:10.1080/02687030701800800.

Rose, T., Worrall, L., & McKenna, K. (2003). The effectiveness of aphasia-friendly principles for printed health education materials for people with aphasia following stroke. *Aphasiology,* 17 , 947 – 963.

Routhier, S., Bier, N. & Macoir, J. (2016) Smart tablet for smart self-administered treatment of verb anomia: two single-case studies in aphasia, Aphasiology, 30, (2-3), 269-289, DOI: 10.1080/02687038.2014.973361

Routhier, S., Bier, N. Macoir, J (2015) The contrast between cueing and/or observation in therapy for verb retrieval in post-stroke aphasia. Journal of Communication Disorders 54 (2015) 43–55. http://dx.doi.org/10.1016/j.jcomdis.2015.01.003

Royal College of Speech and Language Therapists. (undated). Impact of the COVID-19 pandemic on the speech and language therapy profession. Available at: https://www.rcslt.org/wp-content/uploads/media/docs/Covid/RCSLT-Survey---impact-of-COVID---FINAL-(1).pdf [Accessed: 2 October 2021].

Royal College of Speech and Language Therapists. (undated). Self-managed computer therapy guidance. Available at: https://www.rcslt.org/members/delivering-quality-services/self-managed-computer-therapy/self-managed-computer-therapy-guidance/#section-3 [Accessed: 25 March 2022].

Schneider, S. & Thompson, C. (2003). Verb production in agrammatic aphasia: The influence of semantic class and argument structure properties on generalization. *Aphasiology, 17*:3, 213-241, DOI: 10.1080/729255456

Schwartz, M.F. Saffran, E.M., Fink, R.B. Myers, J.L. & Martin, N. (1994) Mapping therapy: A treatment programme for agrammatism. *Aphasiology, 8*, 19-54, DOI: 10.1080/02687039408248639

Scobbie, L., McLean, D., Dixon, D., Duncan, E. & Wyke, S. (2013). Implementing a framework for goal setting in community based stroke rehabilitation: a process evaluation. *BMC Health Services Research, 13,*190 http://www.biomedcentral.com/1472-6963/13/190

Scobbie, L., Dixon, D. & Wyke, S. (2011) Goal setting and action planning in the rehabilitation setting: development of a theoretically informed practice framework. *Clinical Rehabilitation, 25,* 468–482.

Silagi, M.L.,Hirata, F.N., Iracema, L. & de Mendonça, Z. (2014) Sentence production and rehabilitation of agrammatism Sentence production in rehabilitation of agrammatism: A case study. Dementia Neuropsychologia, 8, 297-301.

Speechmark Categories Colorcards (undated). Available at: https://www.routledge.com/Categories-ColorCards-2nd-Edition/Speechmark/p/book/9780815369011. [Accessed: 14 February 2019].

Spell, L. A., Richardson, J. D., Basilakos, A., Stark, B. C., Teklehaimanot, A., Hillis, A. E., & Fridriksson, J. (2020). Developing, implementing, and improving assessment and treatment fidelity in clinical aphasia research. *American Journal of Speech-Language Pathology, 29*(1), 286-298.

doi:10.1044/2019_AJSLP-19-00126

Springer, L., Huber, W., Schlenck, K.J. & Schlenck, C. (2000). Agrammatism: Deficit or compensation? Consequences for aphasia therapy. *Neuropsychological Rehabilitation*, 10 (3), 279–309.

StepByStep Version 5. Available at : https://aphasia-software.com/ [Accessed: 25 March 2022].

SUBTLEXus (undated) Available at: https://www.ugent.be/pp/experimentele-psychologie/en/research/documents/subtlexus [Accessed: 14 February 2019].

Swales, M.A., Hill, A.J. & Finch, E. (2016) Feature rich, but user friendly: Speech pathologists' preferences for computer-based aphasia therapy, *International Journal of Speech-Language Pathology, 18(4)*, 315-328, DOI: 10.3109/17549507.2015.1081283

Swinburn, K. with Byng, S. (2006). *The Communication Disability Profile*. London: Connect.

Swinburn, K., Porter, G., and Howard, D., (2004) *The Comprehensive Aphasia Test.* Hove: Psychology Press

Takizawa, T., Nishida, N., Ikemoto, A., & Kurauchi, N. (2015). Comparison of single-word therapy versus sentence therapy for verb retrieval and sentence production in Broca's aphasia. *Aphasiology, 29*(2), 169-194. doi:10.1080/02687038.2014.961121

Tate, R., L, Perdices, M., Rosenkoetter, U., et al. (2016). The Single-Case Reporting Guideline In BEhavioural Interventions (SCRIBE) 2016 Statement. *Arch Sci Psychol*, 4:1–9

Tate, R.L. Perdices, M., Rosenkoetter, U., Wakim, D., Godbee,K.,M. Togher, L. & McDonald, S., (2013). Revision of a method quality rating scale for single-case experimental designs and n-of-1 trials: The 15-item Risk of Bias in N-of-1 Trials (RoBiNT) Scale. *Neuropsychological Rehabilitation, 23*, 5, 619–638. http://dx.doi.org/10.1080/09602011.2013.824383

Tate, R.L., McDonald, S., Perdices, M., Togher, L., Schultz, R. & Savage, S. (2008) Rating the methodological quality of single-subject designs and n-of-1 trials: Introducing the Single-

Case Experimental Design (SCED) Scale. *Neuropsychological Rehabilitation, 18, (4),* 385–401. DOI:10.1080/09602010802009201

Thiessen, A., & Brown, J. (2021). Personalization of restorative and compensatory treatments for people with aphasia: A review of the evidence. *Topics in Language Disorders, 41*(3), 269-281. doi:10.1097/TLD.0000000000000025

Thompson, C.K., Choy, J.J., Holland, A. & Cole, R. (2010) SENTACTICS: computer-automated treatment of underlying forms. *Aphasiology, 24 (10),* 242-1266.

