

RESEARCH REPORT

Training flexibility in fixed expressions in non-fluent aphasia: A case series report

Claudia Bruns | Suzanne Beeke | Vitor C. Zimmerer | Carolyn Bruce | Rosemary A. Varley

Department of Language & Cognition,
Division of Psychology and Language
Sciences, University College London,
London, UK

Correspondence

Claudia Bruns, Department of Language
& Cognition, Division of Psychology and
Language Sciences, University College
London, Chandler House, 2 Wakefield
Street, London WC1N 1PF, UK.
Email: c.bruns@ucl.ac.uk

Funding information

This research formed part of the first
author's PhD, supported by an Economic
and Social Research Council (ESRC) doc-
toral training grant (award reference num-
ber 1475581).

Abstract

Background: Many speakers with non-fluent aphasia (NFA) are able to produce some well-formed word combinations such as *'I like it'* or *'I don't know'*, although they may not use variations such as *'He likes it'* or *'I don't know that person'*. This suggests that these utterances represent fixed forms.

Aims: This case series investigation explored the impact of a novel intervention aimed at enhancing the connected speech of individuals with NFA. The intervention, motivated by usage-based principles, involved filling open slots in semi-fixed sentence frames.

Methods & Procedures: Five participants with NFA completed a 6-week intervention programme. The intervention trained participants to insert a range of different lexical items into the open slots of high-frequency phrases such as *'I like it'* to enable more productive sentences (e.g., *'they like flowers'*). The outcomes and acceptability were examined: The primary outcome measure focused on changes in connected narrative, and the availability of trained constructions (e.g., *'I like it'*) was explored through a story completion test. Two baseline measures of behaviour were taken prior to intervention, and outcomes assessed immediately after intervention and at a 6-week maintenance assessment.

Outcome & Results: A pre-/post-treatment comparison of connected speech measures showed evidence of enhanced connected speech for two of the five participants (P2 and P5). An analysis of story completion test scores revealed positive change for two participants (P1 and P2). Findings were mixed with regard to baseline stability of outcome measures and post-intervention stability of language changes. The intervention was acceptable to all participants.

Conclusion & Implications: While this pilot study yielded promising findings with regard to the intervention's acceptability and increased connected speech for some participants, the findings were mixed across the sample of five

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *International Journal of Language & Communication Disorders* published by John Wiley & Sons Ltd on behalf of Royal College of Speech and Language Therapists

participants. This research helps inform hypotheses and selection criteria for future studies.

KEYWORDS

case series, formulaic expressions, non-fluent aphasia, sentence processing, therapy, intervention

What this paper adds

What is already known on the subject

- Despite difficulties producing grammatically correct sentences, many speakers with aphasia are able to produce well-formed utterances, often representing familiar expressions such as ‘*I don’t know*’ and ‘*I like it*’. In usage-based Construction Grammar (CxG) theories, familiar utterances are assumed to be processed as one unit and are therefore more resilient to brain damage. CxG assumes that residual utterances such as ‘*I like it*’ map onto more abstract sentence frames (e.g., ‘[REFERENT] *like*-TENSE [THING]’).

What this paper adds to existing knowledge

- Sentence therapy, informed by CxG principles, is novel in aphasiology, and usage-based interventions need to be evaluated with regard to their impact on language processing at the connected speech level. This case series report explores the acceptability and outcomes of a usage-based sentence therapy. We also introduce and explore the value of an automated, frequency-based analysis tool for evaluating connected speech outcomes in aphasia therapy.

What are the potential or actual clinical implications of this work?

- The findings inform further development of usage-based aphasia interventions targeting word combinations.

INTRODUCTION

Despite reduced ability to build grammatically correct sentences, many speakers with non-fluent aphasia (NFA) are able to produce some well-formed word combinations. Many of these residual combinations represent familiar, high-frequency phrases (Bruns et al., 2019; Van Lancker Sidsis, 2012; Zimmerer et al., 2018). Investigating the use of the familiar expression ‘*I don’t know*’, Bruns et al. (2019) found that individuals with Broca’s aphasia employed the phrase for a variety of conversational functions, mostly with a turn-constructive use. However, it was predominantly used with the same linguistic form. The present study investigates whether residual, familiar expressions can be used as a starting point to stimulate greater creative capacity in speakers with NFA, such as from ‘*I don’t know*’ to ‘*I don’t know* [PERSON/LOCATION/SOMETHING]’.

Familiar expressions, often referred to as formulaic expressions, are at the heart of usage-based grammar theories, where linguistic form such as ‘*I don’t know*’ is paired with semantic–pragmatic meaning (e.g., expressing lack of knowledge). This is referred to as a form–meaning pairing or construction. Constructions can be of any length and degree of abstractness. For instance, the familiar phrase ‘*I like it*’ may represent a stored exemplar of a more abstract schema ‘[REFERENT] *like*-TENSE [THING]’. This means that storage is not limited to single words, but whole sentences, parts of sentences and schemas (e.g., “frames with slots”; Dąbrowska, 2014, p. 619) can also be stored as constructions. In usage-based theories such as Construction Grammar (CxG; Goldberg, 2003), lexicon and syntax are not strictly separated. If a construction is encountered frequently, it is more likely to be represented in the speaker’s ‘construct-i-con’ (Goldberg, 2003). This makes CxG a

fruitful framework to inform aphasia therapy as it allows targeting words, word combinations or sentences, depending on a speaker's resources.

Some elements of existing speech and language therapy interventions for NFA are in accordance with usage-based assumptions. For instance, familiar expressions such as *'thank you'* are targeted in Melodic Intonation Therapy (Helm-Estabrooks et al., 1989) and are also employed in Intensive Language-Action Therapy (Stahl et al., 2016); semi-fixed frames with open slots (e.g., *'I'd like X'*) are practised in script training (Bilda, 2011); and Reduced Syntax Therapy (Ruiter et al., 2010; Schlenck et al., 1995; Springer et al., 2000) trains abstract constructional schemas (e.g., the elliptical [DOING/DONE] + [WHAT] construction).

Such sentence-level therapies for aphasia often show improvement of trained structures, but limited generalization to spoken discourse. For example, a summary of 13 script training studies by Kaye and Cherney (2016) concluded that while scripts were successfully acquired in all studies, some of which found maintenance effects, there was limited generalization to untrained scripts or interactions with different communication partners. Similarly, Murray et al. (2007) evaluated Treatment of Underlying Forms (TUF; Thompson & Shapiro, 2005) within a case study and found generalization of treated sentences to connected speech, but limited (cross-modal) generalization to untreated structures.

As yet no intervention programme has explicitly tested application to aphasia of usage-based principles. For example, constructional grounding (Israel et al., 2001; Riches, 2013) refers to a lexically specific and high-frequency construction (e.g., *'I like it'*) which may act as a source construction, representing an instance of a more abstract constructional schema (e.g., [REFERENT] *like*-TENSE [THING]). Superimposition (Dąbrowska, 2014) is a process where lexical items or chunks are inserted into an open slot of a semi-fixed construction (e.g., *'all of us like cake'*). Employing such principles in the design of sentence therapies might help speakers with aphasia produce longer and more flexible utterances.

The present study tested a novel behavioural intervention for NFA aimed at increasing the productivity of multiword expressions. Introducing flexibility into source constructions was approached by aiming to loosen open slots in semi-fixed constructions such as [REFERENT] *like*-TENSE [THING], starting with a high-frequency phrase (e.g., *'I like it'*) and then superimposing lexical items or chunks (e.g., *'You like it' → 'You like watching football'*).

The intervention incorporated psycholinguistic and neuroscientific learning principles such as structural priming (Kaschak et al., 2014) and errorless learning (e.g., Conroy & Scowcroft, 2012). Structural priming is a phe-

nomenon where a speaker is more likely to reuse a structure they have recently encountered (Kaschak et al., 2014). In the structural priming paradigm, a participant is primed with a specific syntactic structure or a schema (e.g., *'he gave me a book'*), for instance by reading aloud. The second part of a structural priming task typically uses a picture description task where the participant, despite no explicit instruction, tends to describe the event employing the structure with which they were primed (e.g., *'she gave him an apple'* as opposed to *'she gave an apple to him'*). Lee and Man (2017) used implicit structural priming training with an individual with Broca's aphasia and found improved sentence production of prepositional dative structures (e.g., *'the boy is giving a guitar to the singer'*). Errorless learning methods employed in aphasia studies include decreasing cues (Conroy & Scowcroft, 2012) or, as in a 'pure' errorless learning method, direct repetition of a target structure. Similar principles were applied in the 'Sheffield WORD—Structured speech therapy' program (SWORD; Varley et al., 2016; Whiteside et al., 2012), which was designed for individuals with acquired apraxia of speech (AoS). Here, fluent and flexible use of constructions (words in isolation and embedded in phrases such as *'Where is ___?'*) is practised in a stepwise way, using error-reduction strategies. In the first step, the participant with AoS is presented with a video clip of a neurotypical speaker producing a target word. This is followed by imagined production of the word. In a third step, the participant attempts production (repetition) of the target. This procedure seeks to facilitate fluent language production by minimizing errors. A description of how these two principles (structural priming and error-reduction principles) were incorporated into the current intervention can be found in the Methods section. Moreover, the current study incorporated social-motivational learning elements (Varley, 2011). Constructions were practised by making use of a participant's own voice: Target phrases were recorded and modified to enhance fluency (which we will refer to as 'self-voice'). Listening to one's own voice producing a more fluent target sentence may be particularly motivating for people with NFA. Moreover, self-voice, which brain imaging data suggests is processed in the right hemisphere (Kaplan et al., 2008), has been found to positively affect learning and remembering in university students (Forrin & MacLeod, 2017) and may have the same benefits for other populations.

