Piloting Building Early Sentences Therapy for pre-school children with low language abilities: an examination of efficacy and the role of sign as an active ingredient

Anastasia Trebacz[12], Cristina McKean[2], Helen Stringer[2], Sean Pert[3]
[1]School of Medicine, University of Sunderland
[2]School of Education, Communication and Language Sciences, Newcastle University
[3]Division of Human Communication, Development and Hearing, University of Manchester

Address correspondence to: Anastasia Trebacz, School of Medicine, University of Sunderland, Chester Road, Sunderland, SR1 3SD, UK.
Email: Anastasia.trebacz@sunderland.ac.uk

Acknowledgements
Professor James Law acted as a co-supervisor to the first Author's PhD from which this study results. We would like to acknowledge his role and contribution to this work. The Authors would like to sincerely thank the schools and families who took part in this research project.

Data availability statement:
Consent was not obtained for data sharing.

Funding statement: This work arose from the first author's PhD studentship which was generously funded by the Newcastle University Research Excellence Academy.

Conflict of interest statement:
The Authors declare no conflicts of interest.

Ethics approval statement: Ethical approval was granted for this work by the Newcastle University Faculty of Humanities and Social Sciences Ethics Committee.

Patient consent statement:
Fully informed consent was obtained from parents on behalf of their children to take part in the study. Additionally, parents were informed that if children became upset or did not want to take part they would immediately be withdrawn from the session and the study without any negative consequences.

Permission to reproduce material from any other sources:
Permission to share materials from the BEST intervention was granted by the second author.

Abstract

Background:
Early intervention is recommended for pre-school children with low language. However, few robustly evaluated language interventions for young children exist. Furthermore, in many interventions the theoretical underpinnings are underspecified and the 'active ingredients' of the interventions not tested. This paper presents a quasi-experimental study to test the efficacy and examine the active ingredients of Building Early Sentences Therapy (BEST): an intervention based on usage-based theory designed to support young children to understand and produce two, three and four clause element sentences. BEST manipulates the input children hear to support them to harness the cognitive mechanisms hypothesized in usage-based theories to promote the development of abstract linguistic representations. One such input manipulation is the use of signing alongside verbal input signaling both content and morphology of target sentences.

Aims:
To examine whether 1) BEST is more effective than Treatment as Usual and 2) signing of content and morphology is an active ingredient of the intervention.

Methods & procedures:
A quasi-experimental study recruited children aged 3;5-to-4;5-years from thirteen schools. Schools were assigned to receive either BEST with sign, BEST without sign or Treatment as Usual (TAU). TAU group received their usual classroom provision. Across arms schools were matched with respect to classroom oral language environment and indices of deprivation. Participants were forty-eight children (twenty-eight boys) with expressive and/or receptive language abilities ≤16th centile measured using the New Reynell Developmental Language Scales (NRDLS). Outcomes, gathered by researchers blind to treatment arm, were NRDLS production and comprehension standard scores and measures of production of targeted sentence structures.

Outcomes & results:
Primary outcomes indicate that BEST with sign was significantly more effective than TAU with respect to NRDLS production standard score but not comprehension. The advantage for production was maintained at follow-up. BEST without sign was significantly more effective than TAU on measures of targeted vocabulary, sentence structure and morphology.

The results from this quasi-experimental study provide evidence for the efficacy of a usage-based intervention on expressive language outcomes for preschool children with low language. There is also evidence to support the inclusion of sign as an active ingredient, and so efforts to train interventionists in its use are worthwhile. Patterns of finding across outcomes suggest signing of content and morphology may support the development of abstract linguistic representations and accelerate language learning. Given these positive results and the scale of this study, a fully powered randomised controlled trial is warranted.

What is already known on this subject:
Robust language skills are crucial for positive social, emotional, academic, and economic outcomes across the lifespan. There is a paucity of robustly evaluated interventions for preschool children with language difficulties. The development of such interventions is crucial for ameliorating language difficulties and promoting positive educational and psychosocial outcomes.

What this paper adds:
This paper evaluates BEST, a novel usage-based language intervention targeting children with language difficulties in the early years. Findings indicate that a usage-based intervention is efficacious for treating language difficulties. In particular BEST benefited expressive language development, bringing benefits to both treated and untreated language structures and improving standard scores. Further evaluation is warranted. The role of sign as an active ingredient is also supported.

Clinical Implications
Findings suggest that BEST may be effective for targeting children who have been identified as having language difficulties. In particular, expressive language may be improved when the intervention is delivered as it was originally manualised, including a signing system to represent content and grammatical morphology. More broadly, these findings also provide preliminary evidence that the use of a signing system does not hinder oral language development in children with language difficulties and may conversely support their expressive language. Future research exploring the role and underpinning mechanisms of sign in language intervention is warranted.

Keywords:

early years; language difficulties; language intervention; signing system; gesture; usage-based theory; quasi-experimental; pilot and feasibility study.

Most children acquire oral language with relative ease, however approximately 7.5% of children have persisting language difficulties that cannot be explained by the presence of another condition and which hinder their academic progress, socio-emotional development and communicative participation (Norbury et al., 2016). As such, these children are eligible for a diagnosis of Developmental Language Disorder (DLD) (Bishop et al., 2017) and require targeted or specialist intervention to reach their potential. Language difficulties are highly unstable in the early years (Reilly et al., 2014) and if they do not resolve by school entry, are likely to persist until at least age 11 (McKean et al., 2017). It is therefore crucial to provide effective early intervention in the pre-school years to support children at risk of or presenting with language disorder (Law et al., 2017). Despite the high level of need for such programs (Bercow, 2018), there are few robustly evaluated interventions for children with language difficulties in the early years (Law et al., 2017). Fewer still have clearly articulated theoretical underpinnings (Roulstone et al., 2012) despite the key role of theory in underpinning the development of interventions (Skivington et al., 2021).

The development and evaluation of complex interventions

Speech and language therapy (SLT) interventions are intrinsically complex in that they comprise numerous interacting components and have outcomes and content tailored to the individual(s) receiving the intervention (Craig et al., 2008). In recent years, Medical Research Council (MRC) and National Institute for Health Research (NIHR) guidance regarding the methods and phases required to develop and evaluate complex health interventions have been widely adopted across the behavioural and health sciences (Craig et al., 2008; Skivington et al., 2021), although are seldom explicitly applied in SLT research. These guidelines recommend that prior to definitive, large-scale trials, interventions require careful *development* such that the key active ingredients are defined, and the underpinning theory is explicitly described. This enables successful implementation with appropriate fidelity, ensuring that practitioners understand the importance of the components of the intervention and so do not 'water down' its implementation. Furthermore, it allows practitioners and researchers to make sense of its rationale: this 'sense-making' being key to novel interventions becoming 'normalised' practice (May et al., 2018).

The present study uses a quasi-experimental design to evaluate the efficacy of Building Early Sentences Therapy (BEST) (McKean et al., 2013), a complex intervention for children with language difficulties that is underpinned by constructivist, or usage-based theories of language development (Tomasello, 2003) when compared to Treatment as Usual. BEST manipulates the language learning context and input children hear to support them to harness the cognitive mechanisms posited by usage-based theorists to underpin language learning. In this way BEST aims to increase children's ability to use a range of simple 2, 3 and 4 clause element sentences flexibly, with a range of verbs and nouns and with appropriate grammatical morphology.

In addition to considerations of efficacy, Frizelle & McKean (2022) have argued that research to identify which aspects of an intervention are 'active ingredients' and so drive change for

the child is also vital for the field to move forward. They suggest that knowledge of precisely which aspects of an intervention drive change can support Speech and Language Therapists to implement research in clinical practice with sufficient fidelity to retain its effectiveness in practice, even when tailoring the approach to the needs of individual children. Furthermore, they suggest this knowledge can drive efficiency such that lower dosages of intervention are likely to be required if the active ingredients are prioritised in delivery. To that end this study also considers whether a key component of BEST, the use of a sign system, is an active ingredient.

Usage-based theory and its potential use in intervention design

In the last 20 years, a new family of explanatory models of typical child language development has emerged, with a large and rapidly growing body of empirical work that supports its assertions (Ambridge et al., 2006; Tomasello, 2003). These 'constructivist' or 'usage-based' theories suggest that the adult end-state of language acquisition is not a set of grammatical rules per se, but rather an inventory of constructions which are linked to the pragmatic and semantic functions which they can communicate (Croft & Cruse, 2004). These constructions vary along a continuum of abstractness and hence flexibility with respect the lexical items which can be placed into them; ranging from the highly concrete and inflexible (e.g. 'How do you do?') to the highly abstract, and flexible (e.g. NOUN1 + VERB + NOUN2 – meaning NOUN1 acts on NOUN2 and NOUN2 is affected), and with other constructions falling somewhere in between (e.g. X wouldn't Y let alone Z). Children's knowledge of these constructions is thought to be learned slowly and incrementally, and the progress of this learning determined both by the nature of the input and the child's cognitive abilities to construct abstract representations. Hence children build or 'construct' their knowledge of grammar over time and in response to their own and to other's use of language for specific communicative purposes (Tomasello, 2003).

Tomasello (2003) described a constructivist, usage-based account of the process of language acquisition from first words to end-state adult 'grammar' suggesting that this process is driven by two human characteristics: intention-reading (the ability to create shared understanding of communicative intentions within an interaction with a person) and pattern finding (the ability to identify regularities and patterns in complex inputs). In this account Tomasello describes five stages of language acquisition which are posited to proceed once multi-word utterances occur: 1) frozen phrases; 2) lexically specific constructions; 3) abstract constructions; 4) paradigmatic categories and 5) retreat from over-generalisation. Importantly, it is not the child who moves through these five stages, such that all their language knowledge 'as a piece' moves from one stage to the next. Rather, individual language constructions take this journey towards a highly abstract end state, progressing at different speeds for different constructions, and with differing endpoints in the journey. Tomasello's account also describes the cognitive mechanisms brought to bear on the learning process which allow children to move from one stage to the next. Appendix 1 provides additional detail regarding these stages and mechanisms.

Despite this large body of evidence these 'constructivist' or 'usage-based' theories have rarely been explicitly applied to the design of interventions for children with language difficulties (see Riches (2013) for a notable exception and complexity theory (Van Horne et al., 2018) which has implicit usage-based underpinning). This is a significant missed opportunity, as these theoretical accounts and their supporting empirical data provide rich detail regarding both the cognitive or learning mechanisms which children harness in the process of language learning and the nature of the input children need to hear to leverage those mechanisms. Hence, they provide numerous candidate 'active ingredients' which could be manipulated for intervention design (see appendix 1).

