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.

Published in the American Journal of Speech Language Pathology,

Available at: https://pubs.asha.org/doi/10.1044/2021_AJSLP-21-00120

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)[1] 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

[1] 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.

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 studies, three of which were included by Conroy et al[2]. 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*

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

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 (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)[3]. 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

[3] 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.

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.

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:

1. What are the levels of evidence for sentence production treatments for people with aphasia?
2. 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)?
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 (within and across level generalization)?

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?
5. 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[4].

[4] 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.

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 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

279 papers published by Baker in 2012 were again useful. Baker points out that the concept of
280 an active ingredient of treatment is again drawn from pharmacology. In terms of
281 *behavioural* treatments, Baker also points out that it is treatment *activities* which are the
282 potential active ingredients of treatments. She makes a useful distinction between
283 treatment activities which comprise *therapeutic inputs* (i.e. acts of the clinician) - such as
284 giving a cue, and those which are *client acts* or responses expected of the client - such as
285 repeating a word. Thus, in order to inform research question 5, these were recorded. Also
286 influential was a paper by Byng, Nickels and Black (1994). They discuss the need for a
287 "theory of therapy" to inform the process of determining what type of aphasia treatment is
288 appropriate for a client. They discuss that this includes not only describing treatment
289 activities but also the nature of the *interaction* between therapist and client since this is also
290 likely to influence the outcome of treatment (e.g. how/whether the rationale for treatment
291 is discussed and whether corrective feedback is given during treatment). Thus these aspects
292 of treatment were also charted (when reported) to inform the potential active ingredients
293 of treatment.

294 The data extracted from the reviewed studies were entered into a Microsoft Excel
295 spreadsheet and are reported in Tables 2 and 3 below. All 33 studies were reviewed by the
296 first author only (JH), with 10% of studies blind reviewed by the second and third author
297 respectively (MC and LD: three studies each). There was 95% agreement between
298 reviewers with any disagreements resolved via discussion.

299 The results of treatment for *individual* participants are described whenever these are
300 reported. When a study reported results for the *group* only, this is reported separately i.e.
301 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.

Insert Figure 1 about here.

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).

Insert Table 1 about here

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[5]. 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 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).

[5] 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

Insert figure 2 about here

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.

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

391 untrained verbs is summarized in Table 2 Columns 4 - 7. Significance levels and effect sizes
392 are given for each individual participant when available.
393
394 Insert Table 2 about here.
395
396 The impact of treatment on the production of a) *trained* and *untrained verbs in isolation* was
397 assessed in only 7 studies (21%) for trained verbs, and 14 (42%) of studies for untrained
398 verbs (see Table 2, columns 4 and 5). The seven studies investigating impact of treatment on
399 production of *trained* verbs in isolation had a total of 33 participants. Individual results are
400 reported for 14 of these participants with 10 individuals showing significant improvement
401 (71%) (confirmed by statistical analysis for three of these participants: $p<.05$). Two studies
402 reported group results with both finding significant improvement: (Hoover et al., 2015
403 (n=12); Schneider & Thompson, 2003 (n=7)). With regard to the production of *untrained*
404 verbs in isolation, this was investigated in 14 studies with 62 participants. Individual results
405 are reported for 42 of these participants, 15 of whom (36%) showed significant
406 improvement (confirmed by statistical analysis for eight of these participants: $p<.05$).
407 Significant improvement was also reported in two group studies: Bazzini et al. (2012) (n=8)
408 and Hoover et al. (2015) (n=12).
409 The impact of sentence treatments on b) *sentence* production involving either
410 trained or untrained verbs is summarised in Table 2 columns 6 and 7. Sentence production
411 involving *trained* verbs was assessed in 20 (67%) studies which involved 75 participants,
412 with individual results reported for 69 of these. Significant improvement in sentence
413 production was reported for 57 participants (83%) (confirmed by statistical analysis for 25 of
414 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[6] 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

[6] 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.