Thompson, C.K., Lange, K.L, Schneider, S.L. & Shapiro, L. P. (1997). Agrammatic and non-brain-damaged subjects' verb and verb argument structure production. *Aphasiology, 11*, 473-490, https://doi.org/10.1080/02687039708248485

van der Meulen, I., van de Sandt-Koenderman, M.E., Duivenvoorden, H.J. & Ribbers, G.M. (2010). Measuring verbal and non-verbal communication in aphasia: reliability,validity, and sensitivity to change of the Scenario Test. *International Journal of Language and Communication Disorders, 45*, 4, 424–435

Van Lancker, D. & Nicklay, C.K.H. (1992). Comprehension of personally relevant versus novel language in two globally aphasic patients. *Aphasiology, 6(1)*, 37-61. DOI: https://doi.org/10.1080/02687039208248576

Wade, J., Mortley, J. & Enderby, P. (2003). Talk about IT: Views of people with aphasia and their partners on receiving remotely monitored computer-based word finding therapy. *Aphasiology, 17*, 1031-1056. DOI: 10.1080/02687030344000373

Wallace, S.J., Worrall, L., Rose, T., Le Dorze, G., Cruice, M., Isaksen, J., Pak Hin Kong, A., Simmons-Mackie, N., Scarinci N. & Alary Gauvreau, C. (2017) 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, 39*(14), 1364-1379, DOI:

10.1080/09638288.2016.1194899

Wallace, G. L., & Canter, G. J. (1985). Effects of personally relevant language materials on the performance of severely aphasic individuals. *Journal of Speech and Hearing Disorders, 50*, 385–390.

Wambaugh, J. L., Mauszycki, S., & Wright, S. (2014). Semantic feature analysis: Application to confrontation naming of actions in aphasia. *Aphasiology, 28*(1), 1-24. doi:10.1080/02687038.2013.845739

Wambaugh, J.l. & Ferguson, M. (2007).Application of semantic feature analysis to retrieval of action names in Aphasia. *Journal of Rehabilitation Research and Development*, 44, 381-394.

Wambaugh, J., Cameron, R., Kalinyak-Fliszar, M., Nessler, C., & Wright, S. (2004). Retrieval of action names in aphasia: Effects of two cueing treatments. *Aphasiology, 18*(11), 979-1004. doi:10.1080/02687030444000471

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, 455-466.

Webster, J. and Whitworth, A. (2012) Treating verbs in aphasia: exploring the impact of therapy at the single word and sentence levels. International Journal of Language & Communication Disorders, 47: 619–636. doi:10.1111/j.1460-6984.2012.00174.x

Webster, J. & Gordon, B. (2009). Contrasting therapy effects for verb and sentence processing difficulties: a discussion of what worked and why. *Aphasiology,* 23 (10), 1231-1251.

Webster, J., Franklin. S. & Howard, D. (2007) An analysis of thematic and phrasal structure in people with aphasia: What more can we learn from the story of Cinderella? Journal of Neurolinguistics, 20, 363–394.

Webster, J., Morris, J. & Franklin, S. (2005). Effects of therapy targeted at verb retrieval and the realisation of the predicate argument structure: a case study. *Aphasiology,* 19 (8), 748-764.

Whitworth, A., Webster, J., & Howard, D. (2015). Argument structure deficit in aphasia: It's not all about verbs. *Aphasiology, 29*(12), 1426-1447. doi:10.1080/02687038.2015.1037823

Whitworth, A. (2010). Using Narrative as a Bridge: Linking Language Processing Models with Real-Life Communication. *Seminars in Speech and Language, 31*,1, 64-75.

Whitworth, A., Perkins, L. and Lesser, R. (1997) Conversation Analysis Profile for People with Aphasia. Whurr: London.

Wilson, L. and Read, J. (2016). Do particular design features assist people with aphasia to comprehend text? An exploratory study. *International Journal of Language and Communication Disorders,* 51,(3), 346–354.

Wilson, S., Roper, A., Marshall, J., Galliers, J. R., Devane, N., Booth, T. and Woolf, C. (2015). Codesign for People with Aphasia Through Tangible Design Languages. *CoDesign*, 11(1), pp. 21-34. doi: 10.1080/15710882.2014.997744

Wilson, S., Galliers, J., Roper, A., Macfarlane, A. & O'Sullivan, D. (undated) Language-Light UX Guidelines Nine Interaction Design Guidelines for Aphasia. Available at: http://languagelightux.org/ [Accessed on: 22 May 2018].

Wilson, L. and Read, J. (2016). Do particular design features assist people with aphasia to comprehend text? An exploratory study. *International Journal of Language and Communication Disorders,* 51,(3), 346–354.

Woolf, C., Caute, A., Haigh, Z., Galliers, J., Wilson, S., Kessie, A., . . . Marshall, J. (2016). A comparison of remote therapy, face to face therapy and an attention control intervention for people with aphasia: A quasi-randomised controlled feasibility study. *Clinical Rehabilitation, 30*(4), 359-373. doi:10.1177/0269215515582074

Wray, F., Clarke, D. and Forster, A., 2018, Post-stroke self management interventions: a systematic review of effectiveness and investigation of the inclusion of stroke survivors with aphasia. *Disability and Rehabilitation,* 40(11), 1237–1251.

Zheng, C., Lynch, L. & Taylor, N. (2016). Effect of computer therapy in aphasia: a systematic review, Aphasiology, 30:2-3, 211-244, DOI: 10.1080/02687038.2014.996521

Appendices.

Appendix A. Prompt cards from the CAPPA.