Building Early Sentences Therapy

Building Early Sentences Therapy (BEST) (McKean et al., 2013), aims to support pre-school children to move through the first three of Tomasello's developmental stages: moving from frozen phrases to item-based constructions to abstract representations for simple 2, 3 and 4 clause element sentences. The goal is to facilitate the development of the child's flexible use of sentence structures, increasing their range of communicative functions. Furthermore, given the finding that the development of abstract constructions may support children to learn other, related structures more readily (Langacker, 2000), we posit that supporting children to create abstract representations has the potential to accelerate future language learning for related, novel constructions.

BEST manipulates the input the child is exposed to and the language learning context to support the child's use of the cognitive mechanisms of intention-reading, cultural learning, schematisation, categorisation, analogy, mapping, and retention (see appendix 1). Although apparently trivial for typically developing children, many of these cognitive and learning mechanisms are challenging for children with or at risk of Language Disorder (e.g. cues within the input such as phonotactic or morphological patterns, which drive pattern finding and thence analogy, are less readily accessible to language impaired children due to underlying phonological processing impairments (Chiat, 2001); see also Riches et al., 2005 with regards to mapping and retention).

To exploit the cognitive mechanisms outlined above, the input presented to children during BEST sessions is designed to exaggerate the features which promote intention reading, cultural learning, schematisation, categorisation, analogy, mapping, and retention, thus supporting the development of abstract representations. The methods of input manipulation applied in BEST and hypothesised to be active ingredients of the intervention are summarised in appendix 1. This study aims to test whether BEST, an intervention underpinned by usage-based theory, is effective in improving expressive and receptive language in pre-school children with low language abilities.

Sign as an active ingredient in BEST

BEST uses a form of signed supported English, based on the signing systems Makaton (see Walker & Armfield, 1981) and Paget Gorman Signed Speech (Rowe, 1981) to mark both content words and grammatical morphology. The role of sign is to highlight cues within the input such as content words to support semantic mapping (Vogt & Kauschke, 2017) and phonotactic and morphological patterns, which drive pattern finding and thence analogy, that typically developing children harness to create abstract representations of arguments and predicate argument structures (Tomasello, 2003). Chiat (2001) posits that children with or at risk of language disorder find the mapping of semantic and syntactic structures harder to access due to underlying phonological impairment and so have less opportunities to bootstrap learning. By drawing attention to these cues, sign makes them more accessible to children with or at risk of language difficulties and may also reduce the overall processing load (Goldin-Meadow, 2011).

McKean et al (2013) hypothesise several benefits of sign in the BEST intervention. Firstly, the signing of content words may support verb semantic mapping, thus also improving representations of predicate argument structure. The signing of content words may also reduce the overall cognitive processing load on children, allowing them to access the cues delivered in the intervention and accelerating their learning. The signing of morphological items may promote mapping of the morphological frame, in turn supporting bootstrapping of semantic roles and the development of abstract representations, and it may also facilitate the use and understanding of grammatical markers.

There is evidence that sign, and iconic gesture, offer benefits for children's language in the early years (see Capone & Mcgregor, 2004 for a review). Despite early benefits, typically developing children seem to rely less on gesture and more on oral language for interpreting meaning from around 2;5-3-years-old. Children with language difficulties may continue to benefit more from sign; specifically it may support children to scaffold meaning when comprehension breaks down (Botting et al., 2010), support semantic encoding (Vogt & Kauschke, 2017) and word learning (van Berkel-van Hoof et al., 2019). Evidence suggests that children with and without language impairment utilise speech and gesture in a similar way, however language impaired children benefit more from gesture to an older age (Lüke et al., 2020). Considering the evidence, researchers and clinicians have advocated for the use of sign and gesture to support children with and without language difficulties (e.g. Dockrell et al., 2012; van Berkel-van Hoof et al., 2019). However, there is little available evidence evaluating the efficacy of a signing system in preschool language interventions and no evidence regarding its effect on the development of morphosyntax or predicate argument structure.

Despite the potential benefits, attitudes to the use of sign in hearing children can, on occasion, be a barrier to acceptability due to parental concern that the use of sign may delay oral language development (Abbott & Lucey, 2005). Furthermore, reliably learning to use a signing system requires an investment from the individual delivering the intervention and barriers of confidence or time may impede its uptake and use. This study therefore seeks to

test the hypothesis that sign is an active ingredient in BEST. To that end this study compares outcomes in expressive and receptive language for pre-school children with low language from three interventions: BEST with sign, BEST without sign and usual treatment. If BEST with sign is more effective than BEST without sign, then sign can be seen as a key 'active ingredient' of BEST and also potentially as a useful adjunct to other interventions. If BEST without sign is more effective than treatment as usual this suggests that the other input manipulations, based on usage-based theory, used in BEST are also active ingredients (see appendix 1). This study therefore addresses the following research questions:

For pre-school children (aged 3;5-to-4;5-years) with low language abilities:

1. Is BEST with sign more effective than treatment as usual in improving expressive and/or receptive language suggesting BEST is an effective intervention?
2. Is BEST without sign more effective than treatment as usual suggesting the contextual and input manipulations used in BEST are active ingredients of the intervention?

Methods

Ethics

The project was granted ethical approval by the Newcastle University Faculty of Humanities and Social Sciences Ethics Committee. Informed consent was obtained from Headteachers, teachers and parents prior to participation.

Study Design

This paper reports a repeated measures quasi-experimental single blind study measuring efficacy outcomes under research conditions. As with true experimental studies, quasi-experimental studies aim to establish a cause-and-effect relationship between independent and dependent variable. However, unlike in true experimental studies, participants in quasi-experimental studies are not randomly assigned to treatment arms (Harris et al., 2006). In the present study, participants were assigned to groups using the matching procedure outlined below, rather than being randomly assigned.

Participant recruitment

A participant flowchart is presented in figure 2 below.

Schools

The study took place in one Local Authority in England between April 2018 and July 2019. Local Speech and Language Therapists (SLTs) selected schools that were likely to have a high prevalence of language difficulties based on their clinical experience and knowledge of the population served acted as gatekeepers to schools, making initial contact with Head Teachers. Following Head Teacher consent, thirteen schools were assigned to one of three treatment arms: BEST with sign, BEST without sign or Treatment as Usual (TAU).

Children

Teachers recommended children about whom they had concerns regarding their language development and/or achieving expected language milestones. With informed parental consent, children were assessed on the BEST eligibility assessment (McKean et al., 2013) and the assessment battery (see measures section below). Fifty children met the following criteria and were eligible to take part: aged 3;5-to-4;5-years-old; monolingual English speaker (based on parent/carer self-report); language abilities falling at or below the 16th centile of the New Reynell Developmental Language Scales (NRDLS) comprehension and/or production subscales (Edwards et al., 2011); and demonstrated triadic attention, symbolic play and imitation in play during the assessment battery. Exclusionary criteria were: sensorineural hearing impairment, severe visual impairment, diagnosed learning disability or Autism Spectrum Disorders.

Matching Procedure

Schools were matched across treatment arms according to two criteria: 1) The classroom oral language environment using the Communication Supporting Classrooms Observation Tool (CSCOT) score (Dockrell et al., 2012); and 2) The level of deprivation affecting the school using the Income Deprivation Affecting Children Index (IDACI) score using school postcode (Department of Communities and Local Government, 2015). The goal was to ensure equivalent distribution in these metrics across treatment arms. This process was repeated three times over the course of the study as further recruitment waves took place. Baseline characteristics can be found in tables 2 and 3 below.

Procedures

The study consisted of 4 phases: 1) baseline assessment; 2) 8 weeks of intervention (either BEST with sign, BEST without sign or TAU); 3) immediate post-intervention outcome assessment; 4) six-week follow-up assessment. Recruitment and procedures were repeated 3 times over 3 school terms to achieve maximum participant numbers within the available resource (23 children in wave 1, 14 children in wave 2 and 13 children in wave 3). The study procedure is outlined in figure 1.