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 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.

Insert figure 3 about here

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

554 treatment studies generally reported treatment protocols to a good level of detail,
555 researchers must improve the fidelity of reporting particularly with regard to the dose of
556 treatment given, and with regard to the views of participants on sentence treatments which
557 has not been investigated to date.
558
559 **Levels of evidence for sentence production treatments**.
560 Research question 1 addressed the levels of evidence for sentence production
561 treatments in aphasia. The evidence was predominantly Level 4 as it was dominated by case
562 series and single case studies. Whilst it is now acknowleged that well-controlled case series
563 can be used to support clinical decision making, this should generally only be the case if high
564 quality systematic reviews are not available (https://www.cebm.ox.ac.uk/resources/levels-
565 of-evidence/explanation-of-the-2011-ocebm-levels-of-evidence). There is, therefore, a need
566 for sentence production treatment studies using designs which constitute higher levels of
567 evidence such as well-designed, larger scale RCTs (e.g. Palmer et al., 2019). However, it
568 should also be noted that single case reports and case series *are* well suited to the current
569 phase of sentence treatment research, which is predominantly Phase I and II (as defined by
570 Robey and Schultz, 1998) in that sentence treatment protocols are still being refined and
571 optimal dosages established, for example. The detailed description of treatment afforded by
572 case reports and series (which is often omitted in RCTs) is also likely to be important in
573 informing the sentence treatments used in such larger scale studies. Indeed, the significant
574 role that case reports have played in advancing medical science in this way is now
575 acknowledged (e.g. Murad et al., 2018).
576
577

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 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

626 procedures in aphasia treatment research. However, there are an increasing number of
627 studies which demonstrate that it is indeed feasible to implement treatment fidelity
628 measures, noting also that these must be planned and accounted for from initial study
629 design (e.g. Carragher, Brooke, Worrall, Thomas, Rose, Simmons-Mackie et al., 2019;
630 Conlon, Braun, Babbitt & Cherney, 2020; Hilari, Behn, Marshall, Simpson, Thomas &
631 Northcott et al., 2019).
632
633 **Evidence of treatment effects and generalization for sentence production**
634 **treatments.**
635 Conclusions regarding the impact of sentence production treatments on production
636 of *trained verbs in isolation* must be treated with caution because this was assessed in only
637 a small proportion of sentence treatment studies, with a similar pattern in relation to the
638 assessment of *untrained* verb production in isolation. Having said this, sentence treatments
639 were effective in improving the production of *trained* verbs in isolation for the majority
640 (71%) of participants. This compares to improvement for 80% of participants in verb-in-
641 isolation treatments (Hickin et al.,2020). Regarding production of *untrained* verbs in
642 isolation, this improved for 36% of participants following sentence treatments compared to
643 15% of participants following verb-in-isolation treatments (ibid). This provides indicative
644 evidence that sentence treatments may generalise to untrained verb-in-isolation production
645 more successfully than verb treatments (within level generalisation). However, as stated
646 earlier, because the number of sentence treatment studies that evaluated production of
647 untrained verbs in isolation was small, this finding must be treated with caution. Given that
648 sentence treatments are *ipso facto* investigating treatment involving the production of
649 treated verbs simultaneously with other sentence constituents (i.e. not in isolation) it is