Finally, the intervention was computerized to allow opportunities for self-managed home practice. This may increase the frequency with which participants engage with intervention activities (Nobis-Bosch et al., 2011). High-intensity aphasia therapy (8.8 h per week for 11 weeks) has been found to be more effective compared with lower intensity therapy, administered over a longer

time span (Bhogal et al., 2003). The principle of increasing therapeutic dose (total number, length and frequency of sessions) through home practice has become a common component of aphasia interventions. Varley et al. (2016), for instance, found a positive relationship between dose (the time spent on a self-administered computerized intervention) and treatment outcome (correctly named words post-intervention) in a large group of participants with AoS.

Change in connected speech production was explored through use of the automated *Frequency in Language Analysis Tool* (FLAT; Zimmerer et al., 2016, 2018). This is the first use and exploration of the value of this automated software as an outcome measure in intervention research. It has particular advantages in that it characterizes features of connected speech in a fast and blind fashion. We examined the value of combination ratio as an outcome measure (number of trigrams (three-word combinations) produced by a speaker divided by the number of words).

The current study set out to evaluate the outcomes of a case series, to test the potential of this computerised, usage-based intervention for improving sentence processing abilities in NFA. In addition, participants with NFA and their regular conversation partners (CPs) were asked to comment on the acceptability of the intervention. The following main research questions guided the analysis:

1. Is there evidence that after intervention participants with NFA demonstrate:
 - a. enhanced connected speech, as measured by a higher proportion of multiword utterances (i.e., combination ratio) in narratives?
 - b. increased use of trained constructions, as measured by performance in a story completion test?
 - c. better spoken sentence comprehension, as measured by performance in the Test for Reception of Grammar—Version 2 (TROG-2; Bishop, 2003)?
 - d. change in the perceived impact of aphasia, as measured by participant ratings on the Aphasia Impact Questionnaire 21 (AIQ-21; Swinburn et al., 2018)?
2. How acceptable is the intervention to participants with NFA and their CPs?

METHODS

Participants

Following ethical approval from the UCL Language & Cognition Departmental Ethics Committee (Project ID: LCRD.2017.01), five participants and their regular CPs were recruited via a university research register. Written

consent to participate in this study was obtained both from participants with aphasia and their CPs.¹

Three participants were female and two were male. They presented with chronic, post-stroke NFA (average of 91 months post-stroke onset, range = 24–165) as identified by experienced SLTs, had no known neurodegenerative illness, used English as their main language and — based on self-report — had sufficient auditory and visual acuity to interact with a laptop. The mean age was 60 years (range = 48–68). They did not receive other speech and language therapy interventions directed at impairment level during their involvement in the study, although they might continue participation in social support groups. Participants' language output was characterized by grammatically impoverished, non-fluent, and effortful spontaneous speech and narrative production. Table 1 presents an overview of participant demographic characteristics and performance on background/profiling assessments. All participants showed relatively preserved auditory lexical comprehension (range = 24–28 out of 30), as determined by the spoken word comprehension subtest of the CAT (Swinburn et al., 2004). However, the group was heterogeneous with regard to object naming, digit span and non-verbal IQ performance, as measured by the BNT (Kaplan et al., 2001), subtest 13 from the PALPA battery (Kay et al., 1992) and the WASI-II Matrix Reasoning subtest (Wechsler, 2011), respectively. The background profiles revealed variability with respect to aphasia severity: P4 presented with the most severe aphasia, followed by P3 and P2 (Table 1). P1's and P5's background profiles indicated less severe aphasia compared with the rest of the group, although P1's object naming was considerably more affected by her aphasia than P5's naming performance.

Design and intervention

Employing a case series design, each participant was involved in a 16-week study, consisting of three phases: baseline (weeks 1–3), intervention (weeks 4–9) and post-intervention probes (weeks 10–16). The intervention was delivered by a researcher with a speech and language therapy qualification (first author). During the 6-week intervention phase, five individual 60-min sessions alternated with five 60-min home visits. Home visits facilitated home practice and ensured that participants could complete self-managed activities on a laptop supplied by the research team. A description of the intervention using the

¹ Participants videotaped weekly 10–20-min conversation samples with their CP during the baseline and post-intervention phases. These data are not presented here.

TABLE 1 Demographic characteristics and background assessments of participants with NFA

Participant ID	Age (years)	Gender	Months post-onset	Highest level of education	Naming objects (BNT) ^a (maximum 60)	Digit span by recognition (PALPA-13) ^b (maximum 7)	Spoken word comprehension (CAT subtest) ^c (maximum 30)	Non-verbal IQ (WASI-II MR) ^d (maximum 30)
P1	48	Female	46	Secondary school	30	6	28	17
P2	63	Male	84	Secondary school	25	3	28	13
P3	60	Male	165	University	10	4	24	12
P4	68	Female	24	College	7	3	24	10
P5	60	Female	137	Postgraduate studies	55	6	28	15
Average	59.8	-	91.2	-	25.4	4.4	26.4	13.4
SD	7.4	-	59.5	-	19.2	1.5	2.2	2.7

Notes: ^aBoston Naming Test (BNT; Kaplan et al., 2001).

^bSubtest 13 from the Psycholinguistic Assessments of Language Processing in Aphasia (PALPA; Kay et al., 1992).

^cComprehension of spoken words subtest from the Comprehensive Aphasia Test (CAT; Swinburn et al., 2004).

^dMatrix Reasoning Subscale from the Wechsler Abbreviated Scale of Intelligence—Second Edition (Wechsler, 2011).

TIDieR checklist (Hoffmann et al., 2014) is presented in Appendix A.

The main aim of the intervention was to train flexible use of 12 constructions. Each represented a familiar expression with high usage frequency (a source construction such as *'I like it'*) which mapped onto a semi-fixed frame with open slots (e.g., [REFERENT] *like* [THING/PERSON/LOCATION/PROCESS], as in *'you like cake'*). The high-frequency source constructions were derived from the spoken British National Corpus (BNC, 2007). The constructions corresponded to overarching interactional functions (e.g., giving an opinion, asking a question) relevant to everyday talk about experiences, opinions and exchanging information. To reduce lexical-semantic demands, open slots were mostly filled by pro-forms (e.g., *'that'* or pronouns such as *'I'*)—words of high frequency and versatile as to the conversational contexts in which they can be used.

The same high-frequency constructions were employed for all participants, although when practising inserting new lexical material into the open slots, personalized items could be used (e.g., *'I like CDs/Elton John'*). The intervention was divided into three phases (Figure 1), and the following underlying techniques and principles were applied: self-voice, structural priming, superimposition, error-reduction principles and increasing dose through home practice. Phase 1 elicited fluent as possible versions of a participant's own productions of the target phrases (self-voice). Each participant recorded the same set of 59 phrases (names were individualized, e.g., *'[NAME] had dinner'*) supported by the researcher (e.g., by direct repetition or reading aloud). If parts of a recording were effortful and non-fluent, the researcher modified the phrase using the sound-editing software Audacity (version 2.1.2; <https://www.audacityteam.org/>). Modifications through waveform editing included deleting fillers (e.g., *'um'* and *'er'*) and reducing long vowel durations or long pauses between words, similar to the procedures described by Harmon et al. (2016). In Phase 2 (structural priming), participants were exposed to variations of the target constructions within a gamified reaction time task using an auditory word monitoring paradigm. In this word monitoring game (WMG), the participant reacts as quickly as possible via button press to a prespecified target word (e.g., *'like'*) once it is encountered in an auditory sentential context (*'I like it when spring is coming'*). The WMG was self-managed (i.e., practised at home) and the program automatically recorded when and for how long the WMG was practised. The aim was to activate constructional schemas by repeated exposure, which may prime and thus facilitate production of these constructions in Phase 3. Phase 3 involved production training. The self-voice sound files from Phase 1 were incorporated into a program (using