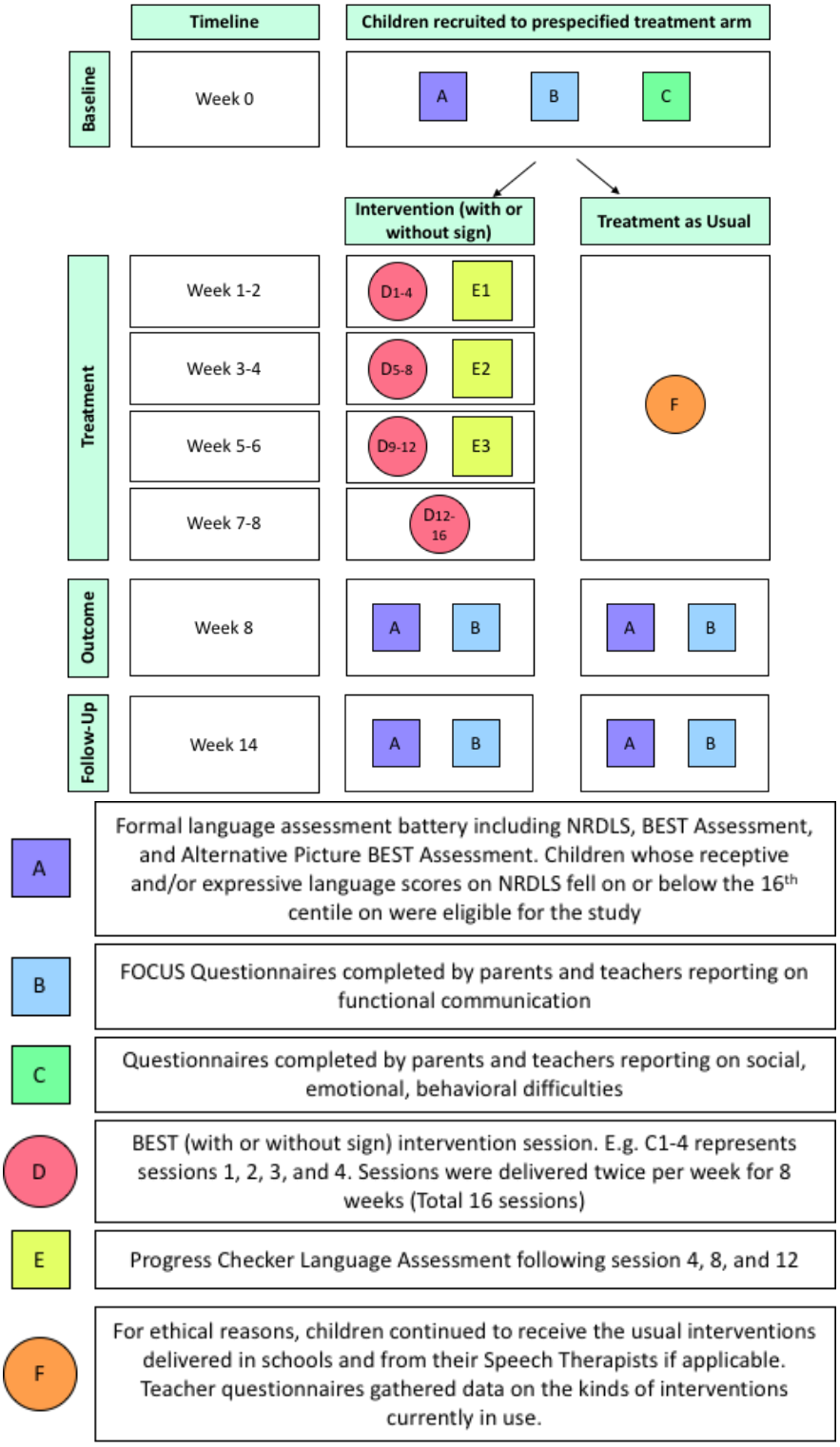


Figure 1: Perera Diagram displaying overall study processes. The cycle is repeated 3 times over 3 consecutive school terms.

Measures

Eligibility measures: children were assessed on an observational measure of symbolic play, triadic attention and imitation in play, to ensure they could access the BEST intervention input (BEST eligibility assessment, McKean et al., 2013) and for language difficulties on or below the 16th centile for comprehension and/or production subscales (NRDLS, Edwards et al., 2011) (also see additional eligibility criteria above).

Outcome measures: the primary outcome measures were standard scores on NRDLS production and comprehension subscales.

Secondary outcomes were 'Targeted BEST Assessment', 'Generalised BEST Assessment' and the FOCUS questionnaire.

'Targeted BEST Assessment'- This assessment probed the children's use of the targeted sentences used in BEST using a picture description task. Children were asked to describe sixteen pictures covering the range of sentence constructions targeted in the intervention. The children's spoken utterances were transcribed and scored with respect to the percentages correct of content words and morphological structures used when describing the pictures yielding a Targeted BEST content score and a Targeted BEST morphology score respectively (see appendix 4).

'Generalised BEST Assessment'- The generalisation of these structures to sentences with differing vocabulary items was also probed in a matched picture description task using non-targeted vocabulary- and scores calculated as above to yield Generalised BEST content and Generalised BEST morphology scores.

Finally, the Focus on the Outcomes of Children under Six (FOCUS- 36 Item Version) (Thomas-Stonell et al., 2012) was completed by parents and teachers[1] to provide a measure of communication activity and participation domains (FOCUS total score).

Assessment Reliability

To avoid bias, outcome and follow up assessments were conducted by research assistants (RAs) who were Undergraduate and Postgraduate SLT and Psychology students with experience delivering assessments to children. RAs were blind to the children's treatment arm. The first and second authors delivered robust assessment battery training and the assessment delivery protocol was manualised. Resolutions to scoring ambiguities were agreed upon and recorded. Assessments were audio recorded and checked for live transcription accuracy by RAs upon completion.

Interventions

BEST with sign

Structure and content: BEST is comprised of sixteen sessions lasting approximately 15-20 minutes delivered twice weekly over 8 weeks to groups of one-to-six children. Over its full course the intervention targets the structures in Table 1. For each structure, a number of early developing verbs are used (Morrison et al., 2003) (see appendix 2 for examples). These verbs are grouped into pairs, which have the same predicate argument structure (PAS), and which can be combined with the same nouns to make semantically plausible sentences.

[1] Communication with the Authors confirmed that teachers who knew the child well could complete the professional version of the FOCUS instead of Speech and Language Therapists as it was originally designed

Table 1: Verbs and Predicate argument structures targeted by the BEST programme and the use of Contrast and Variation in those structures

No. of arguments	Set	Argument Structure	Input	Output
1	**A**	Agent^{2} + Action^{1}	*laughing*	**sitting**
1	**B**	Agent^{2} + Action^{1}	*jumping*	**walking**
2	**C**	Agent^{1} + Action^{1} + Patient^{2}	*eating*	**washing**
2	**D**	Agent^{1} + Action^{1} + Patient^{2}	*riding*	**smelling**
2	**E**	Agent^{1} + Action^{1} + Patient$^{2(A)}$	*kissing*	**hugging**
2	**F**	Agent^{1} + Action^{1} + Patient^{2}	*kicking*	**brushing**
3	**G**	Agent^{1} + Action + Patient^{2} + Locative^{1}	*putting*	**putting**
3	**H**	Agent^{1} + Action + Patient^{1} + Locative^{2}	*pouring*	**pouring**
3	**I**	Agent^{1} + Action^{1} + Patient$^{2(A)}$ + Locative$^{2(B)}$	*putting*	**pouring**
3	**J**	Agent^{1} + Action + Patient^{2} + Benefactive^{1}	*giving*	**giving**
3	**K**	Agent^{1} + Action + Patient^{1} + Benefactive^{2}	*throwing*	**throwing**
3	**L**	Agent^{1} + Action^{1} + Patient$^{2(A)}$ + Benefactive$^{2(B)}$	*giving*	**throwing**

1- Contrast between Input & Output; 2- Variation within Input and Output; 2(A)- Variation within Input only; 2(B)- Variation within Input only

Within each session children are taken through a two-phase process three times:

Phase 1- Input (with variation): The child hears Verb 1 (e.g. eat) of the target structure (e.g. Agent + Action + Patient) used between 3 and 6 times with a 'frame' held constant and one 'slot' varied (e.g. The man is eating an apple, the man is eating an orange the man is eating a banana). Whilst hearing the input the child sees the actions being completed by the adult with miniature toys. Sign is used to provide a visual model of the sentence content and morphology as they are spoken and acted out with the toys. In order to support the child to begin to make links between the two constructions and so facilitate the cognitive process of analogy, the final example in the Input phase (Verb 1) switches to the Agent which will be used in the Output Phase (Verb 2) (see appendix 2).

Phase 2- Output (with variation and contrast): The child then sees the adult act out an event with the same PAS but with a contrasting verb, Verb 2 of the verb pair, and the child is

encouraged to describe what they see. This is repeated a number of times, again with a 'frame' held constant and one 'slot' varied (e.g. The teddy is washing an apple, the teddy is washing an orange, the teddy is eating a banana). The child then is allowed to act out the event with the toys while the adult provides verbal input of the target utterance. The adult then recasts the correct response verbally and with sign before moving on to ask the next child.

Signing system: Signs from the Makaton signing system (Walker & Armfield, 1981) were used to symbolise content words and Paget Gorman Signed Speech (PGSS) (Rowe, 1981) to represent grammatical morphology (Makaton does not offer this capability). BEST was originally designed to use only PGSS, however since the teacher questionnaires showed that most schools used Makaton in classrooms, the content words of Makaton were used to avoid exposing children to a different sign for the same item than one they already knew.

Homework: Following each session, parents receive a homework book which includes pictures for each of the target sentences that occur in the BEST programme (i.e. in both the Input and the Output conditions). For this homework the focus is on input, with parents encouraged to describe the pictures and so provide repeated input of the target sentences. The child is not expected to repeat or imitate these sentences but is praised and rewarded if they do so spontaneously.

Materials: BEST is a manualised intervention with standardised recording forms, homework booklets and reward charts which were used for all groups (LIVELY Group, 2019). A standardised set of toys was also used.

Usual classroom provision: Schools in the intervention arms continued to deliver all usual support to participating children.

BEST Intervention without sign

The BEST without sign intervention omits signing and is otherwise identical to BEST with sign, as described above.

Treatment as Usual (TAU)

Schools assigned to receive TAU continued to deliver usual classroom provision to all children. Usual provision is characterised in the results section below.

Class teachers from all participating schools completed a questionnaire providing information about their usual classroom language support activity prior to the school matching procedure. This was followed by observation by the first author using the CSCOT (Dockrell et al., 2012). From these data it was evident that children in all schools typically received frequent, high-quality language intervention to promote their oral language development. Many teachers reporting targeted one-to-one or small-group intervention daily, or on most days, for approximately 5-10 minutes. Additionally, teachers supported

children via strategies and activities including visual supports such as symbols and signs, storytelling, circle time sessions, teaching meaning and checking understanding.

Treatment fidelity

Treatment fidelity for BEST with and without sign was ensured in the following ways: The first author received training from the designers of BEST to deliver the intervention and supported an SLT to deliver BEST for eight weeks. The BEST manualised procedures were followed by the first author during intervention delivery, and standard materials were used across all sessions.

The fourth author conducted treatment fidelity checks using a checklist on a subset of the intervention sessions. One full video-recorded session from early in the intervention (within the first four sessions) was observed and any deviation from the BEST manual was reported so that this could be addressed. Treatment as usual, as outlined above, was variable and determined by each individual school. In line with other studies, fidelity for TAU was not evaluated.