650 understandable that verb production in isolation is not routinely assessed. However, if a
651 research aim is to investigate the importance of verb retrieval *per se* to sentence production
652 then assessing verb retrieval in isolation as well as in a sentence would seem important. In
653 particular, routine assessment of verb retrieval in isolation (as is the case after verb in
654 isolation treatment) would make comparison of the two types of treatment easier.
655 There are a number of challenges in interpreting the effect of sentence treatments
656 on *sentence* production. These include the differing ways in which generalisation is assessed
657 and how the *significance* of improvement is determined. Thus, with regard to within level
658 generalization, there were two almost equally frequent ways of assessing this: i) assessing
659 sentence production in an *untrained task* (such as a standardised sentence production test),
660 and ii) assessing *untrained exemplars* of sentences (i.e. using matched but untrained *verbs*),
661 with a minority of studies assessing generalisation via production of untreated sentence
662 (syntactic) *structures* (6 studies). Interpretation of the evidence would be assisted by
663 reaching a consensus on how to assess generalisation. The evidence would also be
664 strengthened by the use of inferential statistics to evaluate the significance of treatment
665 effects. Statistics were used to confirm treatment effects just under 50% of the time in
666 relation to both trained and untrained sentence production. With these caveats in mind,
667 sentence treatments appeared effective in improving *trained* sentence production (for 83%
668 of participants) and *untrained* sentence production (for 59% of participants).
669
670 **Broader generalization of sentence production treatments for people with aphasia.**
671 Conclusions regarding the impact of sentence treatments on *functional*
672 *communication* must be tentative because this was not routinely assessed. Nonetheless,
673 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

698 context (in 91% of studies) making it a likely active ingredient. However, this task varied
699 considerably in the way it was implemented. For example, some studies allowed and even
700 encouraged the use of pronouns (as well as full lexical forms) in target sentences with the
701 aim of increasing the variety of utterances attempted and making these potentially more
702 functional (e.g. Carragher et al., 2015; Nickels et al., 1991). Other studies specifically
703 discouraged the production of pronouns, notably those investigating VNeST wherein the
704 rationale is that the treatment is predicated upon strengthening (priming) the semantic
705 network of a verb by requiring production of a verb alongside its arguments, and the
706 production of pronouns circumvents this process (see e.g. Edmonds 2016).
707 Other common ingredients of sentence production treatments were the use of a
708 *picture prompt* to stimulate sentence production, the use of *written* cues and a *focus on the*
709 *thematic roles* of a verb's arguments during treatment. The use of a picture prompt to
710 stimulate sentence production varied across studies. Edmonds and colleagues specifically
711 state that pictures are *not* used in VNeST because they can constrain a verb's meaning to
712 what is imaged and this may limit potential responses. Edmonds et al. regard this as
713 particularly problematic for verbs whose semantic networks are "loose" in comparison to
714 nouns (Edmonds, 2016, p.126). The use of pictures is therefore theorized to potentially
715 constrain the amount of activation (i.e. strengthening) of a treated verb's network, and to
716 limit engagement of autobiographical and episodic memory which are also regarded as
717 active ingredients of VNeST. The use of pictures to prompt sentence production thus
718 warrants further investigation (as does the use of video stimuli – see discussion below).
719 A focus on the thematic roles of a target verb and how they map onto the syntactic
720 structures was a key component of studies investigating mapping and PAS treatments and
721 this raises an important question regarding how key to treatment is raising the

722 *metalinguistic awareness* of participants. Metalinguistic awareness was raised using a
723 variety of strategies including colour cues, wh- questions, icons and written labels of
724 thematic roles accompanied, in all cases, by discussion. In most studies, corrective feedback
725 was also given, but in a minority of studies it was not (e.g. Byng et al., 1994; Nickels et al.,
726 1991) with the specific aim of encouraging participants to self-monitor their production. A
727 study carried out by Webster and Gordon (2009) gives an insight into how important latent
728 metalinguistic awareness may be to the success of treatment. They report two different
729 treatments given to their participant, only one of which was successful in improving verb
730 and sentence production. The first treatment given by Webster and Gordon was a mapping
731 treatment. However, their participant was confused by the use of linguistic terminology
732 during treatment: this caused her to become frustrated and to disengage with the
733 treatment which was ultimately ineffective. The second treatment specifically did *not* aim to
734 improve metalinguistic awareness and used a noun-verb association task, with no linguistic
735 terminology used and no discussion about errors in sentence production. This second
736 treatment was accepted by the participant and resulted in statistically significant
737 improvement in trained verb and sentence production. It may, therefore, be useful to
738 establish the level of metalinguistic awareness of participants prior to planning treatment
739 (e.g. by discussing how much "grammar" they know), and/or to trial metalinguistic
740 treatments to establish their acceptability to a person with aphasia.
741 The least used ingredients of sentence treatments included the use of *conversation-*
742 *based tasks* (to facilitate generalisation of sentence production skill to real life
743 communication), *video* and *gesture cues*. Despite the lack of explicit treatment for
744 conversation and discourse skills, there is evidence of generalization to these contexts
745 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.