People Prompts

People	**Difficult (Y/N)**
Wife/husband/partner	
sons and/or daughters	
grandchildren	
other family members	
friends	
neighbours	
health professionals	
strangers	
others (please specify)	

Situations Prompt

Situation	**Difficult Y/N**
at home with family	
at home with friends	
answering the door	
answering the phone	
ringing people	
at the shops	
attending meetings	
at social clubs	

on the bus	
in the car	
at church	
at mealtimes	
visiting family	
visiting friends	
playing board or card games	
other (please specify)	

Topics Prompt

Topic	Difficult Y/N
news/current affairs	
television	
own ideas	
plans for the future	
work	
hobbies	
holidays	
politics	
religion	
family	
friends	
other people	
past events/times	

sports	
pets	
immediate plans	
daily routine	
finance	
other (please specify)	

Appendix B. Sentence Production Treatment Fidelity Checklist: Phase 1 Exercise 2 Sentence Closure.

Clinician Behaviours	Present	Absent
Clinician establishes and maintains rapport by asking about the participant's experience of home practice during the previous week and addresses any concerns raised, including any which arise during the session (e.g. responding appropriately if participant reports feeling tired).	☐	☐
Clinician explains the purpose of this phase of treatment, setting it in the context of all three phases of treatment by explaining that this first phase focuses on improving retrieval of verbs on there own as these are the keystones of sentence production as well as production of words (nouns) associated with verbs to strengthen the connections between these. The next phase of treatment will focus on producing different types of sentences and the next stage of treatment will focus on helping the participant use practiced verb and sentences in real life communication.	☐	☐
Clinician explains the purpose of the new exercise using appropriate aphasia-friendly strategies, including explaining how each exercise relates to the client's deficit and addresses their goals.	☐	☐
Clinician introduces the participant to new exercise, explaining the purpose of each part of the slide, and demonstrating how to use cues if required. Clinician demonstrates sequentially the different types of cue available and explains the rationale for each. In particular, the clinician explains that this exercise aims to enable the participant to say a verb in a sentence repeatedly to strengthen the links between these words in the brain. The clinician explains that there is a cueing hierarchy which is designed to help the participant say each sentence more and more independently, and draws the participant's attention to the change in each step of the cueing hierarchy.	☐	☐
Clinician gives the participant appropriate feedback about their responses during treatment and encourages the participant to produce responses independently but giving progressive cues before providing correct examples when necessary.	☐	☐
Clinician explains which exercises are to be practiced during the following week, gives the participant (and the carer if appropriate) the opportunity to ask questions about this.	☐	☐
Clinician discusses when the client will practice, including how any potential barriers to practice will be addressed (e.g. reminder texts will be sent; participant to be assertive about completing exercises in the face of other demands).	☐	☐

Participant Behaviours	Present	Absent
Participant demonstrates understanding of the purpose of the session.	☐	☐
Participant appears motivated during the session.	☐	☐
Participant interacts with the clinician during the session, including proactively (e.g. by asking questions).	☐	☐
Participant is able to navigate all exercises chosen for home practice independently (or with help from a carer who is present) before the session is ended.	☐	☐
Participant demonstrates understanding of which exercises are to be practiced during the following week.	☐	☐

Appendix C. Lists of the PR verbs chosen by each participant in PhD study.

P1			**P2**		
Backheel	entertainment		**Aim**	feelings & senses	
Cheer	entertainment		**Change**	feelings & senses	
Cross	entertainment		**Enjoy**	feelings & senses	
Draw	entertainment		**Hate**	feelings & senses	
Head	entertainment		**Help**	feelings & senses	
Kick	entertainment		**Hope**	feelings & senses	
Lose	entertainment		**Love**	feelings & senses	
Miss	entertainment		**Remember**	feelings & senses	
Play	entertainment		**Struggle**	feelings & senses	
Pull	entertainment		**Suffer**	feelings & senses	

Save	entertainment		**Think**	feelings & senses	
Score	entertainment		**Try**	feelings & senses	**total 12 (30%)**
Tackle	entertainment		**Chatter**	communication mediums & modes	
Throw	entertainment		**Listen**	communication mediums & modes	
Volley	entertainment		**Read**	communication mediums & modes	
Win	entertainment	**total 16 (40%)**	**Speak**	communication mediums & modes	
Argue	communication mediums & modes		**Talk**	communication mediums & modes	
Phone	communication mediums & modes		**Understand**	communication mediums & modes	**total 6 (15%)**

Read	communication mediums & modes		**Carve**	food & drink	
Send	communication mediums & modes		**Eat**	food & drink	
Talk	communication mediums & modes		**Marinate**	food & drink	
Text	communication mediums & modes		**Season**	food & drink	
Understand	communication mediums & modes		**Stir**	food & drink	
Write	communication mediums & modes	**total 8 (20%)**	**Whisk**	food & drink	**total 6 (15%)**
Cough	health		**Brush**	nature & gardening	
Sleep	health		**Cuddle**	nature & gardening	

Take	health		**Stroke**	nature & gardening	**total 3 (7.5%)**
Trip	health		**Buy**	shopping	
Walk	health/travel	**total 5 (12.5%)**	**Pay**	shopping	
Catch	nature & gardening		**Shop**	shopping	**total 3 (7.5%)**
Feed	nature & gardening		**Drive**	travel	
Scratch	nature & gardening		**Visit**	travel	
Stroke	nature & gardening	**total 4 (10%)**	**Walk**	travel/health	**total 3 (7.5%)**
Buy	shopping		**Rent**	enertainment	
Pay	shopping		**Watch**	enertainment	**total 2 (5%)**
Shop	shopping	**total 3 (7.5%)**	**Wait**	health	
Drink	food & drink		**Walk**	health/travel	**total 2 (5%)**
Eat	food & drink	**total 2 (5%)**	**Save**	money	