Analysis

Multilevel modelling was used to compare treatment arms as a robust method to account for the hierarchical structure of the data as the children were nested within schools (Finch et al., 2019). Analyses were carried out in R (R Core Team, 2019). The dependent variables for the planned analyses were as follows:

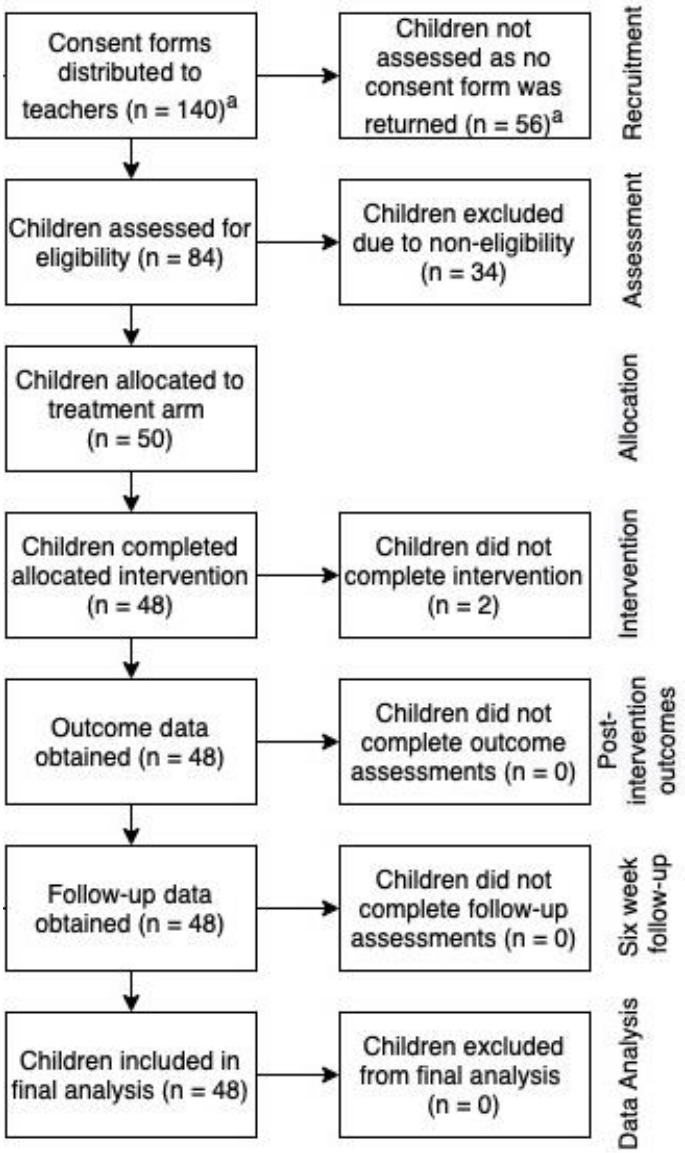
Primary outcomes:	1) NRDLS production; 2) NRDLS comprehension
Secondary Outcomes:	3) Targeted BEST content; 4) Targeted BEST morphology (see appendix 4); 5) Generalised BEST content; 6) Generalised BEST morphology
Functional Outcome:	7) FOCUS total score

Results

Forty-eight children (28 boys) completed the study. Figure 2 shows the participant flowchart. The retention rate was 96%. Two children did not complete the study due to already receiving intensive intervention (n=1) and a school withdrawing (n=1). Baseline characteristics are displayed in table 2 and 3. With regards to IDACI score, four of the twelve schools were in the top 20% least deprived postcode areas while the remaining eight were in the top 40% most deprived postcode areas. The mean IDACI decile was 5 (Department of Communities and Local Government, 2015). There was a significant difference between groups with respect to participants' sex [$\chi^2(2, n = 48) = 6.49, p = 0.039$], with the BEST with sign arm appearing to have a significantly higher proportion of boys than the other groups.

There was a significant difference between groups with respect to CSCOT scores (F(2, 45) = 3.51, p= .04) where BEST without sign CSCOT scores were higher than TAU by around 0.14 points (p=.04). There were no other significant differences at baseline.

Figure 2: Participant flowchart



[a]Ten forms were given to each school. Data was not obtained regarding numbers distributed to parents, therefore the number of children initially approached is unknown.

Table 2: Categorical descriptive characteristics at baseline across entire sample and individual treatment arms

	Full Sample	Treatment Arm	Between Groups comparisons

	N	BEST with sign	BEST without sign	TAU	
		N	N	N	
N	48 -	17	17	14	-
Sex M(F)	28(20)				p = 0.039[1]*
Male	28	14	7	7	
Female	20	3	10	7	

[1] Chi square used for group comparisons indicates significant difference*

Table 3: Continuous descriptive characteristics at baseline across entire sample and individual treatment arms

	Full Sample	Treatment Arm			Between Groups comparisons
	M(SD)	BEST with sign	BEST without sign	TAU	
		M(SD)	M(SD)	M(SD)	
Age in months at entry	46.96 (6.13)	46.65(6.99)	47.71(5.67)	46.43(5.92)	p = 0.951[1]
IDACI	0.26 (0.17)	0.31(0.14)	0.19(0.16)	0.29(0.19)	p = 0.647[1]
CSCOT	0.49 (0.16)	0.52(0.18)	0.54(0.12)	0.40(0.14)	p = 0.048[1]*
Dosage[a]	9.52 (6.47)	13.65(2.47)	13.24(2.22)	-	p = 0.937[2]

*[1] ANOVA used for group comparisons.; [2]Chi square used for group comparisons; [a] Average number of sessions received out of 16. * indicates significant difference*

Table 4 shows baseline descriptive statistics and comparisons for all measures and treatment arms. Between groups ANOVAs revealed no significant differences for the treatment arms on any measures at baseline.

Table 4: Descriptive statistics for measures at each timepoint

	BEST with Sign		BEST without Sign		Treatment as Usual		Baseline Comparisons (ANOVA)
	M	SD	M	SD	M	SD	
Primary Outcomes							
NRDLS Production							
Baseline	76.36	7.91	79.24	10.65	76.94	6.49	(F(2, 45) = 0.51, p = 0.605)
Outcome	76.79	9.02	84.35	9.67	87.06	14.05	
Follow-up	80.07	11.38	83.94	10.59	89.75	13.64	
NRDLS Comprehension							
Baseline	81.21	7.97	81.00	6.84	82.41	7.49	(F(2, 45) = 0.18, p = 0.839)
Outcome	83.79	9.67	87.41	10.39	91.59	12.42	
Follow-up	87.71	13.8	86.88	9.49	93.44	12.91	
Secondary Outcomes							

BEST Content (%)							
Baseline	48.98	15.82	43.49	11.51	44.96	14.08	(F(2, 44) = 0.62, p = 0.541)
Outcome	49.85	14.07	78.43	18.37	77.86	16.54	
Follow-up	58.63	17.62	74.20	11.09	78.68	15.29	
BEST Morphology (%)							
Baseline	36.54	21.11	34.42	19.54	37.56	19.32	(F(2, 44) = 0.11, p = 0.900)
Outcome	48.83	22.46	69.97	24.13	74.53	22.06	
Follow-up	55.82	24.12	71.98	22.11	77.23	18.54	
Generalised Content (%)							
Baseline	40.48	16.30	32.22	14.06	36.00	13.76	(F(2, 43) = 1.15, p = 0.327)
Outcome	41.22	14.89	40.68	10.80	45.59	14.51	
Follow-up	40.33	17.59	49.78	14.31	54.04	18.19	
Generalised Morphology (%)							
Baseline	33.37	22.26	32.14	19.23	33.79	16.31	(F(2, 43) = 0.03, p = 0.969)
Outcome	46.80	22.65	53.81	23.03	57.59	25.35	
Follow-up	51.07	26.01	63.33	18.90	68.07	20.78	

Group comparisons

Longitudinal multilevel models were employed to determine the efficacy of BEST with sign and BEST without sign when compared to Treatment as Usual. The models are robust to small amounts of missing data and as such, incomplete cases could be included (Finch et al., 2019).

We intended to include functional communication as an outcome measure using the FOCUS (Focus on the Outcomes for Children Under Six) (Thomas-Stonell et al., 2012) questionnaire. Due to poor response rate for this measure (<40%), analysis of functional communication was not carried out. We had close to complete data for all direct measures (see appendix 3).

Analyses were therefore completed for the dependent variables displayed in table 5 at both outcome and follow-up:

Table 5: primary and secondary outcomes

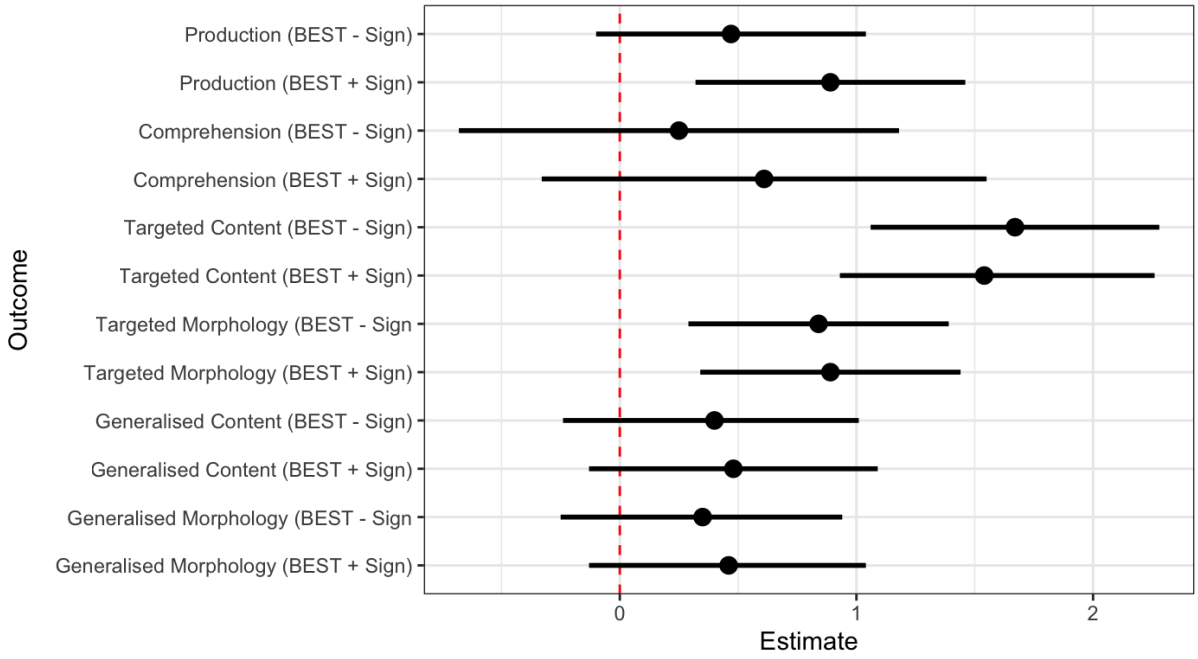
	Outcome measure	**Score type**
Primary outcomes	NRDLS production	Standard score
	NRDLS comprehension	Standard score
Secondary Outcomes	Targeted BEST content	Percentage correct
	Targeted BEST morphology	Percentage correct
	Generalised BEST content	Percentage correct
	Generalised BEST morphology	Percentage correct

In the maximal converging model score for each outcome measure, the following fixed effects were included: time point (baseline, outcome and follow-up), allocated intervention arm, interaction between timepoint and intervention arm, age, the number of sessions received, IDACI score and language profile (expressive, receptive or mixed difficulties) determined by scores on the NRDLS. This model was designed to include all theoretically relevant factors and to account for the clustering of repeated measures, as well as at the level of the individual child and the school.