794 **References.**

795 Adelt, A., Hanne, S., & Stadie, N. (2018). Treatment of sentence comprehension and
796 production in aphasia: Is there cross-modal generalisation? *Neuropsychological*
797 *Rehabilitation, 28*, 937-965. doi:10.1080/09602011.2016.1213176

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

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

803 Ballard K.J. and Thompson C.K. (1999) Treatment and generalization of complex sentence
804 production in agrammatism. *Journal of Speech, Language & Hearing Research, 42*,
805 690-707.

806 Bastiaanse, R., Hurkmans, J. & Links, P.(2006) The training of verb production in Broca's
807 aphasia: A multiple‐baseline across‐behaviours study, *Aphasiology, 20*(2-4), 298-
808 311, DOI: 10.1080/02687030500474922

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

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

815 Blankestijn-Wilmsen, J., Damen, I., Voorbraak-Timmerman, V., Hurkmans, J., Koning, B. D.,
816 Pross, A., & Jonkers, R. (2017). The effect of static versus dynamic depictions of
817 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

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

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.

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.

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.

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

Marshall, J. (1995). The mapping hypothesis and aphasia therapy, *Aphasiology, 9*(6), 517-539 doi: http://dx.doi.org/10.1080/02687039508248712

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.

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

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

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.

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)

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

Figure Legends.

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

Figure 2. The Number of Hours of Sentence Production Treatment Reported in the Reviewed Studies. (Schneider & Thompson (2003); Fink et al., (1992) 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).

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

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 argument on the speed and accuracy of production.	Verb Argument Structure (VAS) treatment	6/8 Ps had 30 hours treatment, 1P had 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	to establish if VNeST is effective..	VNeST	unclear but appears to vary from 4 - 6 weeks, twice a

	(n=4) (2 TMA, 2 conduction)			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 & Budson, (2015)	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 group context or combined	6.75 hours, over two days, for 6 weeks (2.25 hours a week for each treatment 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

1068

1069 **Table 2. Sentence treatments: ingredients, number of verbs/sentences treated and impact of treatment at each level of communication.**