Feel	feelings & senses	**total 1 (2.5%)**	**Withdraw**	money	**total 2 (5%)**
Walk	travel/health	**travel 1 (2.5%)**	**Clean**	housework	**total 1 (2.5%)**
Work	work	**total 1 (2.5%)**	**Wash**	personal care	**total 1 (2.5%)**

P3			**P4**		
Argue	communication mediums & modes		**clip**	nature & gardening	
Listen	communication mediums & modes		**Dig**	nature & gardening	
Phone	communication mediums & modes		**Grow**	nature & gardening	
Read	communication mediums & modes		**mow**	nature & gardening	
Talk	communication mediums & modes		**Plant**	nature & gardening	

Text	communication mediums & modes		**sow**	nature & gardening	
Understand	communication mediums & modes		**Water**	nature & gardening	
Write	communication mediums & modes	**total 8 (20%)**	**weed**	nature & gardening	**total 8 (20%)**
Enjoy	feelings & senses		**Drive**	travel	
Feel	feelings & senses		**fly**	travel	
Hate	feelings & senses		**see**	travel	
Help	feelings & senses		**swim**	travel/health	
Hope	feelings & senses		**Visit**	travel	
Struggle	feelings & senses		**Walk**	travel/health	**total 6 (15%)**
Think	feelings & senses		**dive**	work	

Try	feelings & senses	**total 8 (20%)**	**fight**	work	
Boil	food & drink		**hit**	work	
Cook	food & drink		**sail**	work	
Drink	food and drink		**serve**	work	
Eat	food & drink		**shoot**	work	**total 6 (15%)**
Fry	food & drink		**phone**	communication mediums & modes	
Peel	food & drink	**total 6 (15%)**	**Read**	communication mediums & modes	
Buy	shopping		**Talk**	communication mediums & modes	
Order	shopping		**Understand**	communication mediums & modes	
Pay	shopping		**Write**	communication mediums & modes	**total 5 (12.5%)**

Shop	shopping	**total 4 (10%)**	**Enjoy**	feelings & senses	
Brush	nature & gardening		**Feel**	feelings & senses	
Catch	nature & gardening		**think**	feelings & senses	
Grow	nature & gardening	**total 3 (7.5%)**	**Try**	feelings & senses	**total 4 (10%)**
Drive	travel		**Buy**	shopping	
Visit	travel		**Order**	shopping	
Walk	travel/health	**total 3 (7.5%)**	**Pay**	shopping	
Watch	entertainment		**sell**	shopping	**total 4 (10%)**
Wear	entertainment	**total 2 (5%)**	**Cook**	food & drink	
Sleep	health		**Drink**	food & drink	
Walk	health/travel	**total 2 (5%)**	**Eat**	food & drink	**total 3 (7.5%)**
Cut	work		**meet**	entertainment	
Work	work	**total 2 (5%)**	**Watch**	entertainment	**total 2 (5%)**
Clean	housework	**total 1 (2.5%)**	**swim**	health/travel	
Move	house	**total 1 (2.5%)**	**Walk**	health/travel	**total 2 (5%)**

Wash	personal care	**total 1 (2.5%)**	**move**	house	
			pack	house	**total 2 (5%)**

P5			**P6**		
Bake	food & drink		**Blend**	food & drink	
Blend	food & drink		**Boil**	food & drink	
Boil	food & drink		**Chop**	food & drink	
Chop	food & drink		**Cook**	food & drink	
Cook	food & drink		**Eat**	food & drink	
Drink	food & drink		**Fry**	food & drink	
Eat	food & drink		**Grate**	food & drink	
Grate	food & drink		**Grill**	food & drink	
Melt	food & drink		**poach**	food & drink	
Mix	food & drink		**Roast**	food & drink	
Peel	food & drink		**Steam**	food & drink	
Roll	food & drink		**Stirfry**	food & drink	**total 12 (30%)**
Season	food & drink		**Chip**	entertainment	
Weigh	food & drink		**Download**	entertainment	
Whisk	food & drink	**total 15 (37.5%)**	**Draw**	entertainment	

Catch	nature & gardening		**Lose**	entertainment	
Cut	nature & gardening		**Play**	entertainment	
Feed	nature & gardening		**Putt**	entertainment	
Grow	nature & gardening		**Score**	entertainment	
Plant	nature & gardening		**Sunbathe**	emtertainment	
Prune	nature & gardening		**Throw**	entertainment	
Water	nature & gardening	**total 7 (17.5%)**	**Win**	entertainment	**total 10 (25%)**
Enjoy	feelings & senses		**Email**	communication mediums & modes	
Feel	feelings & senses		**Listen**	communication mediums & modes	
Help	feelings & senses		**Phone**	communication mediums & modes	

Hope	feelings & senses		**Talk**	communication mediums & modes	
Love	feelings & senses		**Understand**	communication mediums & modes	
Think	feelings & senses	**total 6 (15%)**	**Write**	communication mediums & modes	**total 6 (15%)**
Phone	communication mediums & modes		**Drive**	travel	
Read	communication mediums & modes		**Fly**	travel	
Speak	communication mediums & modes		**See**	travel	
Talk	communication mediums & modes		**Swim**	travel/heallth	

Understand	communication mediums & modes	**total 5 (12.5%)**	**Visit**	travel	**total 5 (12.5%)**
Buy	shopping		**Feel**	feelings & senses	
Pay	shopping		**Like**	feelings & senses	
Shop	shopping	**total 3 (7.5%)**	**Think**	feelings & senses	**total 3 (7.5%)**
Watch	entertainment	**total 1 (2.5%)**	**Buy**	shopping	
Clean	housework	**total 1 (2.5%)**	**Pay**	shopping	**total 2 (5%)**
Visit	travel	**total 1 (2.5%)**	**Present**	work	
Work	work	**total 1 (2.5%)**	**Work**	work	**total 2 (5%)**
			Swim	heallth/travel	**total 1 (2.5%)**

Appendix D. List and categorisation of the personally relevant verbs chosen by the six participants in this PhD case series.