Random intercepts for school were predicted by the following random slopes: timepoint, number of sessions received, and age in months at entry. Random intercepts for participant were predicted by random slopes for timepoint.

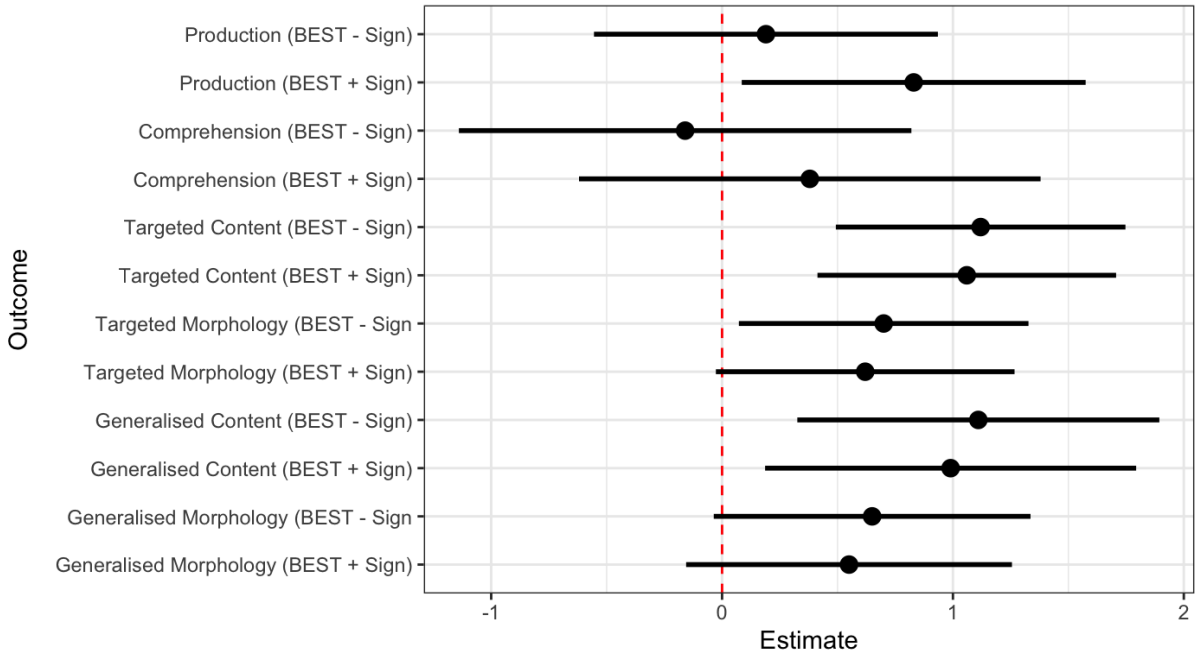
The following figures summarise the main findings at outcome (figure 3) and follow-up (figure 4). Results are summarised below (see appendix 5 for full model results). To address the research questions TAU was the reference condition, so all comparisons are between BEST (with or without sign) and TAU.

Figure 3*: Standardised effects by outcome measure and intervention arm at outcome*



Note. Scores are standardised and compared with TAU at outcome. Red dashed line depicts zero change. Lines that do not cross the dashed line represent statistically significant differences from TAU. BEST + Sign denotes BEST with sign; BEST - Sign denotes BEST without sign.

Figure 4*: Standardised effects by outcome measure and intervention arm at follow-up*



Note. Scores are standardised and compared with TAU at outcome. Red dashed line depicts zero change. Lines that do not cross the dashed line represent statistically significant differences from TAU. BEST + Sign denotes BEST with sign; BEST - Sign denotes BEST without sign.

NRDLS Production

There was a significant effect of BEST with Sign at outcome (β = 10.14, 95% CI [3.68, 16.60], $p < .01$) and follow up (β = 9.44, 95% CI [0.93, 17.94], $p < .05$) when compared to TAU. Large effect sizes were detected at both timepoints; ($d = 1.11$) and ($d = 1.45$) respectively. The effect of BEST without sign was not significant at outcome (β = 5.32, 95% CI [-1.12, 11.77], $p = .11$) or follow-up (β = 2.12, 95% CI [-6.31, 10.56], $p = .62$) when compared to TAU. The effect size at outcome was large ($d = 1.08$) whilst at follow-up there was a negligible effect size ($d = .18$).

NRDLS Comprehension

There were no significant effects of BEST with or without sign when compared to TAU. For BEST with sign at outcome and follow-up, medium effect sizes were detected ($d = .50$) and ($d = .50$) respectively. At outcome the effect size for BEST without Sign was small ($d = .35$) and at follow-up there was negligible negative effect ($d = -0.12$).

Targeted BEST Content

There was a significant effect of BEST with Sign at outcome (β= 32.09, 95% CI [19.41, 44.77], $p < .001$) and at follow-up (β = 21.95, 95% CI [8.58, 35.33], $p < .01$) when compared to TAU. Effect sizes for these measures were large; ($d = 1.78$) and ($d = 2.14$) respectively. At outcome the effect of BEST without Sign was significant (β = 34.72, 95% CI [22.12, 47.32], $p < .001$) and the effect size was very large ($d = 3.60$). There was also a significant effect of BEST without Sign at follow-up (β = 23.27, 95% CI [10.19, 36.35], $p < .001$) with a large effect ($d = 1.25$) when compared to TAU.

Targeted BEST Morphology

The effect of BEST with Sign was significant at outcome (β= 23.82, 95% CI [9.02, 38.61], $p < .01$) but not at follow-up (β = 16.57, 95% CI [-0.69, 33.84], $p = .06$), when compared to TAU. Large effect sizes were detected at both outcome and follow-up ($d = 1.13$) and ($d = 1.25$) respectively. The effect of BEST without Sign was significant at outcome (β = 22.42, 95% CI [7.67, 37.17], $p < .01$) and at follow-up (β = 18.85, 95% CI [2.04, 36.66], $p < .05$) when compared to TAU. The effect sizes were also large for each time-point ($d = 1.99$ and $d = 0.79$ respectively).

Generalised BEST Content

The effect of BEST with Sign was not significant at outcome (β= 7.59, 95% CI [-1.98, 17.15], $p = .12$) with a medium effect size ($d = .56$). There was a significant effect of BEST with Sign at follow-up (β = 15.52, 95% CI [2.93, 28.11], $p < .05$) with a large effect size ($d = 1.61$) when compared to TAU. The effect of BEST without Sign was not significant at outcome (β = 6.08, 95% CI [-3.72, 15.8;7], $p = .23$), although the effect size was large ($d = .81$). There was a significant effect of BEST without Sign at follow-up (β = 17.54, 95% CI [5.26, 29.81], $p < .01$) and a large effect ($d = 1.01$) when compared to TAU.

Generalised BEST Morphology

There were no significant effects of BEST with or without sign for morphology on the generalisation sentences at outcome or follow up when compared to TAU. BEST with Sign was not significant at outcome (β= 11.21, 95% CI [-3.18, 25.59], $p = .13$) or follow-up (β = 13.40, 95% CI [-3.82, 30.62], $p = .13$). A medium effect size was detected at outcome ($d = .55$), and a large at follow-up ($d = 1.02$). BEST without Sign was not significant at outcome (β = 8.51, 95% CI [-6.10, 23.13], $p = .26$) and the effect size was medium ($d = .79$). It was also not significant at follow-up (β = 15.93, 95% CI [-0.94, 32.79], $p = .07$) with a medium effect ($d = .66$).

Summary

Outcome: There were statistically significant differences with large effect sizes in favour of BEST with sign when compared to TAU on the following outcome measures: NRDLS production; targeted BEST content and targeted BEST morphology. There were no statistically significant differences between BEST with sign and TAU for the following: NRDLS comprehension, generalised BEST content, or generalised BEST morphology.
There were statistically significant differences with large effect sizes in favour of BEST without sign compared to TAU only for targeted BEST content and targeted BEST morphology. No significant differences were present for NRDLS production, NRDLS comprehension, generalised BEST content, or generalised BEST morphology.

Follow-up: Effects were maintained for NRDLS production and targeted BEST content. For targeted BEST morphology a significant effect was not maintained for BEST with sign, but was maintained for BEST without sign. Finally, for BEST with and without sign, generalised

BEST content was significantly better than TAU at follow-up, despite being non-significant at outcome for BEST with and without sign, indicating that within the six weeks follow-up period, children may have had time to generalise their learning from taught structures.

Discussion

This study sought to assess the efficacy of a usage-based intervention for children in the early-years presenting with low language abilities defined as falling ≤16th centile on the NRDLS. It also sought to assess the role of a signing system signalling content and morphology as an active ingredient of the intervention (McKean et al., 2013). The BEST intervention as initially developed, including signing of content and morphology, was significantly more effective than TAU on the primary outcome NRDLS production at outcome and follow up with large effect sizes. A large positive effect in this standardised primary outcome, which measures broad gains across the language system, which is corrected for age, and which is maintained after follow-up is particularly promising. It suggests the BEST intervention can accelerate language progress to a rate faster than expected for a child's age, generalise learning to non-targeted structures and that this catch-up rate of progress continues post intervention. In addition to these promising clinical implications, it also aligns with the hypothesis that BEST is not simply teaching constructions but changing the nature of the child's underlying abstract linguistic representations allowing learning of other new morphosyntactic elements to be achieved more readily (McKean et al., 2013) and accelerating learning via generalisation across novel, similar structures (Langacker, 2000; Tomasello, 2003).

The lack of an effect for NRDLS comprehension on a standardised outcome measure is disappointing and to some degree challenges this claim regarding changes in underlying linguistic representations. However, comprehension tests relying on behavioural responses such as NRDLS have been criticised as not being pure measures of comprehension but rather overlaid with confounds relating to attention, memory, and other executive functions. These issues are particularly important for preschool children and could mask underlying changes (Frizelle et al., 2017). However, this null finding also aligns with a number of other studies which suggest interventions are effective in the receptive domain only for vocabulary and for structures targeted in the intervention rather than generalising to wider receptive language abilities (e.g. Law et al., 2018). We must acknowledge issues of power in this study and so null findings must be considered with care. A non-significant trend in favour of BEST with sign is present, with a medium effect size (*d* = .50) at outcome and follow-up, but with wide confidence intervals. The results provide preliminary evidence of the efficacy of the BEST intervention for expressive language and indicate that further exploration of effects for production and comprehension with an increased sample size and randomisation procedure is warranted (Craig et al., 2008).