1.Study	2.Ingredients of Treatment	3.Number of Verbs/Sentences Treated and Type	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
Bastiaanse, Hurkmans, & Links (2006) (n=2)	Verb Production at the Word and Sentence Level (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**)	60 verbs treated, half transitive and half intransitive	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)	Verb Argument Structure (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**)	Unclear: ?20 sentences demanded for each of 16 stages of treatment	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)	Predicate Argument Structure (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**)	60 verbs (20 unergative, 20 transitive and 20 ditransitive); SV, SVO & SVOC sentences treated	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**)	number not stated; all agentive verbs; SVO sentences	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**)	20 sentences (plus 20 control sentences verb matched for frequency, transitivity & valency)	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**)	30 sentences in 3 sets of 10	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)	Verb Network Strengthening Treatment (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**)	10 Vs (plus 10 controls)	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**)	10 Vs (plus 10 controls)	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**)	10 Vs (plus 10 controls)	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**)	10 verbs	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**)	passive , active & object cleft sentences (n not reported)	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)	Constraint Induced Language Treatment (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**)	57 verbs (32 personally relevant, 5 light; 20 unspecified); SVO -> SVO + adjunct	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)	Constraint Induced Language Treatment (HELPSS): **picture prompt; story completion; +/- a spoken sentence prime;** followed by a **spoken prompt question; spoken production of verb; external** & **internal argument** (**7**)	11 sentence types	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**)	81 transitive verbs, divided into 9 functional sets (27 verbs in 3 sets treated in each treatment condition)	**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**)	number not reported; material verbs + be & have	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**)	32 Vs in drill phase; specific verbs not targeted in communication phase	**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)	24 verbs in 41 sentences	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**)	60 verbs (30 intransitive, 30 transitive)	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**)	10 verbs (3 change of possession, 4 reverse role and 3 communication verbs) in 3 argument structures	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**)	18-29 verbs	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**)	number of verbs not stated; active versus passive sentences	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**)	no target list of verbs	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)**	number of verbs not stated; all action verbs; SV -> non-reversible SVO -> reversible SVO	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**)	32 (plus 32 personally/ functionally relevant verbs which were not assessed)	**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**)	6 verbs (with 144 exemplar sentences); active, passive, subject cleft & object cleft sentences trained.	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)	mapping treatment: verb component: **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**)	10 verbs in 30 passive sentences, each sentence in past, present & future tense (total = 90 sentences)	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**)	40 verbs (10 each of 2 or 3 place; motion or change of state verbs)	**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 * 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**)	30 (in 3 sets of 10); transitive action verbs -> experiential verbs in canonical sentences -> action verbs in passive, cleft subject, cleft obect, subject relative clauses & object relative clause sentences	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**)	8 sentence types from the HELPSS with 15 exemplars of each (total = 120 sentences)	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**)	30 verbs (2 sets of 15) (40 in 2 sets of 20 for 1P)	**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**)	48 personally relevant verbs	**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**)	80 everyday verbs (in 2 sets of 40)	**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**)	100, 2 argument, transitive, functionally relevant verbs (in 2 sets of 50)	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)**