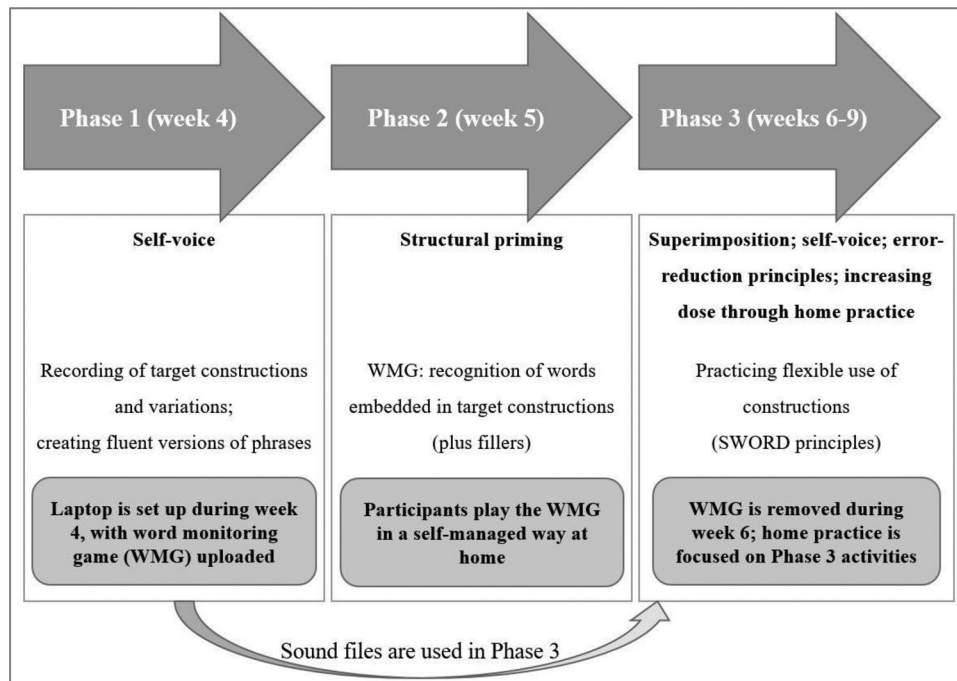


FIGURE 1 Three phases of the intervention

Microsoft PowerPoint) designed to practise more flexible use of constructions.

The main underlying usage-based principle in Phase 3 was superimposition (Dąbrowska, 2014), where lexical items (e.g., 'Claire') were superimposed over an open slot (e.g., 'Where is ___?' → 'Where is Claire?'). Each target sentence involved three steps applying SWORD (White-side et al., 2012) error-reduction strategies: (1) seeing the written form of the target phrase while listening to an audio-recording of it; (2) imagining saying the sentence; and (3) saying the sentence and making a recording of own attempt. During the 'listening' step, variations of a core construction were animated. For instance, in the 'you like it'-'listen'-slide, the animation started with the text box '_____ like it', accompanied by a sound file where the open slot was filled by a beep ('*beep* like it', created with Audacity). Next, the pronoun 'you' option was flown into the open slot, accompanied by the sound file of the whole phrase 'you like it'. Open slots in source constructions were often filled with pro-forms (e.g., pronouns and deictic expressions such as 'that') as well as high-frequency verbs and nouns (e.g., 'dinner', 'swim') to reduce lexical demands.

Outcome measures

Table 2 shows the assessment battery. All assessments were conducted by the first author who was not masked to phase

(i.e., baseline versus outcome). Personal narratives (e.g., 'Can you tell me about the last time you went on holiday, or the last trip you took?') and picture-based narratives ('Dinner Party' and 'Jogging' cartoon series; retelling eight black-and-white line drawings that together depicted a complex event; Fletcher & Birt, 1983) were audio-recorded and served as primary outcome measures to assess participants' language production at the multiword level. These were administered at two baselines (approximately 2 weeks apart), once after the intervention and after a 6-week no treatment period to assess the stability of any communicative change.

Outcome variables on narratives were frequency-based measures from the FLAT (version 2; Zimmerer & Wibrow, 2015; Zimmerer et al., 2018). The main outcome variable was combination ratio, which quantifies the amount of connected speech by dividing the number of well-formed trigrams by number of words. For instance, the sentence 'A man phones his friend' includes five words and three trigrams (*a man phones; man phones his; phones his friend*; combination ratio = 0.6). A higher combination ratio indicates a higher proportion of trigrams in a speaker's connected speech. Since several samples were collected at each time point (Table 2), average combination ratio per assessment point were calculated across each participant's samples.

Secondary outcome measures were an adaptation of the Story Completion Test (Goodglass et al., 1972), the TROG-2 (Bishop, 2003) and the AIQ-21 (Swinburn et al., 2018).

TABLE 2 Overview of outcome and control measures

Week	Baseline			Intervention 4–9	Post-intervention probes						Variable(s) of interest		
	1	2	3		10	11	12	13–15	16				
<i>Primary outcome measure: Quantity of connected speech</i>													
Dinner party ^a	✓				✓							✓	Combination ratio: Number of well-formed trigrams divided by the number of words
Jogging ^a			✓		✓							✓	Normative values, taken from Zimmerer et al. (in prep.) are 0.80 (SD = 0.06) for 'Dinner Party' narrative and 0.73 (SD = 0.05) for personal narrative 'Last Holiday'
Personal narratives ^b	✓		✓		✓							✓	
<i>Secondary outcome measures</i>													
Story completion test ^c	✓		✓		✓							✓	Expected answers
TROG-2 ^d				✓								✓	Number of well-formed utterances
AIQ-2I ^e			✓									✓	Number of blocks/items correct
Conversation videos	✓		✓		✓							✓	Participant ratings
												✓	Not presented here
<i>Control measure</i>													
Written synonym matching task ^f	✓		✓		✓							✓	Number of items correct

Notes: ^aFletcher and Birt (1983).

^bQuestions to elicit monologues; week 1: Last holiday; week 3: The most frightening experience you've had; week 10: The most interesting person you've met; and week 16: The best film you've seen (or most recent one).

^cAdapted from Goodglass et al. (1972).

^dTest for Reception of Grammar—Version 2 (TROG-2; Bishop, 2003).

^eAphasia Impact Questionnaire (AIQ-2I; Swinburn et al., 2018).

^fShortened form of a written Synonym Matching Task (40% low-imageability and 40% high-imageability versus 10% low-frequency word pairs), taken from the Action for Dysphasic Adults Auditory Comprehension Battery (ADA; Franklin et al., 1992).

The Story Completion Test was originally developed by Goodglass et al. (1972) to study the availability of abstract syntactic structures in NFA. It was adapted to probe a participant's ability to produce the target constructions and was administered twice at baseline and immediately post-treatment and at a 6-week maintenance assessment. As in Goodglass et al. (1972), the researcher presented a brief scenario where the final sentence or phrase was missing, for example, to elicit the trained 'I like it' construction: 'I'm fond of ball games. My friend asks me: "What are your thoughts on football?", so I say...?'. Participants were asked to complete each of the 12 stories with a single sentence with the aim of eliciting the trained source construction. Responses were scored by the research team (unmasked) by applying two criteria: An 'expected answer' and a 'well-formed utterance'. As an example, for the item where the expected construction was 'I like it', the answer 'it's fun' would be scored with a 0 in the 'expected answer' category, while a 1 would be noted in the 'well-formed utterance' category.

Although story situations were designed to elicit trained constructions within predictable sentential contexts, a norming sample of 41 native speakers of English revealed that the probabilities with which the target constructions were produced varied considerably across items (range = 7–100%).² The 'expected answer' scores were therefore weighted according to the normative cloze probability (i.e., the proportion of participants that finished a probe sentence with the expected response). For instance, if an item elicited a normative cloze probability of 39% (i.e., 39% of neurotypical speakers answered with the target construction), the score that a participant with NFA would receive for producing the expected construction for that item was 39, while a normative cloze probability of 100% would mean that a score of 100 would be assigned if the participant answered with the target construction. The score was given for the target construction or a construction with the same grammatical structure, disregarding semantics (e.g., 'I like it'/'I enjoy football'/'I hate football'), and a score of zero was assigned if the answer represented another structure (e.g., 'it's fun' instead of 'I like it'). The maximum score was 708 across the 12 items.

Similarly to Goodglass et al. (1972), in addition to the 'expected answer' score, the number of well-formed utterances was recorded in the following way. For every answer reflecting a well-formed utterance (disregarding lexical semantics), a score of 1 was given. Ellipsis (a one-word response) for target 'I like it' was allowed (e.g., 'fantastic'). Formulas such as 'oh dear' and 'come on' were not scored

as well-formed utterances. Other formulas ('I don't know', '(I'm) sorry', 'no way') however, counted as a well-formed utterance for certain items (for more details, see Bruns, 2018).