BEST without sign was not significantly different from TAU on the primary NRDLS outcomes but was effective in improving the vocabulary, sentence structures and morphology targeted in the intervention. Taken together with the significant effects for BEST with sign, this indicates that sign is an active ingredient of the BEST intervention. When designing BEST, McKean et al. (2013) hypothesised four potential benefits of sign: 1) content word signing would support verb semantic mapping and consequently development of predicate argument structure representations; 2) content word signing reduces overall processing load, making other cues more accessible to the child, accelerating learning of structures heard; 3) signing of morphology would promote mapping of the morphological frame, which in turn would support children to bootstrap semantic roles and develop abstract representations; 4) signing of morphology would enable children to use and understand grammatical markers. Although the results cannot clearly identify whether all or some of these are in play, NRDLS outcomes signify benefits to wider language learning. This indicates that the mechanisms described in hypotheses 1 and 3 may be supported as they suggest positive effects on underlying linguistic representations, particularly with regards to predicate argument structure and the development of abstract representations (Chiat, 2001; Tomasello, 2003; van Berkel-van Hoof et al., 2019; Vogt & Kauschke, 2017).

Consideration of the pattern of results in the secondary outcomes (Targeted BEST; Generalised BEST) offer additional insights. Here children in both BEST with and without sign treatment arms progressed in the Targeted BEST language structures (sentence structures, vocabulary, and morphology) better than TAU. This perhaps suggests hypotheses 2 and 4 are not supported and children can learn the specific vocabulary, sentence structures and morphology equally well with and without sign and sign acting to support the processes of schematisation and analogy (see appendix 1). Further research tracking rate of learning and with a larger sample would be needed to test this definitively.

There is certainly no evidence to suggest sign within an oral language intervention impedes children's expressive language learning and so parents can be reassured of the absence of harm and the presence of benefit (Abbott & Lucey, 2005).

The pattern of BEST assessment and generalised BEST assessment scores at outcome and follow-up support the suggestion made earlier with respect to NRDLS outcomes, that changes to the child's linguistic knowledge and representations may continue after the intervention. The generalised BEST assessment, which tests the sentence structures and morphology targeted in BEST but with differing vocabulary items, found better content scores than TAU for both BEST with and without sign at follow up but not at outcome. Content scores do not simply represent larger vocabulary knowledge but more consistent use of all arguments within a sentence's predicate argument structure (see appendix 4). Although weaker evidence than the NRDLS follow up, these finding add support to claims of changes to underlying representations (McKean et al., 2013) and catalysing of language change (Langacker, 2000).

Strengths and limitations

This study is situated in the piloting and feasibility phases of the MRC guidance for evaluating complex interventions (Craig et al., 2008; Skivington et al., 2021).
The study was conducted with a high level of methodological rigour relevant to these stages of efficacy research including intervention delivery training for the first author and assessment training for research assistants, blinding of assessors and treatment fidelity checks.

The sample size, although appropriate for the preliminary stage in the research process, is a limitation for the statistical analysis due to low statistical power. Future evaluation of BEST should increase the statistical power to increase the robustness and generalisability of results. Effect sizes reported here can inform sample size calculations for a fully powered trial. Likewise, there is a degree of homogeneity within the sample, for example children are monolingual English speakers and come from the same geographical region. Future trials should also assess the effectiveness of BEST for different populations. There is an inherent risk of regression to the mean with repeated measures intervention studies which must be considered, however the carefully controlled design aimed to mitigate this risk insofar as possible. That is, in matched group studies and RCTs, as long as children are similar at baseline and/or this is statistically controlled then the probability of regression to the mean is the same in both the control and the treatment groups. Hence group differences rather than overall progress is the key test as to whether the intervention is effective.

Differences in classroom environment at baseline as measured by the CSCOT present a possible confound. Average scores for TAU were lower than those of BEST with and without sign. It must be noted that the CSCOT was designed as a tool for teachers and school staff to evaluate and improve the classroom oral language environment and was not designed or recommended as a research tool for comparing different classrooms (Dockrell et al., 2012). The use of the CSCOT as a tool for matching was therefore a novel application in the present study. Scores were systematically lower at each wave suggesting that experience with the tool resulted in more stringent scoring calling the validity of the scores into question. Whilst it is not possible to draw conclusions on the validity of the CSCOT as a matching tool, matching with respect to IDACI scores was successful with no statistically significant differences across schools in differing treatment arms or data waves.

Conclusions

The study suggests that BEST, an intervention underpinned by usage-based theory is effective for supporting expressive language development of preschool children with language difficulties, and that it can accelerate language learning more broadly. Caution is required due to the scale of the study. The next stage of efficacy research, a fully powered randomised controlled trial informed by the findings of this study is therefore warranted. The results also suggest that a signing system that represents both vocabulary and predicate argument structure is an active ingredient of the BEST intervention. The use of sign clearly

supports expressive language development and this may occur through the highlighting of morphological 'frames' into which 'slots' can be placed, supporting schematization and categorisation, and enabling analogy across sentences to abstract predicate argument structure representations across verbs (Tomasello, 2003). Although clearly not a definitive test of usage-based theory, our findings align with these explanatory models of language acquisition. Importantly, these findings suggest that more research is warranted to explore their potential to inform intervention development; specifically the ways in which the language input manipulation promotes the use of cognitive mechanisms to support learning of target structures and change across the linguistic system.

References

Abbott, C., & Lucey, H. (2005). Symbol communication in special schools in England: The current position and some key issues. *British Journal of Special Education*, *32*(4), 196–201. https://doi.org/10.1111/j.1467-8578.2005.00397.x

Ambridge, B., Theakston, A. L., Lieven, E. V. M., & Tomasello, M. (2006). The distributed learning effect for children's acquisition of an abstract syntactic construction. *Cognitive Development*. https://doi.org/10.1016/j.cogdev.2005.09.003

Bercow, J. (2018). *Bercow: Ten Years On An independent review of provision for children and young people with speech, language and communication needs in England*.

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & Consortium, T. C.-2. (2017). Phase 2 of CATALISE: a multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *The Journal of Child Psychology and Psychiatry*. https://doi.org/10.1111/jcpp.12721

Botting, N., Riches, N., Gaynor, M., & Morgan, G. (2010). Gesture production and comprehension in children with specific language impairment. *British Journal of Developmental Psychology*, *28*(1), 51–69. https://doi.org/10.1348/026151009X482642

Capone, N. C., & Mcgregor, K. K. (2004). Gesture Development: A Review for Clinical and Research Practices. *Journal of Speech Language and Hearing Research*, *47*, 173–186.

Chiat, S. (2001). Mapping theories of developmental language impairment: Premises, predictions and evidence. *Language and Cognitive Processes*. https://doi.org/10.1080/01690960042000012

Childers, J. B., & Tomasello, M. (2001). The role of pronouns in young children's acquisition of the English transitive construction. *Developmental Psychology*. https://doi.org/10.1037/0012-1649.37.6.739

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., Petticrew, M., Health, P., Unit, S., Michie, S., Nazareth, I., & Petticrew, M. (2008). Developing and evaluating complex interventions: new guidance. In *Medical Reseach Council*. https://doi.org/10.1136/bmj.a1655

Croft, W., & Cruse, D. A. (2004). *Cognitive linguistics* (D. A. Cruse & D. A. Cruse, Eds.). Cambridge, U.K. : Cambridge University Press.

Department of Communities and Local Government. (2015). The English Indices of Deprivation 2015: Research Report. In *DCLG*.

Dockrell, Bakopoulou, I., Law, J., & Spencer, S. (2012). *Developing a communication supporting classrooms observation tool*.

Edwards, S., Letts, C., & Sinka, I. (2011). *The New Reynell Developmental Language Scales*. GL Assessment.

Eldridge, S. M., Chan, C. L., Campbell, M. J., Bond, C. M., Hopewell, S., Thabane, L., Lancaster, G. A., Altman, D., Bretz, F., Campbell, M., Cobo, E., Craig, P., Davidson, P., Groves, T., Gumedze,

F., Hewison, J., Hirst, A., Hoddinott, P., Lamb, S. E., … Tugwell, P. (2016). CONSORT 2010 statement: Extension to randomised pilot and feasibility trials. *BMJ*. https://doi.org/10.1136/bmj.i5239

Finch, W. H., Bolin, J. E., & Kelley, K. (2019). *Multilevel Modeling Using R*. Chapman and Hall/CRC Press.

Frizelle, P., & McKean, C. (2022). Using Theory to Drive Intervention Efficacy: The Role of Dose Form in Interventions for Children with DLD. *Children*, *9*(6). https://doi.org/10.3390/children9060859

Frizelle, P., O'Neill, C., & Bishop, D. V. M. (2017). Assessing understanding of relative clauses: A comparison of multiple-choice comprehension versus sentence repetition. In *Journal of Child Language* (Vol. 44, Issue 6, pp. 1435–1457). Cambridge University Press. https://doi.org/10.1017/S0305000916000635

Goldin-Meadow, S. (2011). Learning through gesture. In *Wiley Interdisciplinary Reviews: Cognitive Science* (Vol. 2, Issue 6, pp. 595–607). https://doi.org/10.1002/wcs.132

Gomez, R. (2002). *Variability and Detection of Invariant Structure*. *13*(5), 431–436. https://doi.org/10.1007/s10162-016-0594-4

Harris, A. D., McGregor, J. C., Perencevich, E. N., Furuno, J. P., Zhu, J., Peterson, D. E., & Finkelstein, J. (2006). The use and interpretation of quasi-experimental studies in medical informatics. *Journal of the American Medical Informatics Association*, *13*(1), 16–23. https://doi.org/10.1197/jamia.M1749

Hirsh-Pasek, K., Golinkoff, R. M., & Hollich, G. (2000). An Emergentist Coalition Model for Word Learning: Mapping Words to Objects Is a Product of the Interaction of Multiple Cues. *Becoming a Word Learner: A Debate on Lexical Acquisition*, 136–164. https://doi.org/10.1093/acprof:oso/9780195130324.003.006

Langacker, R. W. (2000). A dynamic usage-based model. In *Usage-based models of language*.