1070

Figure 1

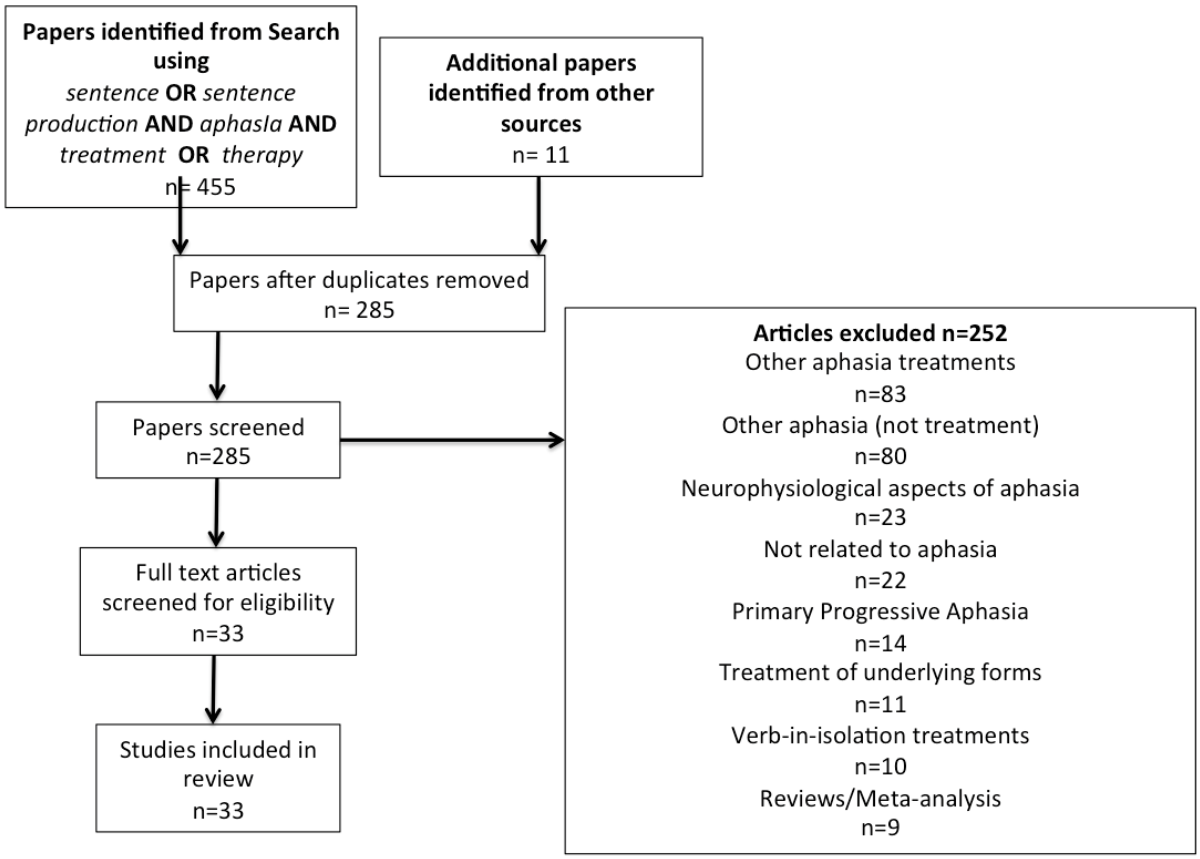


Figure 2.