The TROG-2 (Bishop, 2003), a receptive language test to identify the areas of grammatical comprehension difficulty, was administered once before and after the intervention, to investigate whether there was change in spoken sentence comprehension (number of blocks correct, out of 20). We selected the TROG-2 as the phase 2 and 3 intervention components involved listening to (and reading) sentences, and as it probes a wide range of sentence structures and is less subject to ceiling effects than screening assessments of sentence comprehension. Additionally, the AIQ-21 (Swinburn et al., 2018) was used as a participant-reported measure of communication abilities, emotional state/well-being, and participation. It was administered at baseline 2 and immediately after the intervention. The AIQ-21 uses a Likert-type scale from 0 (most positive rating) to 4 (most negative rating) for all questions.

Finally, a shortened form of a written Synonym Matching Task, taken from the ADA (Franklin et al., 1992), was created as a control measure. This was administered twice at baseline, post-intervention and at maintenance. It was devised to assess written word semantic processing which was not directly targeted in the intervention. Performance was measured according to number of items correct (out of 40).

In addition, participants videotaped weekly 10–20-min conversation samples with a regular CP during the baseline and post-intervention phases. These conversation samples are not analysed as part of this report.

Acceptability of the intervention

Acceptability of the intervention to both participants with NFA and their CPs was investigated through post-intervention study-specific questionnaires (see Appendix B). There were separate questionnaires for participants with NFA and CPs. These were designed to capture the views of participants regarding different aspects of the intervention (e.g., which elements were helpful/unhelpful; the overall usefulness of the intervention). Open questions allowed both participants with NFA and their CPs to add further comments. Participants filled in the questionnaires in their own time without the researcher present, and participants with NFA were encouraged to fill in the questionnaire together with their CP. The questionnaires were returned anonymously in pre-stamped addressed envelopes directly to the second author.

² This norming sample consisted of 41 university students (average age = 24.3, SD = 2.5, 38 female).

Data analysis

Audio-recorded connected speech samples (personal and picture-based narratives) were orthographically transcribed by a research assistant (MSc student of Linguistics) blinded as to the sample collection point. Each participant's pre-, post-intervention and maintenance data were analysed by the first author using the FLAT version 2 (Zimmerer & Wibrow, 2015; Zimmerer et al., 2018). Before performing the frequency-based analysis, the first author checked all transcriptions for accuracy and annotated them in line with FLAT transcription conventions. For instance, clause boundaries (e.g., in the case of false starts or ungrammatical word combinations) were marked with separators ('<.>') to ensure that only grammatically well-formed multiword utterances were included in the FLAT analysis (Bruns et al., 2019; Zimmerer et al., 2018). For example, the utterance 'upstairs, him is, erm, tie and everything' is separated into 'upstairs <.> him is (erm) <.> tie and everything'.

Since the group represented a relatively small, heterogeneous sample of individuals with NFA, all results were evaluated at the individual rather than the group level. Group means were plotted for the main outcome variables and the control measure, to identify patterns across individuals. Individual recurrent phrases (e.g., 'oh dear me', frequently produced by P4) were included in analysis of spontaneous speech samples as each participant acted as their own control. For the primary outcome measure, the mean combination ratio was compared within participants across the four assessment points (baseline 1, baseline 2, post-intervention and maintenance) using descriptive statistics. The mean combination ratio at each time point was calculated by averaging values for each narrative task (e.g., time point 'baseline 1' consisted of the average of the two mean combination ratios of the 'Last Holiday' and 'Dinner Party' narratives). In terms of the effectiveness of the intervention, we were interested in the difference between baseline 2 and outcome 1. This difference was compared against the pre-therapy stability (difference between baseline 1 and 2). Stability of post-intervention and maintenance performance was also assessed.

Performance on secondary outcome measures was analysed by comparing pre- and post-intervention scores for each participant. For the adapted story completion test, McNemar tests for related samples were conducted individually to investigate whether baseline performance was stable (baseline 1 versus 2). Where baseline performance was found to be stable, Cochran's Q tests were used to examine each participant's performance prior to and following the intervention (baseline 2 versus post-intervention versus maintenance). For these comparisons,

baseline 2 measurement was used as the pre-intervention measure as participants were more familiar with the tasks and the researcher compared with baseline 1. The same analysis was applied to the control measure.

RESULTS

All five participants completed the intervention study with an intervention phase of 6–7 weeks (P5 had to cancel two therapy sessions which prolonged their intervention phase by 1 week). Results are reported measure by measure, evaluating individual change patterns after intervention.

Narratives

For the primary outcome variable, combination ratio, it was expected that the intervention would lead to increased connected speech, that is, higher values at post-intervention and maintenance probes. Figure 2 displays participants' scores over time as well as a group mean for each assessment point. As shown, all five participants presented with combination ratios below normative values (Table 2), indicating less connected speech compared with neurotypical speakers (i.e., lower proportion of multiword utterances).

While within-participant comparisons indicate an increase in combination ratios between baseline 2 and the first post-intervention probe for four participants (P2–P5), these differences did not exceed the baseline variation in three of these participants (P2–P4). For P5, however, the increase in values between baseline 2 and post-intervention was higher than the baseline variation and remained stable across post-intervention and maintenance probes. P1's combination ratio was 0.46 and 0.51 at baseline 1 and 2, respectively, dropping back to 0.46 after intervention, with a value of 0.48 at maintenance. Thus, over time, P1's amount of connected speech appeared relatively stable.

Story completion test

Figure 3 shows 'expected answers' scores in the four assessment probes. As noted for the narratives, participants' raw scores at baseline varied considerably, and only P5's performance after the intervention showed a marked increase compared with pre-intervention scores.

To statistically investigate change, raw scores in the 'expected answer' category were converted into a binary scale (1 = expected construction, 0 = unexpected construction) to perform McNemar and Cochran's Q tests.