Law, J., Charlton, J., Dockrell, J., Gascoigne, J., McKean, M., & Theakston, C. (2017). *Early Language Development: Needs, provision and intervention for preschool children from socio-economically disadvantage backgrounds Early Language Development: Needs, provision, and intervention for preschool children from socio-economically disadvantag*.

Law, J., Charlton, J., Mckean, C., Beyer, F., Fernandez-Garcia, C., Mashayekhi, A., & Rush, R. (2018). *Parent-child reading to improve language development and school readiness: A systematic review and meta-analysis*. www.nuffieldfoundation.org

Leonard, L. B. (2007). Processing limitations and the grammatical profile of children with specific language impairment. *Advances in Child Development and Behavior*, *35*, 139.

Leonard, L. B., Deevy, P., Karpicke, J. D., Christ, S. L., & Kueser, J. B. (2020). After initial retrieval practice, more retrieval produces better retention than more study in the word learning of children with developmental language disorder. *Journal of Speech, Language, and Hearing Research*, *63*(8), 2763–2776. https://doi.org/10.1044/2020_JSLHR-20-00105

LIVELY Research Group. (2019). *Language Intervention in the Early Years*. www.research.ncl.ac.uk/lively/interventions/best/

Lüke, C., Ritterfeld, U., Grimminger, A., Rohlfing, K. J., & Liszkowski, U. (2020). Integrated Communication System: Gesture and Language Acquisition in Typically Developing Children and Children With LD and DLD. *Frontiers in Psychology*, *11*(February), 1–13. https://doi.org/10.3389/fpsyg.2020.00118

May, C. R., Cummings, A., Girling, M., Bracher, M., Mair, F. S., May, C. M., Murray, E., Myall, M., Rapley, T., & Finch, T. (2018). Using Normalization Process Theory in feasibility studies and process evaluations of complex healthcare interventions: a systematic review. *Implementation Science*, *13*(1), 1–42.

McKean, C., Stow, C., & Pert, S. (2013). *Building Early Sentences Therapy: The BEST Manual*. Newcastle University.

McKean, C., Wraith, D., Eadie, P., Cook, F., Mensah, F., & Reilly, S. (2017). Subgroups in language trajectories from 4 to 11 years: the nature and predictors of stable, improving and decreasing language trajectory groups. *Journal of Child Psychology and Psychiatry*, *58*(10), 1081–1091. https://doi.org/10.1111/jcpp.12790

Morrison, C. M., Hirsh, K. W., & Duggan, G. B. (2003). Age of acquisition, ageing, and verb production: Normative and experimental data. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*. https://doi.org/10.1080/02724980244000594

Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., Vamvakas, G., & Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: evidence from a population study. *The Journal of Child Psychology and Psychiatry*, *57*(11), 1247–1257. https://doi.org/10.1111/jcpp.12573

Piaget, J. (1952). *The origins of intelligence in children*. International Universities Press.

R Core Team. (2019). *R: A language and environment for statistical computing* (1.2.5001). R Foundation for Statistical COmputing.

Reilly, S., McKean, C., & Levickis, P. (2014). Late talking: can it predict later language difficulties? *Centre for Research Excellence in Child Language, Research Snapshot 2*.

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

Riches, N. G., Tomasello, M., & Conti-Ramsden, G. (2005). Verb Learning in Children With SLI: Frequency and Spacing Effects. *Journal of Speech, Language & Hearing Research*. https://doi.org/10.1044/1092-4388(2005/097)

Roulstone, S., Wren, Y., Bakopoulou, I., Goodlad, S., & Lindsay, G. (2012). Exploring Interventions For Children And Young People With Speech, Language And Communication Needs: A Study Of Practice. *Better Communication Research Programme*, 86.

Rowe, J. (1981). The Paget-Gorman Sign System. *British Journal of Special Education*, *8*(4), 25–27.

Skivington, K., Matthews, L., Simpson, S. A., Craig, P., Baird, J., Blazeby, J. M., Boyd, K. A., Craig, N., French, D. P., McIntosh, E., Petticrew, M., Rycroft-Malone, J., White, M., & Moore, L. (2021). A new framework for developing and evaluating complex interventions: Update of Medical Research Council guidance. *The BMJ*, *374*(2018), 1–11. https://doi.org/10.1136/bmj.n2061

Thomas-Stonell, N Robertson, B., Walker, J., Oddson, B., & Washington, K Rosenbaum, P. (2012). *FOCUS©: Focus on the Outcomes of Communication Under Six*. Holland Bloorview Kids Rehabilitation Hospital.

Tomasello, M. (2003). Constructing a language. In *A usage based theory of language acquisition*. https://doi.org/10.1017/CBO9781107415324.004

van Berkel-van Hoof, L., Hermans, D., Knoors, H., & Verhoeven, L. (2019). Effects of Signs on Word Learning by Children With Developmental Language Disorder. *Journal of Speech, Language, and Hearing Research*, *62*(6), 1798–1812.

Van Horne, A. J. O., Curran, M., Larson, C., & Fey, M. E. (2018). Effects of a complexity-based approach on generalization of past tense –ed and related morphemes. *Language, Speech, and Hearing Services in Schools*, *49*(3S), 681–693. https://doi.org/10.1044/2018_LSHSS-STLT1-17-0142

Vogt, S. S., & Kauschke, C. (2017). With some help from others' hands: Iconic gesture helps semantic learning in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *60*(11), 3213–3225. https://doi.org/10.1044/2017_JSLHR-L-17-0004

Walker, M., & Armfield, A. (1981). What is the Makaton vocabulary? *Special Education: Forward Trends*, *8*(3), 19—20. http://europepmc.org/abstract/MED/6458105

Appendix 1: Stages of multi-word utterance development, relevant cognitive mechanisms, and their use as active ingredients in BEST (McKean et al., 2013; Tomasello, 2003)

Stage	Relevant cognitive mechanisms		
	Name	Description	Relevant BEST 'active ingredient'
1. Frozen Phases Rote learned, and therefore inflexible, utterances paired with a pragmatic function and a communicative context/cultural routine. The child cannot combine the elements of the structure productively with other words. e.g. "*eat it*" bound to a meal time social routine or "*shoes off*" bound to an undressing routine	*Intention reading*	To create the form-function mapping required for the development of frozen phrases, the child must 'read' the communicative intentions of the person from whom they are learning the phrase. The child's ability to read the intentions of others within the scaffolding of joint attentional frames (Tomasello, 2003 p.21).	A structured and repetitive 'joint action routine' is established, creating a joint attentional frame between the child and the adult, which 'scaffolds' the child's ability to infer the communicative intention of the utterances they hear (i.e. describing an event within a play activity). Hence the child quickly becomes able to infer the communicative intentions of the adult at the level of the attentional frame (which objects and actions are we both attending to and what is the global purpose of this joint attention); and so is supported to make such inferences at the level of the individual communicative acts within that frame (which objects and actions is the adult referring to with a specific utterance) (Tomasello, 2003).
	Cultural learning	A process by which young children learn through imitation (and later through instructed and collaborative learning) of others in their social group (Tomasello, 2003 p.290). The child must not only mirror the communicative behaviour of the adult, but also understand that the roles within the triadic attentional frame (e.g. adult - child - object/action) reverse when they imitate the adult. Through this process, the child comprehends that when they are the speaker, imitating the communication of the adult, the communicative intention of the adult that was directed to the child instead becomes directed to the adult (Tomasello, 2003 p.26).	Role reversal is used within the 'joint action routine' to promote cultural learning and hence the creation of symbolic linguistic representations (Tomasello, 2003).

Stage	Relevant cognitive mechanisms		
	Name	Description	Relevant BEST 'active ingredient'
2. *Lexically specific constructions* Partially productive/flexible utterances with a 'slot and frame' construction where only one element can vary (e.g. "*X fall down*" or "*I'm ACTIONing it*"). 'Frame' categories might be: "X fall down" – the category of events in which animate or inanimate objects unintentionally drop to a lower place "I'm ACTIONing it" – the category of events in which the child is performing an action on an object. 'Slot' categories might be: X fall down – where X is the category of animate or inanimate objects which can fall I'm ACTIONing it – where ACTION is the category of the things I can do to objects Eat X – where X is the category of objects which can be eaten.	*Schemat-isation*	A general cognitive strategy that facilitates the identification of rules and patterns or *schemas*, or within the child's environment, supporting them to rely on mental abstractions (Piaget, 1952). In the case of communication, multiple exposures to the same utterances where one component is varied across exposures (e.g. X fall down, where X is *the girl* and then *the boy* and then *the teddy)* enable children to create rules or *schemas* which represent the aspects of the construction that remain the same across iterations (the 'frame'), and which components vary across iterations (the 'slot', in this case X) (Gomez, 2002; Tomasello, 2003 p.122).	The cognitive processes of *schematisation* and *categorisation* both depend on the quantity and distribution of types and tokens within the input heard by the child. BEST provides multiple presentations of highly similar exemplar sentences in which one element is varied systematically (Gomez, 2002; Tomasello, 2003). (e.g. The baby is laughing; The woman is laughing; The girl is laughing; The teddy is laughing).
	Categori-sation	In order to effectively use the 'slot and frame' constructions emerging from the process of schematisation, children must also form mental *categories* of which items can be put into each 'slot'. At the stage of *lexically specific constructions*, the child's categories are still functional and relatively concrete (e.g. in the construction "*X fall down*, X might consist of 'animate objects which involuntarily move from a high place to a lower place', and the category ACTION in "*He's ACTIONing it*" would consist of 'actions 'he' can perform' (Tomasello, 2003 p.124).	
3. *Abstract constructions*	*Analogy*	Children identify patterns and commonalities between phenomena, in the case of linguistic	Repeated exposure to sentence construction pairs that have the same predicate argument structure but contrasting verbs (e.g.