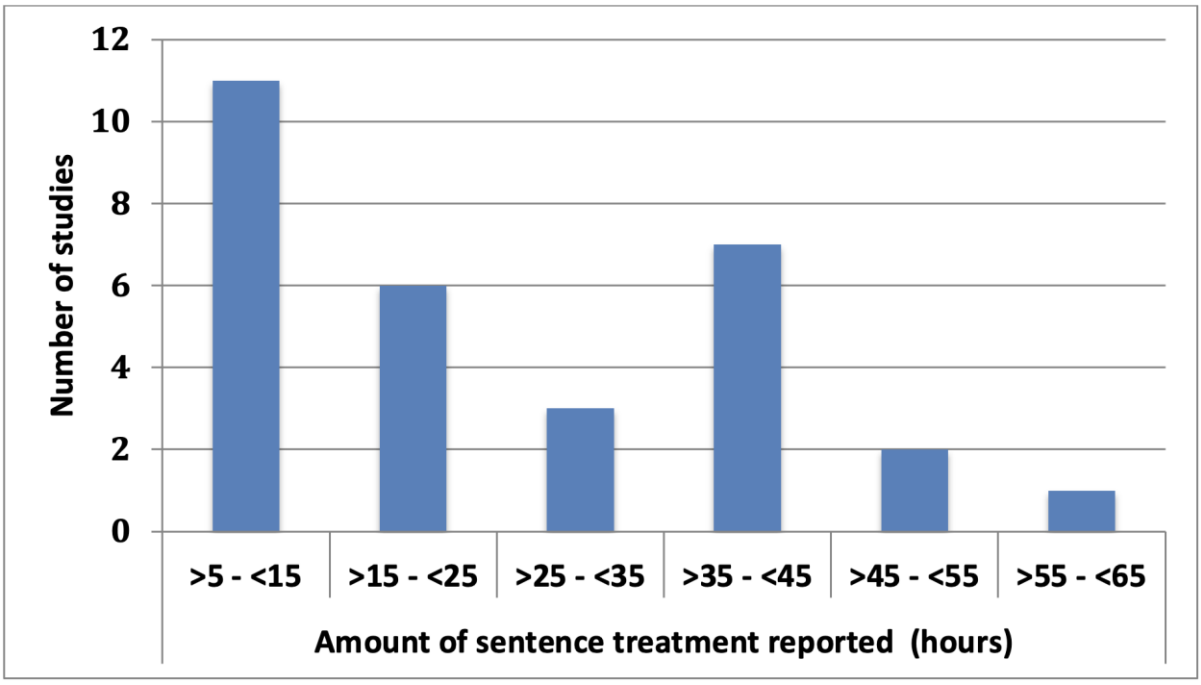
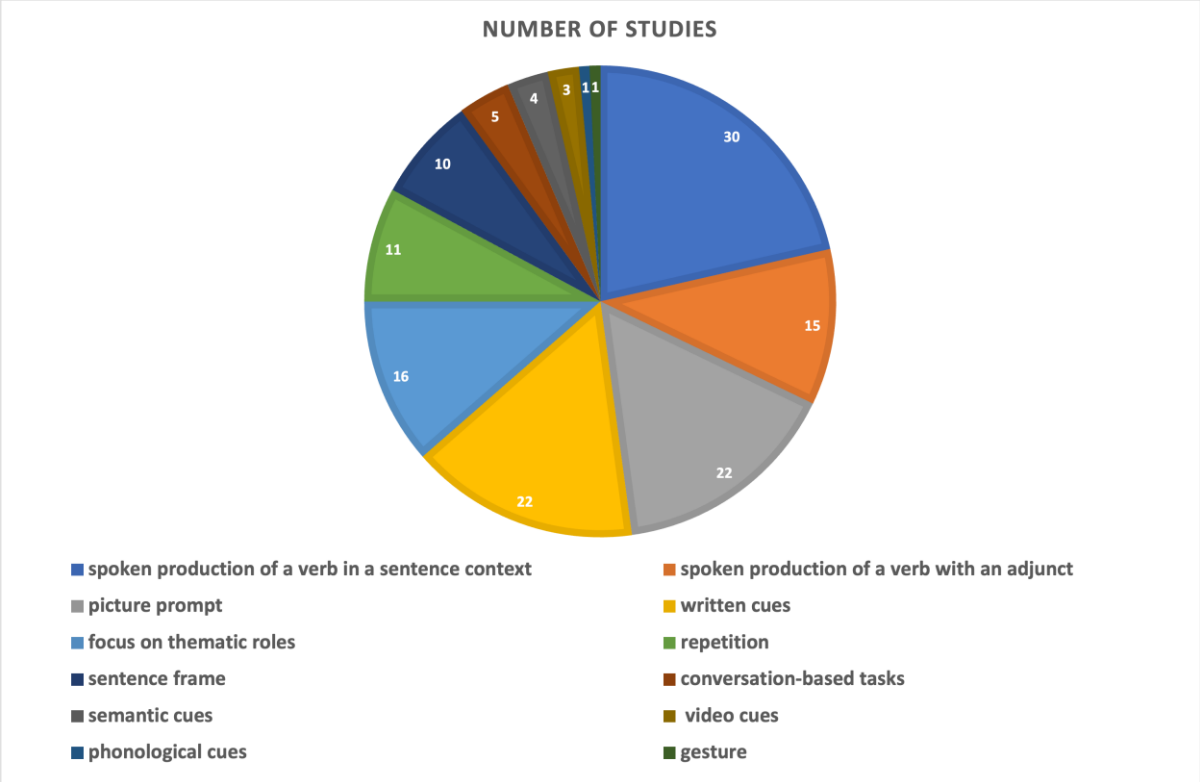


Figure 3.


NUMBER OF STUDIES
30
15
22
22
16
11
10
5
4
3
1
1
spoken production of a verb in a sentence context
spoken production of a verb with an adjunct
picture prompt
written cues
focus on thematic roles
repetition
sentence frame
conversation-based tasks
semantic cues
video cues
phonological cues
gesture

1082 **Preferred Reporting Items for Systematic reviews and Meta-Analyses**
1083 **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			

SECTION	ITEM	PRISMA-ScR CHECKLIST ITEM	REPORTED ON PAGE #
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
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;169:467–473. doi: 10.7326/M18-0850.