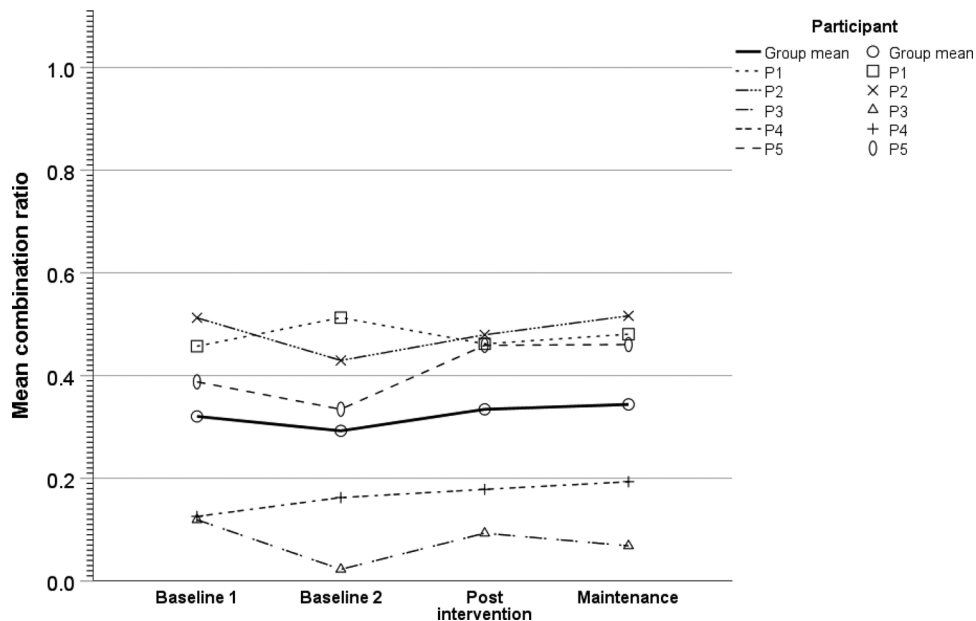


FIGURE 2 Mean combination ratio by participant over time

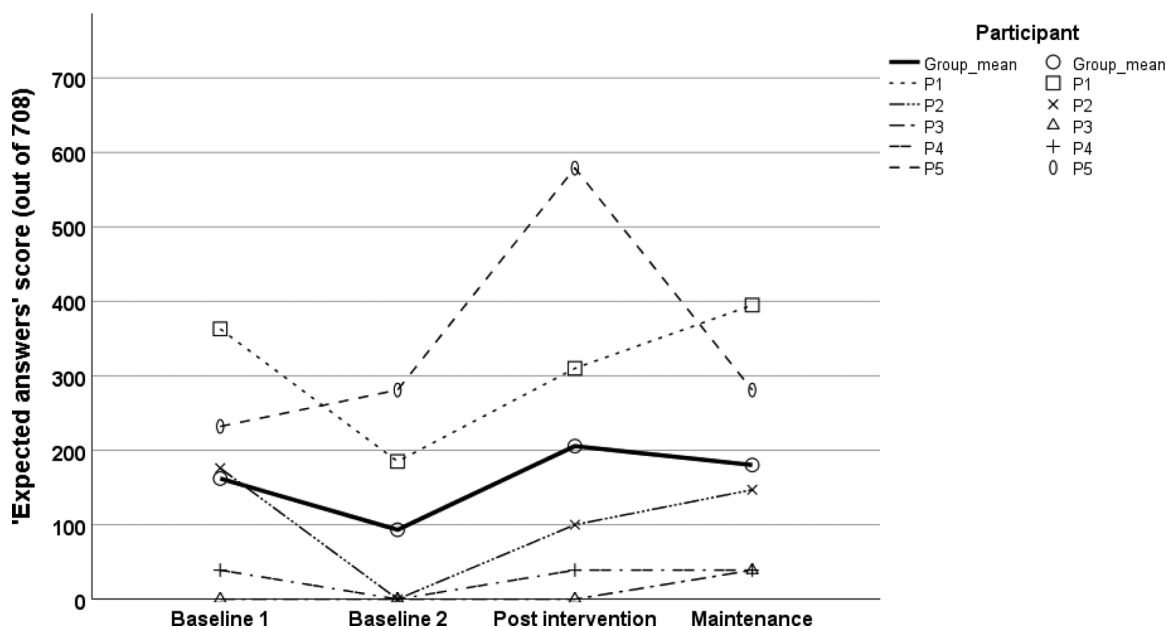


FIGURE 3 Story completion test: Overall score based on expected answers by participant over time

McNemar tests revealed no significant difference between baselines 1 and 2 for P1, P2, P4 and P5 (P1: $p = 0.25$; P2: $p = 0.50$; P3: stable 0 baseline; P4: $p = 1.00$; P5: $p = 1.00$). There was a significant difference between P5's baseline 2 versus post-intervention and maintenance scores, as indicated by Cochran's Q test, $Q = 8.0, p = 0.018$. This result points to an increase in overall score immediately after intervention which dropped back to baseline level at maintenance, as shown in Figure 3. For the remaining participants, there were no significant differences between baseline 2 versus

post-intervention and maintenance scores (P1: $Q = 3.50, n.s., p = 0.174$; P2: $Q = 2.00, n.s., p = 0.368$; P3: $Q = 2.00, n.s., p = 0.368$; P4: $Q = 2.00, n.s., p = 0.368$). P3 was included here since both baselines resulted in a score of zero, which was regarded as stable.

In terms of well-formed utterances, participants' patterns over time are shown in Figure 4. P1's and P2's proportions of well-formed utterances increased significantly after the intervention. P3 was unable to retrieve any constructions at baseline ('*dunno*' answers for all items,

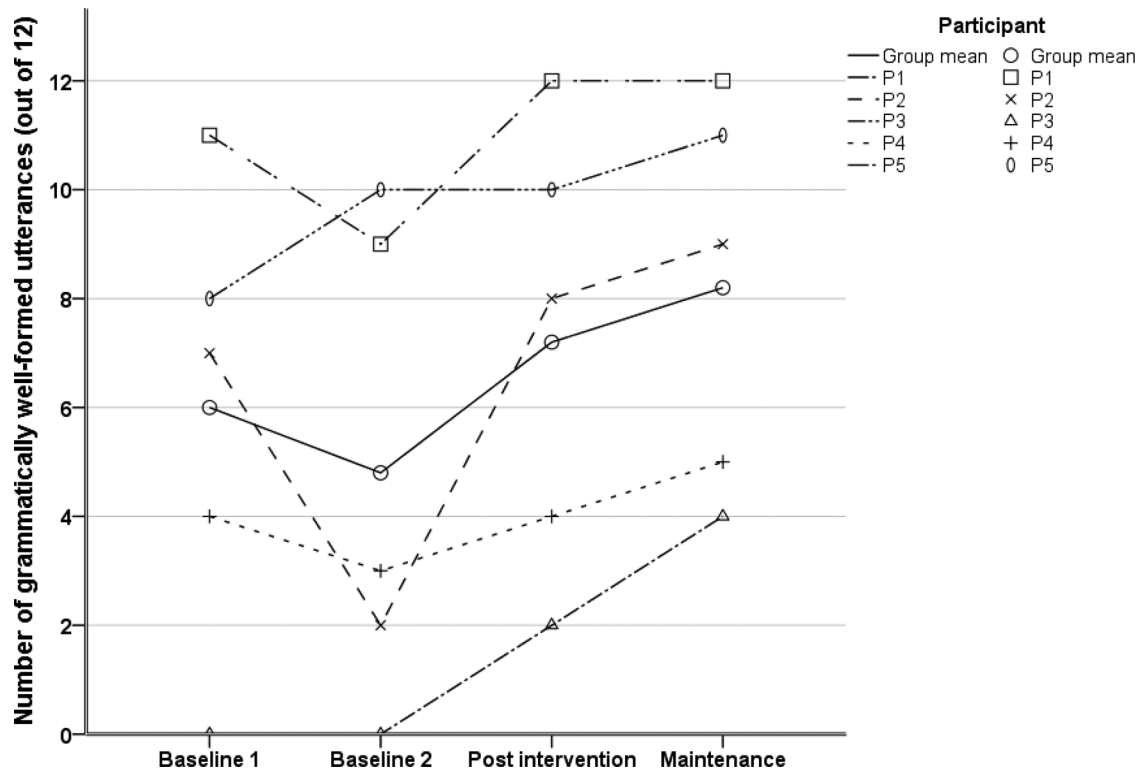


FIGURE 4 Story completion test: Number of well-formed utterances by participant over time

resulting in scores of 0 at baseline 1 and 2), but did so following intervention. Although P3 was able to give differentiated responses after the intervention and produced more well-formed utterances across post-intervention and maintenance probes, this difference was not statistically significant.

Again, stability at baseline for P1, P2, P4 and P5 was examined using McNemar tests, where non-significant results were found for P1 ($p = 0.625$), P4 ($p = 1.00$) and P5 ($p = 0.500$) and a trend toward significance for P2 ($p = 0.063$). This indicates that performance at baseline 1 versus baseline 2 was not significantly different for these participants. Note that P3's scores were zero prior to the intervention (explained above). Change in the proportion of well-formed utterances following the intervention was analysed using Cochran's Q tests (where baseline 2 was used as the pre-intervention score). While there was no significant change in performance over time for P3–P5 (P3: $Q = 4.8$, n.s., $p = 0.091$; P4: $Q = 1.2$, n.s., $p = 0.549$; P5: $Q = 0.500$, n.s., $p = 0.779$), P1 and P2's proportion of well-formed utterances changed significantly (P1: $Q = 6.0$, $p = 0.05$; P2: $Q = 9.6$, $p = 0.008$). It should be noted, however, that for both P1 and P2, post-intervention scores exceeded baseline 1 probes by only 1 point. As shown in Figure 4, P1 and P2's increase from baseline 2 to post-intervention remained stable at maintenance, or improved further by 1 point (P2).

TROG-2

A comparison of TROG-2 scores before and after intervention, as shown in Table 3, revealed a relatively stable performance for P2. While P4's and P1's scores decreased by one block, P3's score increased by two blocks and P5's score by three blocks. This tentatively suggests that some participants with NFA (P3 and P5) may show improved spoken sentence comprehension after intervention.

AIQ-21

AIQ-21 ratings (sampled at weeks 3 and 16), presented in Table 4, suggest an increase in perceived well-being and emotional state following intervention. However, with regard to communication, the ratings became more negative for some participants which may point to increased awareness of communication difficulties after intervention.

Four out of five participants showed improved ratings in emotional state/well-being (P1, P3–P5), while there was no change in this area for P2. In the area of participation, no change was detected in three participants (P1, P3, P4). P5, however, showed an improved average rating and P2's average rating decreased. In terms of communication, three participants rated their communicative skills more

TABLE 3 TROG-2 scores before and after intervention

Comprehension of spoken sentences (TROG-2)			
Participant	Pre (week 2)	Post (week 10)	Change pre- versus post-intervention (number of blocks)
	Blocks	Blocks	
P1	16	15	-1
P2	2	2	0
P3	6	8	+2
P4	5	4	-1
P5	12	15	+3

TABLE 4 AIQ-21 ratings before intervention (week 3) and at maintenance (week 16)

		P1		P2		P3		P4		P5	
		Before	After	Before	After	Before	After	Before	After	Before	After
Average ratings	Communication	1.83	1.83	1.83	1.50	2.17	2.33	0.50	0.83	1.67	1.83
	Participation	0.00	0.00	0.75	1.00	1.25	1.25	0.25	0.25	1.50	1.25
	Emotional state/well-being	1.36	1.00	2.36	2.36	1.27	0.73	0.55	0.36	1.64	1.55
Difference score	Communication	0.00		-0.33		0.17		0.33		0.17	
	Participation	0.00		0.25		0.00		0.00		-0.25	
	Emotional state/well-being	-0.36		0.00		-0.55		-0.18		-0.09	

Note: The AIQ-21 uses a Likert-type scale from 0 (most positive rating) to 4 (most negative rating). Difference scores were expected to be zero or negative, e.g., 1.00 - 1.36 = -0.36.

negatively after the intervention (P3–P5), and for P1 there was no change. P2, on the other hand, showed improved ratings.