1. PR Verbs Chosen	2. No., of participants choosing PR Verb (n=6)	3. Topic (after Palmer et al., 2017 re categorisation of nouns)	4. Sub-topic (after Palmer et al., 2017 re categorisation of nouns)
Aim	1	feelings & senses	feelings & senses
Argue	2	communication mediums & modes	verbal
Backheel	1	entertainment	Sport (football)
Bake	1	food/drink	cooking
Blend	2	food/drink	cooking
Boil	3	food/drink	cooking
Brush	2	nature & gardening	pets
Buy	6	shopping	shopping
Carve	1	food/drink	cooking
Catch	3	nature & gardening	pets (cat – mouse)
Change	1	feelings & senses	feelings & senses
Chatter	1	communication mediums & modes	verbal
Cheer	1	entertainment	sport
Chip	1	entertainment	sport (golf)
Chop	2	food/drink	cooking
Clean	3	house	housework

Clip	1	nature & gardening	pets
Cook	4	food/drink	cooking
Cough	1	health	health
Cross	1	entertainment	Sport (football)
Cuddle	1	nature & gardening	pets
Cut	2	work	hairdressing
Dig	1	nature & gardening	gardening
Dive	1	work	navy
Download	1	entertainment	music
Draw	2	entertainment	Sport (football)
Drink	4	food/drink	food/drink
Drive	4	travel	travel
Eat	6	food/drink	food/drink
Email	1	communication mediums & modes	technology
Enjoy	4	feelings & senses	
Feed	2	nature & gardening	pets
Feel	5	feelings & senses	
Fight	1	work	navy
Fly	2	travel	
Fry	2	food/drink	cooking
Grate	2	food/drink	cooking
Grill	1	food/drink	cooking
Grow	3	nature & gardening	gardening

Hate	2	feelings & senses	
Head	1	entertainment	sport (football)
Help	3	feelings & sensnes	
Hit	1	work	navy (torpedo)
Hope	3	feelings & senses	
Kick	1	entertainment	sport (football)
Like	1	feelings & sensnes	
Listen	3	communication mediums & modes	verbal
Lose	2	entertainment	sport
Love	2	feelings & senses	
Marinate	1	food/drink	cooking
Meet	1	work	colleagues/clients
Melt	1	food/rink	cooking
Miss	1	entertainment	sport
Mix	1	food/drink	cooking
Move	2	house	
Mow	1	nature & gardening	gardening
Order	2	shopping	
Pack	1	house	
Pay	6	shopping	
Peel	2	food/drink	cooking
Phone	5	communication mediums & modes	verbal
Plant	2	nature & gardening	gardening

Play	2	entertainment	sport/music
Poach	1	food/drink	cooking
Present	1	work	using Powerpoint
Prune	1	nature & gardening	gardening
Pull	1	entertainment	sport
Putt	1	entertainment	sport
Read	5	communication mediums & modes	verbal/technology
Remember	1	feelings & senses	
Rent	1	entertainment	holidays
Roast	1	food/drink	cooking
Roll	1	food/drink	cooking
Sail	1	work	navy
Save (goal)	1	entertainment	sport (football)
Save (money)	1	money	
Score	2	entertainment	sport (goal football)
Scratch	1	nature & gardening	pets
Season	2	food/drink	
See	2	travel	
Sell	1	shopping	
Send	1	communication mediums & modes	verbal/technology
Serve	1	work	navy
Shoot	1	work	navy (anti-aircraft gun)

Shop	4	shopping	
Sleep	2	health	
Sow	1	nature & gardening	gardening
Speak	2	communication mediums & modes	verbal
Steam	1	food/drink	cooking
Stir	1	food/drink	cooking
Stirfry	1	food/drink	cooking
Stroke	2	nature & gadening	pets
Struggle	2	feelings & senses	
Suffer	1	feelings & senses	
Sunbathe	1	travel	
Swim	2	health/travel	
Tackle	1	entertainment	sport (football)
Take	1	health	
Talk	6	communication mediums & modes	verbal
Text	2	communication mediums & modes	technology
Think	5	feelings & senses	
Throw	2	entertainment	sport (football)
Trip	1	health (trip over)	
Try	3	feelings & senses	
Understand	6	communication mediums & modes	verbal
Visit	5	travel	
Volley	1	entertainment	sport (football)

Wait	1	health	
Walk	4	health/travel	
Wash	2	house	
Watch	4	entertainment	TV
Water	2	nature & gardening	gardening
Wear	1	entertainment	fashion
Weed	1	nature & gardening	gardening
Weigh	1	food/drink	cooking
Whisk	2	food/drink	cooking
Win	2	entertainment	sport
Withdraw	1	money	
Work	4	work	
Write	4	communication mediums & modes	verbal/technology

Appendix E. CAT Spoken Picture Description Transcriptions.

(Bracketed content not analysed as it represents content subsequently self corrected or extraneous comments. **Bold content** = ICWs. *Italicised content* = inappropriate ICWs. UTS = undetermined thematic structure). Discourse markers are omitted from the PAS analysis as per Webster et al. (2007)).

P1. Assessment 1.

The **man** is er **sleeping** = verb + 1 argument

Elderly erm UTS

The **boy** is erm = verb + 1 argument

The **car** UTS

the **coffee table** UTS

erm the **mug** UTS

erm **paper** UTS

Erm the **book** is **under** the coffee table = verb + 2 arguments

Erm the **cat** is **getting** the (finger spells) **goldfish**. = verb + 2 arguments

Erm the **books on** the UTS

throwing all over the place = verb + 1 argument

The *flower* is there = verb + 1 argument

CD and everything else UTS

Sofa UTS

The *man* (target: boy) is **sitting** up. = verb + 1 argument

P1. Assessment 2.