Control measure: Synonym matching task

The results of the synonym matching task (maximum score = 40), administered across baselines, post-intervention and at maintenance displayed relatively stable performance across participants (Figure 5). Baseline scores were stable (as indicated by McNemar tests, P1: $p = 0.375$; P2: $p = 0.625$; P3: $p = 0.625$; P4: $p = 0.754$; P5: $p = 0.500$). When comparing baseline 2, post-intervention and maintenance using Cochran's Q tests, there was no significant change for any of the participants' scores (P1: $Q = 1.2$, n.s., $p = 0.549$; P2: $Q = 1.0$, n.s., $p = 0.607$; P3: $Q = 2.0$; n.s., $p = 0.368$; P4: $Q = 2.8$, n.s., $p = 0.247$; P5: $Q = 3.5$, n.s., $p = 0.174$). We note that with the exception of P3, most participants were near ceiling level. We attempted to avoid this possibility by biasing word pairs towards low-imageability and low-frequency stimuli. However, the results reveal that the task was insufficiently challenging to act as a control measure.

Acceptability

The overall perceived usefulness of the intervention was rated on a scale from 1 ('not at all helpful') to 5 ('very

helpful'). The resulting average rating from participants with NFA was 4.7, while CPs rated the intervention's overall usefulness for their friend/family member with an average of 4.0.

All five participants found 'one-to-one sessions to practise speaking' particularly helpful, and four participants also selected 'practising speaking at home'; 'making recordings of own voice and listening back to them'; and 'three steps to practise speaking' as helpful elements. Participants' answers regarding the frequency of using target constructions included 'every day' (one participant with NFA) and 'a few times a week' (four participants with NFA). This is suggestive of the constructions' everyday relevance.

For perceived effects of therapy (self-reported changes in communication), one participant with NFA circled 'my speaking got **much** easier' (P2), while the remaining four participants circled 'my speaking got **a bit** easier'. CPs reported that, after intervention, the expressive abilities of their friend or family member 'got a bit easier' (two CPs) or 'no change' (three CPs).

Summary of the results by participant

For P1, combination ratio values did not change in a meaningful way following intervention. P1's outcomes in the adapted story completion test were more positive, with a greater number of well-formed utterances and some evidence for increased availability of target constructions

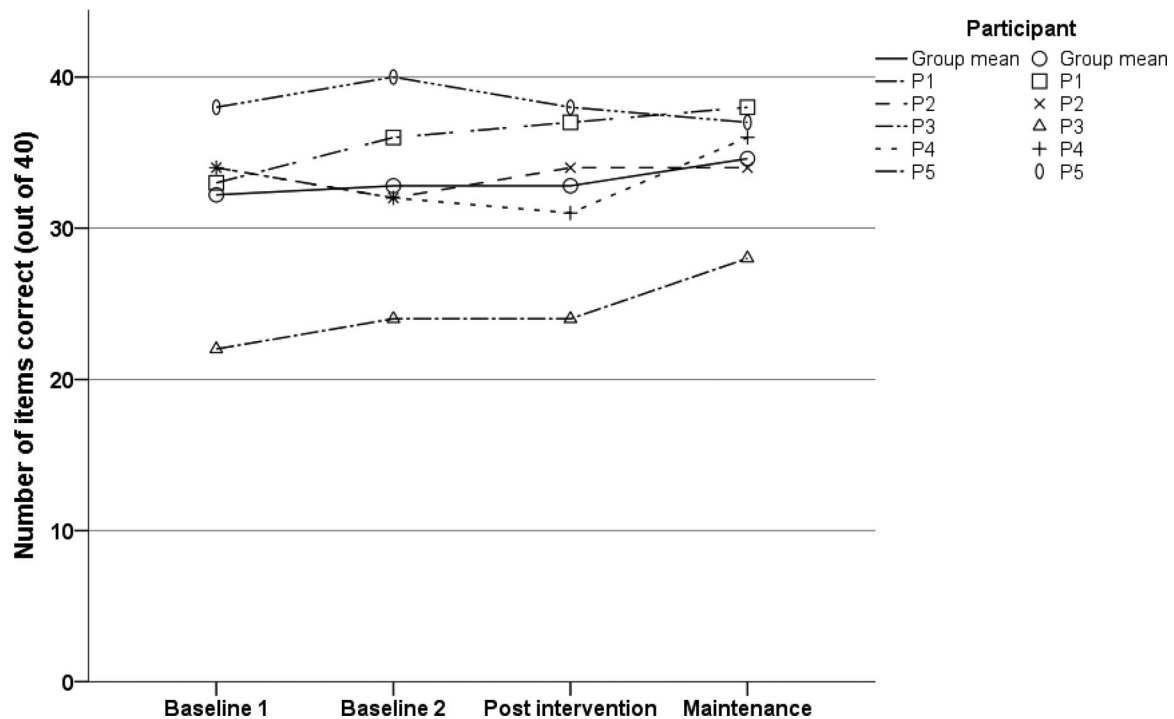


FIGURE 5 Synonym matching task: Number of correct items by participant over time

following intervention. Performance in the TROG-2 slightly decreased post-intervention, while post-therapy AIQ-21 ratings indicated improved emotional state/well-being.

P2 showed positive change in combination ratio post-intervention. On the story completion test, P2 displayed significant gains in number of well-formed utterances from baseline through to maintenance. P2's TROG-2 scores showed no change following intervention. Increased communication ratings on the AIQ-21 were evident, while participation ratings decreased after the intervention.

P3 had more severe aphasia, resulting in a restricted inventory of word combinations. P3 displayed considerable variation in combination ratio across assessment points. With regard to the adapted story completion test, P3 was unable to retrieve any answers other than 'dunno' prior to the intervention, but could give differentiated responses after the intervention, where he showed promising gains, especially with regard to well-formed utterances (increased number of well-formed utterances across post-intervention probes). TROG-2 scores improved after the intervention, as did his AIQ-21 ratings for well-being/emotional state. However, P3 rated his communication more negatively after the intervention.

Despite more impaired language output compared with other participants, P4 showed higher values in combination ratio post-intervention. P4's performance in the adapted story completion test revealed little evidence of availability of target constructions after the intervention,

but a slight increase in the number of well-formed utterances. There was a slight decrease in number of TROG-2 blocks correct after the intervention. AIQ-21 post-therapy ratings were more positive for well-being/emotional state, but more negative for communicative skills.

P5 showed increased combination ratio between the baseline and post-intervention probe, and this performance stayed stable until the final maintenance assessment. There were significant gains in well-formed utterances and expected answers in the story completion test. However, this result was not stable across post-intervention and maintenance assessment points. In addition to these positive results, P5's TROG-2 performance was enhanced after the intervention, while AIQ-21 ratings of communicative skills decreased after the intervention.

DISCUSSION

The present investigation piloted a novel computerized intervention to enhance the connected speech of individuals with NFA. The intervention was motivated by usage-based principles, training semi-fixed constructions with high functional value and then extending them by the process of superimposition of new lexical information into construction 'slots'. Moreover, the three intervention phases employed psycholinguistic and neuroscientific learning elements (e.g., structural priming, error-reduction strategies) and self-voice, a social-motivational

learning component. Findings from a case series of five individuals with NFA revealed promising changes in some participants which encourage further development of the intervention programme and future testing of its effectiveness in a larger trial. This will allow exploration of the profiles of individuals who benefit most (and least) from the intervention. The intervention is acceptable to this client group. Acknowledging this is a small case series, participants with NFA and their CPs rated the intervention as helpful, with slightly higher ratings from participants with NFA.

This study, for the first time, used the FLAT variable 'combination ratio' (amount of connected speech) to evaluate outcomes of aphasia intervention. Regarding the research question, 'Is there evidence that after intervention participants with NFA demonstrate enhanced connected speech, as measured by a higher proportion of multiword utterances in narratives', the intervention showed potential to enhance the ability to combine words into well-formed utterances for a subset of participants. However, since there was considerable variability in the baseline probes and between the participants, further investigations should explore how pre/post-difference scores correspond to small/medium/large effects.

Combination ratio is a useful variable to characterize an individual's ability to combine single words into well-formed word combinations. It is therefore well suited to profile spontaneous speech before and after therapy. FLAT offers the advantage that once a sample is transcribed and formatted, it analyses words and combinations automatically (and therefore objectively). Aside from combination ratio, the FLAT also provides other frequency-based variables, for instance association strength between the component words of two- and three-word combinations. Such variables may be useful to detect change in the creativity and flexibility of combinations. Since frequency-based measures are well-suited for naturalistic, spontaneous speech samples, they are worth further exploration as an outcome measure in usage-based interventions.

Analysis of combination ratio values does not answer the question of whether participants showed an increased use of trained constructions after intervention. To address this question, performance on an adapted story completion test was analysed. The results showed trends towards greater availability of target constructions in the majority of participants, but was only significant for one out of the five participants. Results revealed more evidence for positive change with regard to the number of well-formed utterances.

The limited outcomes on the 'expected answers' category of the adapted story completion test might be related

to the variable cloze probabilities that the items elicited in the norming sample. Despite adjusting the scoring procedure to account for this, 'expected answer' scores fluctuated considerably across baseline and post-intervention probes. The findings suggest that the second layer of scoring, number of well-formed utterances, was more sensitive to grammatical change following intervention. While this task was well-suited for the purposes of studies by Goodglass et al. (1972) and Gleason et al. (1975), the adapted version was not ideal to evaluate the availability of the target constructions used in the present intervention. Additionally, the nature of the instruction (completing a scenario with a phrase or sentence) may not be suitable for participants with global aphasia and/or AoS. However, despite the challenges in the present study, a story completion format might be the only valid method available of eliciting target constructions. In future studies, it would be desirable to devise an improved story completion task and extend the method to include both trained and untrained items, as is common in the assessment of naming performance (e.g., Nardo et al., 2017).

The availability of constructions and their variations could be compared before and after intervention and would thereby add a valuable layer to an evaluation of whether this type of intervention facilitates more productive use of target constructions. For example, P1's narrative revealed the presence of the '[REFERENT] *is*-TENSE [PROPERTY]' construction, with variations including: '*he's really funny*' and '*it's okay*'. Such a qualitative analysis would provide insight into the availability and use of specific constructions in spontaneous speech. Such a procedure would be similar to the techniques applied in Dąbrowska and Lieven (2005) and Lieven et al. (2009) who analysed fixed phrases and frames with slots in the language of children.

Whitworth et al. (2018) suggest that one baseline probe might be sufficient for some connected speech measures (e.g., narratives). If sampling takes place twice within a baseline period, Whitworth et al. (2018) showed that samples were more stable when taken 3 weeks apart than when they are taken within a shorter interval (1 week). In contrast to Whitworth et al. (2018), the current case series indicates considerable intra-individual variation in connected speech across two baseline assessment points (2 weeks apart). However, the present study reported descriptive statistics, while Whitworth et al. (2018) applied inferential statistics. Future evaluations of this intervention should further examine intra- and inter-individual stability of connected speech probes (potentially comparing picture-based narratives with monologic speech elicited by open questions), ideally using a baseline of at least 4 weeks to ensure sampling occurs with a minimum interval of 3 weeks.

The intervention resulted in some change in TROG-2 in two participants, indicating that with further development it might facilitate gains in both sentence comprehension and production. Furthermore, there is tentative evidence that the intervention may raise awareness of communicative disabilities in some participants, as reflected in more elevated scores on the AIQ-21 communication domain (i.e., communication was judged to be worse after intervention). The results also reflect positive changes in participants' perceived well-being and emotional state, a finding in line with the suggestion that practising familiar expressions might have non-linguistic effects such as improved well-being and quality of life (Stahl & Van Lancker Sidtis, 2015).

Finally, Bhogal et al. (2003) recommend at least 8.8 h over 11 weeks for speech and language therapy to be effective. The present intervention was delivered at a lower dose (eight 60-min intervention sessions over 4 weeks). This was supplemented by self-managed home practice. However, the exact amount of time that participants spent on these self-managed activities was not recorded in the current version of the software. Future investigations could measure therapeutic dose more accurately and explore the relationship between dose and outcome.

A larger trial based on the current, promising findings is warranted. Ideally, a researcher blind to the assessment point should conduct the assessment battery. A future trial should consider including standardized assessments as suggested by the ROMA consensus statement (Wallace et al., 2019), to support the availability of big data enabling the exploration of effectiveness of aphasia interventions.

CONCLUSIONS

This novel intervention for NFA, underpinned by usage-based CxG, was based on common constructions and their communicative functions. It was acceptable to participants and their CPs. The intervention shows promise as an effective therapy for some individuals with NFA, with gains in both sentence comprehension and production. Results appear to reveal potential for transfer to participants' connected speech, an important outcome given that it is often difficult for aphasia therapies to achieve transfer to the multiword level (Webster et al., 2015). Outcome measures for future trials could examine further the value of frequency-based variables and refine the story completion task, probing treated and untreated constructions. This study can inform larger trials in which hypotheses, generated by the present case series, could be tested, for instance with regard to the relationships between aphasia severity, dose and outcome.

ACKNOWLEDGEMENTS

The authors thank all the participants who took part in this project. Special thanks to Mike Coleman for programming the word monitoring game. Thanks also to Yana Arkhipova for transcribing the narrative speech samples; and Anna Volkmer for her input in creating the acceptability questionnaires.

CONFLICT OF INTEREST

There is no potential conflict of interest.

REFERENCES

- Bhogal, S.K., Teasell, R. & Speechley, M. (2003) Intensity of aphasia therapy, impact on recovery. *Stroke; A Journal of Cerebral Circulation*, 34(4), 987–992. <https://doi.org/10.1161/01.STR.0000062343.64383.D0>
- Bilda, K. (2011) Video-based conversational script training for aphasia: A therapy study. *Aphasiology*, 25(2), 191–201. <https://doi.org/10.1080/02687031003798254>
- Bishop, D.V.M. (2003) *Test for Reception of Grammar, Version 2 (TROG-2)*. London: Pearson.
- BNC. (2007) BNC, version 3 (BNC XML Edition). Distributed by Oxford University Computing Services on behalf of the BNC Consortium. Retrieved from <http://www.natcorp.ox.ac.uk/>
- Bruns, C. (2018) *A usage-based approach to language processing and intervention in aphasia. Doctoral Thesis*. University College London.
- Bruns, C., Varley, R., Zimmerer, V.C., Carragher, M., Brekelmans, G. & Beeke, S. (2019) I don't know": A usage-based approach to familiar collocations in non-fluent aphasia. *Aphasiology*, 2(33), 140–162. <https://doi.org/10.1080/02687038.2018.1535692>
- Conroy, P. & Scowcroft, J. (2012) Decreasing cues for a dynamic list of noun and verb naming targets: A case-series aphasia therapy study. *Neuropsychological Rehabilitation*, 22(2), 295–318. <https://doi.org/10.1080/09602011.2011.641434>
- Dąbrowska, E. (2014) Recycling utterances: A speaker's guide to sentence processing. *Cognitive Linguistics*, 25(4), 617–653. <https://doi.org/10.1515/cog-2014-0057>
- Dąbrowska, E. & Lieven, E. (2005) Towards a lexically specific grammar of children's question constructions. *Cognitive Linguistics*, 16(3), 437–474. <https://doi.org/10.1515/cogl.2005.16.3.437>
- Fletcher, M. & Birt, D. (1983) *Storylines: Picture sequences for language practice*. London: Longman.
- Forrin, N.D. & MacLeod, C.M. (2017) This time it's personal: The memory benefit of hearing oneself. *Memory (Hove, England)*, 26(4), 574–579. <https://doi.org/10.1080/09658211.2017.1383434>
- Franklin, S., Turner, J. & Ellis, A. (1992) *The A.D.A. Auditory Comprehension Battery*. York, UK: University of York.
- Gleason, J.B., Goodglass, H., Green, E., Ackerman, N. & Hyde, M.R. (1975) The retrieval of syntax in Broca's aphasia. *Brain and Language*, 2, 451–471. [https://doi.org/10.1016/S0093-934X\(75\)80083-6](https://doi.org/10.1016/S0093-934X(75)80083-6)
- Goldberg, A.E. (2003) Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences*, 7(5), 219–224. [https://doi.org/10.1016/S1364-6613\(03\)00080-9](https://doi.org/10.1016/S1364-6613(03)00080-9)
- Goodglass, H., Gleason, J.B., Ackerman Bernholtz, N. & Hyde, M.R. (1972) Some Linguistic Structures in the Speech of a Broca's