The **man** is **snoring** = verb + 1 argument

The **baby** UTS

the baby is = verb + 1 argument argument omitted

baby is **car** (but that's not true) = UTS

The **plant**'s = verb + 1 argument (omitted argument)

The **cat** is **putting** the **paw in** the **fish** = verb + 3 arguments (incorrect DET)

The **books** are *tossing* about the *sofa* = verb + 2 arguments

The man is putting the **legs up** = verb + 2 arguments (incorrect DET)

Coffee table UTS

cup erm UTS

The book is on the table = verb + 2 arguments

books are **on** the table. = verb + 2 arguments (omitted DET)

Erm **radio** is blasting away = verb + 1 argument (omitted DET)

P1. Assessment 3.

The **baby** is = verb + 1 argument (argument omitted)

erm the baby is **playing** = verb + 1 argument

Erm the **radio** Is blaring = verb + 1 argument

Man is **snoring** = verb + 1 argument

and **feet up** UTS

The **cat** is **catching** the *bowl*. = verb + 1 argument

The *flower* is (unintelligible). = verb + 1 argument

The **books** is erm **tumbling** (dialectal use of "is") = verb + 1 argument

The *settee* UTS

erm erm **coffee table** UTS

mug UTS

Erm **books** are **on** the table erm = verb + 1 argument (omitted DET) and then books again. (So that's it) UTS

P1. Assessment 4.

Erm the **man** UTS

elderly man is **snoring** = verb + 1 argument

Feet up UTS

Erm coffee in in coffee **coffee table** UTS

The **baby** is = verb + 1 argument (argument omitted)

baby is erm **car** = UTS

Erm the **plant** is **OK** = verb + 2 arguments

but the erm the **cat** is **getting** the **goldfish out** of the **bowl** = verb + 3 arguments

The **books** are **tumbling** off the **shelf** = verb + 2 arguments

Erm the **radio** is OK. = verb + 2 arguments

P2. Assessment 1.

The **man** is **sleeping** = verb + 1 argument

resting his **legs** *in* the **table** = verb + 2 argument (incorrect PREP)

And the **cat** . . . UTS

the cat is **trying** to /fi/ **fish =** verb + 1 argument

catch a **fish** = verb + 1 argument

And the **books** are **falling** on the er man = verb + 2 arguments

The **little boy** is **playing car one** car = verb + 2 arguments (omitted PREP)

. . . the the **toddler** he **notice =** verb + 1 argument (omitted argument) (omitted morpheme)

the the the the he she he **look** at (I can't say it) = verb + 1 argument (omitted argument) (omitted morpheme) he **notice** the the cat erm er erm = verb + 2 arguments (omitted morpheme)

wake up wake up because it's (I can't say it). = verb (abandoned utterance)

P2. Assessment 2.

The **man** is **sleeping** with with (no sorry again). = verb + 1 argument (omitted argument)

The man is sleeping = verb + 1 argument

The **cat** is **trying** to **get** a **fish**, fishes, fish, fish. = verb + 2 arguments

The the cat UTS

The *bookcas*e the the bookcase UTS

The **books** are **falling** on the man = verb + 2 arguments

going to fall down = verb

The **toddler** is **playing** with a **car** = verb + 2 arguments

and the man er the the the he /intonate/er the er UTS

is er er the the toddler is **look** the cat = verb + 2 arguments (omitted morpheme) (omitted PREP)

and he **said** look the (I can't say it) look at cat = verb + 2 arguments (omitted DET)

is the the *picture booklets* which em = UTS

books are falling = verb + 1 argument (omitted DET)

In the **shelves** are a **plant** and **record player** and two **stereo** er oh **speakers** = verb + 2 arguments (incorrect PREP)

The **table top** with a **cup of tea** UTS

and the the /t/ the - in the **living room** I notice the **coffee table** with cup of tea and books and **magazine** on the in the **centre** = verb + 3 arguments (omitted DET)

P2. Assessment 3.

The **man** is **snoozing** = verb + 2 arguments

The **cat** is **playing** with the **fish** = verb + 2 arguments

The the the **plant pot** is on the **shelf** = verb + 2 arguments

The **toddler** is toddler is **trying** to **distract** the man = verb + 2 arguments

The toddler is = verb + 1 argument (abandoned utterance)

the man is = verb + 1 argument (abandoned utterance)

the toddler is distract the man = verb + 2 arguments (omitted morpheme)

The the cat is = verb + 1 argument (abandoned utterance)

the cat UTS

The **books** are **falling down**. = verb + 1 argument

The man is **resting** her his **head** = verb + 2 arguments

The man is resting his **feet on** the **table** = verb + 3 arguments

The man no the person is = verb + 1 argument (abandoned utterance)

the man is = verb + 1 argument (abandoned utterance)

the man is = verb + 1 argument (abandoned utterance)

the man is /koti/cold tea = UTS

The man is = verb + 1 argument (abandoned utterance)

the man is distract = verb + 1 argument (omitted morpheme) (omitted argument)

The man is = verb + 1 argument (abandoned utterance)

The **tea** is cold = verb + 2 arguments

The **toddler** is playing on = verb + 1 argument (omitted argument)

The toddler is playing with a **car** = verb + 2 arguments

P2. Assessment 4.

The **man** is **sleeping** and with a **cup of tea** *in* the **table** = verb + 3 arguments

The **cat** is **trying to fish** = verb + 1 argument

The cat is trying to trying to **catch** the **fish** = verb + 2 arguments

but the cat is the cat is **not** = UTS

/*vertently*/ (T: inadvertently) is the the the **books** are **falling down** on the **head** of the sleeping man

= verb + 2 arguments

And the **toddler** toddler is = verb + 1 argument (abandoned utterance)

/wi/ er the man UTS

the toddler is man = UTS

the toddler is trying **to help** no **point out** the **danger** with the the books = verb + 3 arguments

(Maybe may be may be the – no (pointing to plant)).