- Aphasic. *Cortex; A Journal Devoted to the Study of the Nervous System and Behavior*, 8(2), 191–212. [https://doi.org/10.1016/S0010-9452\(72\)80018-2](https://doi.org/10.1016/S0010-9452(72)80018-2)
- Harmon, T.G., Jacks, A., Haley, K.L. & Faldowski, R.A. (2016) Listener perceptions of simulated fluent speech in nonfluent aphasia. *Aphasiology*, 30(8), 922–942. <https://doi.org/10.1080/02687038.2015.1077925>
- Helm-Estabrooks, N., Nicholas, M. & Morgan, A. (1989) *Melodic Intonation Therapy. Manual*. Austin, TX: Pro-Ed.
- Hoffmann, T.C., Glasziou, P.P., Boutron, I., Milne, R., Perera, R., Moher, D., et al. (2014) Better reporting of interventions: Template for intervention description and replication (TIDieR) checklist and guide. *BMJ (Online)*, 348(March), 1–12. <https://doi.org/10.1136/bmj.g1687>
- Israel, M., Johnson, C. & Brooks, P.J. (2001) From states to events: The acquisition of English passive participles. *Cognitive Linguistics*, 11(1–2), 103–130. <https://doi.org/10.1515/cogl.2001.005>
- Kaplan, E., Goodglass, H. & Weintraub, S. (2001) *Boston Naming Test*. 2nd Ed. Philadelphia, PA: Lippincott Williams & Wilkins.
- Kaplan, J.T., Aziz-Zadeh, L., Uddin, L.Q. & Iacoboni, M. (2008) The self across the senses: An fMRI study of self-face and self-voice recognition. *Social Cognitive and Affective Neuroscience*, 3(3), 218–223. <https://doi.org/10.1093/scan/nsn014>
- Kaschak, M.P., Kutta, T.J. & Coyle, J.M. (2014) Long and short term cumulative structural priming effects. *Language, Cognition and Neuroscience*, 29(6), 728–743. <https://doi.org/10.1080/01690965.2011.641387>
- Kay, J., Coltheart, M. & Lesser, R. (1992) *Psycholinguistic Assessments of Language Processing in Aphasia*. Hove, UK: Lawrence Erlbaum Associates.
- Kaye, R.C. & Cherney, L.R. (2016) Script Templates: A practical approach to script training in Aphasia. *Topics in Language Disorders*, 36(2), 136–153. <https://doi.org/10.1097/TL.D.0000000000000086>
- Lee, J. & Man, G. (2017) Language recovery in aphasia following implicit structural priming training: A case study. *Aphasiology*, 31(12), 1441–1458. <https://doi.org/10.1080/02687038.2017.1306638>
- Lieven, E., Salomo, D. & Tomasello, M. (2009) Two-year-old children's production of multiword utterances: A usage-based analysis. *Cognitive Linguistics*, 20(3), 481–507. <https://doi.org/10.1515/COGL.2009.022>
- Murray, L., Timberlake, A. & Eberle, R. (2007) Treatment of Underlying Forms in a discourse context. *Aphasiology*, 21(2), 139–163. <https://doi.org/10.1080/02687030601026530>
- Nardo, D., Holland, R., Leff, A.P., Price, C.J. & Crinion, J.T. (2017) Less is more: Neural mechanisms underlying anomia treatment in chronic aphasic patients. *Brain*, 140(11), 3039–3054. [https://doi.org/10.1044/1092-4388\(2010/09-0204\)](https://doi.org/10.1044/1092-4388(2010/09-0204))
- Nobis-Bosch, R., Springer, L., Radermacher, I. & Huber, W. (2011) Supervised home training of dialogue skills in chronic aphasia: A randomized parallel group study. *Journal of Speech, Language, and Hearing Research*, 54(4), 1118–1136. [https://doi.org/10.1044/1092-4388\(2010/09-0204\)](https://doi.org/10.1044/1092-4388(2010/09-0204))
- Riches, N.G. (2013) Treating the passive in children with specific language impairment: A usage-based approach. *Child Language Teaching and Therapy*, 29(2), 155–169. <https://doi.org/10.1177/0265659012466667>
- Ruiter, M.B., Kolk, H.H.J. & Rietveld, T.C.M. (2010) Speaking in ellipses: The effect of a compensatory style of speech on functional communication in chronic agrammatism. *Neuropsychological Rehabilitation: An International Journal*, 20(3), 423–458. <https://doi.org/10.1080/09602010903399287>
- Schlenck, K.-J., Schlenck, C. & Springer, L. (1995) *Die Behandlung des schweren Agrammatismus: Reduzierte Syntax-Therapie (REST)*. Stuttgart: Georg Thieme Verlag.
- Springer, L., Huber, W., Schlenck, K.-J. & Schlenck, C. (2000) Agrammatism: Deficit or compensation? Consequences for aphasia therapy. *Neuropsychological Rehabilitation: An International Journal*, 10(3), 279–309. <https://doi.org/10.1080/096020100389165>
- Stahl, B., Mohr, B., Dreyer, F.R., Lucchese, G. & Pulvermüller, F. (2016) Using language for social interaction: Communication mechanisms promote recovery from chronic non-fluent aphasia. *Cortex; A Journal Devoted to the Study of the Nervous System and Behavior*, 85, 90–99. <https://doi.org/10.1016/j.cortex.2016.09.021>
- Stahl, B. & Van Lancker Sidtis, D. (2015) Tapping into neural resources of communication: Formulaic language in aphasia therapy. *Frontiers in Psychology*, 6, 1526. <https://doi.org/10.3389/fpsyg.2015.01526>
- Swinburn, K., Best, W., Beeke, S., Cruice, M., Smith, L., Pearce-Willis, E., et al. (2018) A concise patient reported outcome measure for people with aphasia: The aphasia impact questionnaire 21. *Aphasiology*, <https://doi.org/10.1080/02687038.2018.1517406>
- Swinburn, K., Porter, G. & Howard, D. (2004) *The Comprehensive Aphasia Test*. Hove, UK: Psychology Press.
- Thompson, C.K. & Shapiro, L.P. (2005) Treating agrammatic aphasia within a linguistic framework: Treatment of underlying forms. *Aphasiology*, 19, 1021–1036. <https://doi.org/10.1080/02687030544000227>
- Van Lancker Sidtis, D. (2012) Formulaic language and language disorders. *Annual Review of Applied Linguistics*, 32, 62–80. <https://doi.org/10.1017/S0267190512000104>
- Varley, R. (2011) Rethinking aphasia therapy: A neuroscience perspective. *International Journal of Speech–Language Pathology*, 13(1), 11–20. <https://doi.org/10.3109/17549507.2010.497561>
- Varley, R., Cowell, P.E., Dyson, L., Inglis, L., Roper, A. & Whiteside, S.P. (2016) Self-administered computer therapy for apraxia of speech: Two-period randomized control trial with crossover. *Stroke; A Journal of Cerebral Circulation*, 47(3), 822–828. <https://doi.org/10.1161/STROKEAHA.115.011939>
- Wallace, S.J., Worrall, L., Rose, T., Le Dorze, G., Breitenstein, C., Hilari, K., et al. (2019) A core outcome set for aphasia treatment research: The ROMA consensus statement. *International Journal of Stroke*, 14(2), 180–185. <https://doi.org/10.1177/1747493018806200>
- Webster, J., Whitworth, A. & Morris, J. (2015) Is it time to stop “fishing”? A review of generalisation following aphasia intervention. *Aphasiology*, 29(11), 1240–1264. <https://doi.org/10.1080/02687038.2015.1027169>
- Wechsler, D. (2011) *Wechsler Abbreviated Scale of Intelligence—Second Edition (WASI-II)*. San Antonio, TX: Psychological Corporation.
- Whiteside, S.P., Inglis, A.L., Dyson, L., Roper, A., Harbottle, A., Ryder, J., et al. (2012) Error reduction therapy in reducing struggle and grope behaviours in apraxia of speech. *Neuropsychological Rehabilitation*, 22(2), 267–294. <https://doi.org/10.1080/09602011.2011.639614>
- Whitworth, A., Budiman, P., Beales, A., Wynn, S., Cartwright, J. & Claessen, M. (2018) Just how stable is our discourse? A hurdle

- for measuring therapeutic change (or not). *Aphasiology*, 32(sup1), 246–247. <https://doi.org/10.1080/02687038.2018.1493427>
- Zimmerer, V.C., Coleman, M., Hinzen, W. & Varley, R.A. (in prep.). Reliance on common word combinations correlates with degree of syntactic impairment in aphasia.
- Zimmerer, V.C., Newman, L., Thomson, R., Coleman, M. & Varley, R. (2018) Automated analysis of language production in aphasia and right hemisphere damage: Frequency and collocation strength. *Aphasiology*, 32(11), 1267–1283. <https://doi.org/10.1080/02687038.2018.1497138>
- Zimmerer, V.C. & Wibrow, M. (2015) Frequency in Language Analysis Tool.
- Zimmerer, V.C., Wibrow, M. & Varley, R.A. (2016) Formulaic language in people with probable Alzheimer’s Disease: A frequency-based approach. *Journal of Alzheimer’s Disease*, 53(3), 1145–1160. <https://doi.org/10.3233/JAD-160099>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Bruns C, Beeke S, Zimmerer VC, Bruce C, Varley RA. Training flexibility in fixed expressions in non-fluent aphasia—a case series report. *International Journal of Language & Communication Disorders*. 2021;56:1009–1025. <https://doi.org/10.1111/1460-6984.12652>