The **CD** and the **radio** on the **shelf** *is* **safe** because because is the the cat is no = verb + 2 arguments

(abandoned utterance) (incorrect inflection)

the cat is = verb + 1 argument (abandoned utterance)

the cat is safe (no sorry) = verb + 2 arguments

OK the CD and the radio *is* safe because the the the cat is trying to try to (oh god)

the cat is trying to catch a fish but yeah = verb + complex VP (incorrect inflection)

The man is **recently** recently **had** a cup of tea = verb + 2 arguments

(And may and and the the maybe)

is a er books as well in the er the in in the books er *in* the table = UTS

Er recently had a cup of tea = verb + 1 argument (omitted argument)

and the because the the man is = verb + 1 argument (abandoned utterance)

table is books UTS

books on the feet UTS

man is snoozing and **resting** his **weary legs on** the the books = verb + 3 arguments Omitted DET

P3. Assessment 1.

Er er **cat** er is er. = verb + 1 argument (omitted argument) (omitted DET)

The cat is er oh her oh oh god um er **fishing**. = verb + 2 arguments

Erm er the cat is no = verb + 1 argument (abandoned urtterance)

the **books** are **falling** on the **man** = verb + 2 arguments

Er er **sleeping**. = verb

Er er **plant** UTS

Er **playing** = verb

Erm er **feet radio** on it. UTS

Speaker /speet/ **speakers** UTS

bald head UTS

erm **cup of tea UTS**

P3. Assessment 2.

Erm er the **cat** is **fishing** = verb + 1 argument

Erm **books** are **falling** on the **man** = verb + 2 arguments (omitted DET)

The **baby** is **playing** = verb + 1 argument

Er **sleeping** = verb

er er er **sit** erm = verb

The man is er sleeping er (sigh) er er er er (sigh) er oh god er (sigh) um oh erm er = verb + 1 argument

socks on /hof/ (Target: off) UTS

P3. Assessment 3.

Er er the **plant** is **on** the **shelf**. = verb + 2 arguments

Er the **cat** is **trying** to **fish** /*kat*/ (target: catch) the **fish** = verb + 2 arguments

Er we um er the **books** are **falling** on the **man** = verb + 2 arguments

The man is **asleep** = verb + 1 argument

Er er the **child** is **playing** with the **toy**= verb + 2 arguments

Er er er er the er the er the **table** is = verb + 1 argument (argument omitted)

the table is = verb + 1 argument (argument omitted)

the table is on the **floor** = verb + 2 arguments

Erm the er the **cup** and the **book** on erm are **on** the table = verb + 2 arguments

Erm the leg um the the the /l/ um UTS

the man has the man has his **legs crossed** = verb + 3 arguments

(Mmm the er mmm no).

P3. Assessment 4.

The man **baby** is **playing** with the **toy** = verb + 2 arguments

Er the **man** is **sleeping** on the **chair** = verb + 2 arguments

Er the **books** are **falling** on **him** = verb + 2 arguments

Er the man has has a oh er has er (shakes head) = verb + 1 argument (abandoned utterance)

The /gereo/ (target: stereo) is on the **shelf** = verb + 2 arguments

The **cat** is playing = verb + 1 argument

erm er er the cat is **trying** to **catch** the **fish** = verb + 2 arguments

Er the man is is the man has a **bald head** = verb + 2 arguments

Er the the man is = verb + 1 argument (argument omitted)

the man is = verb + 1 argument (argument omitted)

no the the man is **wearing** erm a **tie** = verb + 2 arguments

The er **coffee table** has a **mug on** it = verb + 3 arguments

P4. Assessment 1.

There's a **boy** er **playing** with a **car** = verb + 2 arguments

And there's a **man sleeping** with his **feet up** *with* a *chair* = verb + 3 arguments

And there's a erm **cat** er **fishing** the *fight* /fe fai/ erm er /fai/ fight fight no = verb + 1 argument

pushing the **books over** the top of er erm (hell erm) = verb + 2 arguments (argument omitted)

There's a *music* UTS

(and erm that's it).

Toffee /tot/ toffee (Target: coffee) UTS

(that's it)

P4. Assessment 2.

Er **cat** is **fishing** = verb + 1 argument (omitted DET)

The **fish** fish UTS

And the **book fall** on the **head** of the *head* **while** he's sl **sleeping** = verb + complex phrase

And the **boy playing** with the *chair* erm **car**. (erm . . . . that's it) = verb + 2 arguments (omitted IS)

he's the erm UTS

He's fast asleep. = verb + 1 argument

His **feet** on the *chair*. UTS

Er **cup** on the UTS

Fall down = verb

looks like it could fall down = verb +1 argument

(Erm the well that's it).

P4. Assessment 3.

The **cat** is **fishing** for */fat/* /fi/ is **fish** = verb + 2 arguments

And he **knocks over books** *with* **mistake** = verb + 3 arguments (omitted DET) (incorrect PREP)

And the book on the **man** who's **fast asleep** verb + complex phrase

On top of him UTS

The the **little boy** his **car** erm UTS

And at the same time it's to get his er it erm no his erm (abandoned utterance "to get his attention"??) = UTS (abandoned utterance)

He get's away = verb + 1 argument

He get's erm = verb + 1 argument (omitted argument)

But the oh the man is asleep = verb + 2 arguments

So he **sleep** on er the **chair** with his **feet up** = verb + 3 arguments (omitted morpheme)

And the /fo/ *toffee* **tea** on the **table** UTS

He's asleep = verb + 2 arguments

but the **music played** = verb + 1 argument

(And erm that's it).

P4. Assessment 4.

There's a *car* a **cat** = verb + 2 arguments

And he's *booking* (T: hooking) book book the **fish** = verb + 2 arguments

And after also over the **books** UTS

Over the **old man** who's **asleep** = verb + complex phrase

He's **sleeping** with his **feet up** = verb + 2 arguments

And er **book underneath** his *bed* UTS

And there was a **little boy trying to play** the **car** = verb + complex phrase (omitted PREP)

And er the man **listen** to the **music** (unintelligible) = verb + 2 arguments (omitted morpheme)

There's erm a **flower** on the **shelf** of the car the */kuf/* = verb + 3 arguments

(That's it).

P5. Assessment 1.

The **man asleep** in a (*car* in a couch no) **armchair** = UTS

Cat is **after** the **fishes** = verb + 2 arguments (omitted DET)

The **books falling down going on** his **head** = verb + 2 arguments (omitted ARE)

Little boy is **playing** with a **car** = verb + 2 arguments (omitted DET)

The **book** is **under** the **table** = verb + 2 arguments

and the **cup** is on the side on the *chair* = verb + 2 arguments

(I can't think what they are – chair not chair. I go back to that one).

That's a **plant** up there = verb + 3 arguments

This is a **radiogram** = verb + 2 arguments

His **feet** are on the /kut/ = verb + 2 arguments

P5. Assessment 2.

The **boy's playing** with **toys** = verb + 2 arguments

The **cat** is **after fish** = verb + 2 arguments

The **books** are **fallen down** = verb + 2 arguments

He he he's playing = verb + 1 argument

he's play (abandonded utterance) = verb + 1 argument

(no let's go on to)

He's **sleepin**g = verb + 1 argument

He's got **feet on towels** = verb + 3 arguments (omitted possessive PRO)

He's got a **cup** and a **book** = verb + 2 arguments

This is a speaker – **speakers** and a **radio** = verb + 2 arguments

(And – did I say book?)

He’s asleep in a **armchair**. (“a” as opposed to “an” = dialectal) = verb + 3 arguments

(prompt: what’s happening here June?)

Cat’s did UTS

The cat’s **catching** the fish = verb + 2 arguments

There’s a **plant** = verb + 2 arguments

P5. Assessment 3.

The **cat** is **trying** to **get** the **fish** = verb + 2 arguments

The **man** is **resting** = verb + 2 arguments

Erm A **plant pot, radio** and **speakers** UTS

Little boy playing with his **lorry** or a (unintelligible) = verb + 2 arguments (omitted DET) (omitted IS)

A **cup** UTS

And erm something is **falling on** his **head** = verb + 2 arguments

A **book** is falling on his head = verb + 2 arguments

He’s **got** his feet on some **towels** = verb + 3 arguments

Table UTS

A **book** UTS

P5. Assessment 4.

The **cat** is **after** the **fish** = verb + 2 arguments

The **books** are **falling down** = verb + 2 arguments

The **little boy**’s **got** a **toy** = verb + 2 arguments

The **old man asleep** UTS

He’s got his **feet** on the **cushions** = verb + 3 arguments

He’s got the *bottle* – (not the bottle) – **book** under the *chair* = verb + 3 arguments

He’s got a **cup** and **saucer** on the table = verb + 3 arguments

He’s got a **hifi** = verb + 2 arguments

(Prompt: What’s happening here?)

The boy’s got a little **motor car** = verb + 2 arguments

(Prompt: What’s happening here?)

The cat’s getting in the **fishbowl** = verb + 2 arguments

P6. Assessment 1.

The oh the ah the erm oh the er the ah erm **car** UTS

the **cat** UTS

fish UTS

erm **stereo** UTS

man UTS

cup UTS

erm **book** UTS

erm er oh the the p p p UTS

P6. Assessment 2.

The the erm the **cat** is = UTS

erm erm **fish** UTS

guy is (pointing to books falling) UTS

guy is **books** books books UTS

Cup UTS

He is erm (pointing to boy) = UTS

guy is no = UTS

the **boy** is **car** = UTS

Guy no (pointing to stereo) UTS

P6. Assessment 3.

Books. UTS

Cat is = UTS

So it's cat = UTS

Pl **plant** UTS

Speakers UTS

(Writes "CA" Target = cassette).

Boy UTS

Car UTS

Man UTS

Cup UTS

Book UTS

Table UTS

Chair UTS

Book UTS

This one (pointing to stereo) UTS

Book is is erm = UTS

Book is dunno = UTS

P6. Assessment 4.

Cat is erm = UTS

cat is erm (writes "FISH") = UTS

(Pointing to plant writes "PLANT") /p p/

Erm **boy** UTS

Car UTS

He erm UTS

Mug UTS

Table UTS

Erm **book** UTS

Chair UTS

Man UTS

Chair no book. UTS

(Writes “ST”) -> **stereo** UTS

(Writes “SPEAKERS”) -> speakens no **speakers** UTS

Appendix F. Exit interview rating scale questions PwA.

Did you have previous experience of using a computer?

✗ ✓

0 None	1 A little	2 Average	3 A lot	4 Very experienced

How did you find doing therapy on a computer?

✗ ✓

0 very difficult	1 A little difficult	2 OK	3 Quite easy	4 Very easy

Did it work?

✗ ✓

0 No change	1 A little change	2 Not sure	3 Some change	4 A lot of change

Have you used the words practised on the computer in your daily life?

✗ ✓

0 No	1 Yes a few days	2 Yes some days	3 Yes most days	4 Yes everyday

How much help did you need with the computer therapy?

✗ ✓

0 Always	1 A lot	2 Some	3 A little	4 None

Would you use it again?

✗ ✓

0 No	1 Perhaps	2 Not sure	3 Very likely	4 Yes definitely