Stage	Relevant cognitive mechanisms		
	Name	Description	Relevant BEST 'active ingredient'
Flexible, abstract representations allowing children to use any relevant lexical items in the appropriate role in the sentence and so use the sentence structures productively. Analogy across functional relationships supports the creation of semantic categories (e.g. AGENT, PATIENT) and constructions (e.g. AGENT + ACTION + PATIENT) Analogy across construction form (but not function) (e.g. ***The*** *girl like**s** cake;* ***The*** *rabbit eat**s** lettuce)* results in the construction of syntactic categories (e.g. VERB, OBJECT) and constructions (e.g. SUBJECT + VERB + OBJECT).		abstract constructions, the shared *functional* relationships between items. Such categories are analogous because the functional relationships are the same across constructions, e.g. *'The **A** is **B**ing the **C**'* is analogous to 'The ***D*** *is **E**ing the **F**'* (Tomasello, 2003 p163). In the above example, **A** and **D** are doing the action, **B** and **E** are actions, and **C** and **F** are the recipients of the actions. "When an analogy is made the objects involved are effaced; the only identity they retain is their relational structure" (Tomasello, 2003 p.164).	The teddy is eating the apple; The man is washing the apple) provides children with multiple opportunities to identify the similarities in functional relationships and abstract semantic *categories* (e.g. AGENT, ACTION, PATIENT) and semantic *constructions* (e.g. AGENT + ACTION + PATIENT) (Tomasello, 2003). For each sentence construction pairing the items in each argument structure role are non-overlapping sets, providing a level of consistency thought to facilitate analogy (McKean et al., 2013) (e.g. AGENTS are never PATIENTS and vice versa). The use of toys to act out the target sentences support the identification of predicate argument structure roles (e.g. making distinctions between agent and patient more tangible). Input rotates through the different constructions targeted by BEST. This results in distributed exposure to a range of constructions across which the child can find analogies (Ambridge et al., 2006). For each sentence construction pairing the morphological frame remains constant (e.g. **The** boy **is** jumping; **The** woman **is** sitt**ing; The** X **is** Y**ing the** Z) providing an additional structural cue regarding the similarity between constructions (Tomasello, 2003). The use of a signing system which marks both lexical items and grammatical morphology. The marking of these items drives pattern finding and thence analogy, supporting children with language difficulties to create abstract representations of predicate argument structure that might otherwise be difficult due to phonological and morphological processing difficulties. Sign also supports semantic mapping and reduces processing

Stage	Relevant cognitive mechanisms		
	Name	Description	Relevant BEST 'active ingredient'
			load, rendering cues in the input more accessible (Chiat, 2001; Leonard, 2007; Tomasello, 2003; Rowe, 1981; Walker & Armfield, 1981).
All stages	*Mapping*	Establishing a representation in memory of a new meaning-construction pairing which is essential for learning words and early multi-word constructions and their corresponding meanings (Hirsh-Pasek et al., 2000).	Many repetitions of the same and similar constructions are provided alongside visual referents (toys and signs) to facilitate mapping which often requires more exposures for children with language difficulties than their typically developing peers. Other verbal input is avoided (Riches et al., 2005).
All stages	*Retention*	The formation of robust representations of newly learned constructions in long-term memory for future retrieval (Leonard et al., 2020).	Exposure to constructions is distributed over multiple sessions to leverage spacing effects thought to facilitate long term retention of learning (Riches et al., 2005). Multiple opportunities to use the target construction expressively, facilitating long term retention (Frizelle & McKean, 2022).

Appendix 2: Examples of BEST input and output sentences

Predicate Argument Structure		Input		Output
	1	The baby is laughing	1	The man is **sitting**
SET A:	2	The woman is laughing	2	The woman is **sitting**
	3	The boy is laughing	3	The boy is **sitting**
Agent + Action	4	The girl is laughing	4	The girl is **sitting**
	5	The teddy is laughing	5	The teddy is **sitting**
	6	The man* is laughing	6	The baby is **sitting**
	1	The man is eating an apple	1	The **teddy** is **washing** an apple
SET C:	2	The man is eating an orange	2	The **teddy** is **washing** an orange
	3	The man is eating a banana	3	The **teddy** is **washing** a banana
Agent + Action + Patient	4	The man is eating a carrot	4	The **teddy** is **washing** a carrot
	5	The man is eating a lolly	5	The **teddy** is **washing** a spoon
	6	The teddy* is eating a banana	6	The **teddy** is **washing** a cup
SET G:	1	The baby is putting a spoon on the table	1	The **man** is putting a spoon on the **bed**
Agent + Action + Theme + Locative	2	The baby is putting a cup on the table	2	The **man** is putting a cup on the **bed**
	3	The baby is putting a flower on the table	3	The **man** is putting a flower on the **bed**
	4	The baby is putting a key on the table	4	The **man** is putting a key on the **bed**
	5	The man* is putting a phone on the table	5	The **man** is putting a phone on the **bed**
SET H:	1	The man is pouring milk into a cup	1	The **baby** is pouring **juice** into a cup
Agent + Action + Theme + Locative	2	The man is pouring milk into a shoe	2	The **baby** is pouring **juice** into a shoe
	3	The baby* is pouring milk into a box	3	The **baby** is pouring **juice** into a box
SET I:	1	The man is putting a spoon on the bed	1	The **baby** is **pouring** juice in the shoe
Agent + Action + Theme + Locative	2	The man is putting a cup on the bed	2	The **baby** is **pouring** juice in the box
	3	The man is putting a flower on the bed	3	The **baby** is **pouring** juice in the cup
	4	The man is putting a key on the bed		
	5	The baby* is putting a phone on the bed		

Appendix 3 shows assessment completion rate for each measure at each timepoint.

Appendix 3: Percentage of completed assessments at each timepoint

Outcome Measure	Baseline N[a] (%)	Outcome N (%)	Follow up N (%)
NRDLS (production and comprehension)	48 (100%)	48 (100%)	47 (98%)
Targeted BEST Assessment (Content and Morphology)	47 (98%)	47 (98%)	43 (90%)
Generalised BEST Assessment (Content and Morphology)	46 (96%)	48 (100%)	43 (90%)
Teacher FOCUS	26 (54%)	17 (35%)	43 (90%)
Parent FOCUS	24 (50%)	10 (21%)	9 (19%)

[a]Number of assessments completed at each timepoint out of al possible 48 (one per child).

Appendix 4: BEST Assessment Scoresheet

BEST Assessment and Therapy Recording Form

Show the child one picture at a time from the **BEST Assessment Picture Booklet: Pre-Treatment Assessment**and ask the child "Tell me what's happening here?" Record the child's first response only. Do not try and elicit a full response.

No.	Verb	Target	Child's response	Scoring Content (PAS)	☒☑	Scoring morphology	☒☑
1	sit	The teddy is sitting		Agent (X)		Det (the)	
				Verb (sit)		Aux(is)	
						Inflection(ing)	
2	walk	The boy is walking		Agent (X)		Det (the)	
				Verb (walk)		Aux(is)	
						Inflection(ing)	
3	laugh	The woman / lady islaughing		Agent (X)		Det (the)	
				Verb (laugh)		Aux(is)	
						Inflection(ing)	
4	jump	The woman / lady isjumping		Agent (X)		Det (the)	
				Verb (jump)		Aux(is)	
						Inflection(ing)	
5	ride	The boy is riding the/abike.		Agent (X)		Det 1(the X)	
				Verb (ride)		Aux (is)	
				Patient (Y)		Inflection (ing)	
						Det 2 (the/a Y)	
6	smell	The baby is smellingthe/a sock.		Agent (X)		Det 1(the X)	
				Verb (smell)		Aux (is)	
				Patient (Y)		Inflection (ing)	
						Det 2 (the/a Y)	
7	wash	The teddy is washingthe/a banana		Agent (X)		Det 1(the X)	
				Verb (wash)		Aux (is)	
				Patient (Y)		Inflection (ing)	
						Det 2 (the/a Y)	

No.	Verb	Target	Child's response	Scoring Content (PAS)	☒☑	Scoring morphology	☒☑
8	eat	The teddy is eatingthe/a banana		Agent (X)		Det 1(the X)	
				Verb (eat)		Aux (is)	
				Patient (Y)		Inflection (ing)	
						Det 2 (the/a Y)	
9	kick	The woman / lady iskicking the/a apple.		Agent (X)		Det 1(the X)	
				Verb (kick)		Aux (is)	
				Patient (Y)		Inflection (ing)	
						Det 2 (the/a Y)	
10	brush	The boy is brushingthe/a cat		Agent (X)		Det 1(the X)	
				Verb (brush)		Aux (is)	
				Patient (Y)		Inflection (ing)	
						Det 2 (the/a Y)	
11	hug	The girl is hugging the/ateddy		Agent (X)		Det 1(the X)	
				Verb (hug)		Aux (is)	
				Patient (Y)		Inflection (ing)	
						Det 2 (the/a Y)	
12	kiss	The boy is kissing the/ahorse		Agent (X)		Det 1(the X)	
				Verb (kiss)		Aux (is)	
				Patient (Y)		Inflection (ing)	
						Det 2 (the/a Y)	
13	put	The man is putting the/aspoon on the/a bed		Agent (X)		Det 1 (the X)	
				Verb (put)		Aux (is)	
				Patient (Y)		Inflection (ing)	
				Locative (Z)		Det 2 (the/a Y)	
						Prep (on)	
						Det 3 (the/a Z)	
14	pour	The baby is pouring juice into the/a box		Agent (X)		Det 1 (the X)	
				Verb (pour)		Aux (is)	

No.	Verb	Target	Child's response	Scoring Content (PAS)	☒☑	Scoring morphology	☒☑
				Patient (Y)		Inflection (ing)	
				Locative (Z)		Prep (into)	
						Det 2 (the/a Z)	
15	give	The girl is giving the/a phone to the/a woman /lady		Agent (X)		Det 1 (the X)	
				Verb (give)		Aux (is)	
				Patient (Y)		Inflection (ing)	
				Benefactive (Z)		Det 2 (the/a Y)	
						Prep (to)	
						Det 3 (the/a Z)	
16	throw	The boy is throwing the/a banana to the/ababy		Agent (X)		Det 1 (the X)	
				Verb (throw)		Aux (is)	
				Patient (Y)		Inflection (ing)	
				Benefactive (Z)		Det 2 (the/a Y)	
						Prep (to)	
						Det 3 (the/a Z)	
			Please transfer these percentage scores to the table on page 3, and plotthem onto the **Progress Tracker Chart** on page 4.	**TOTAL CONTENT**	**/48**	**TOTAL MORPHOLOGY**	**/67**
				PERCENTAGE (Raw score/total score) x 100		**PERCENTAGE** (Raw score/total score) x 100	

Appendix 5: Multilevel Results for NRDLS Production and Comprehension, Targeted Content and Morphology, and Generalised Content and Morphology Outcomes (Beta [95% CIs])

	Maximal Multilevel Model					
	NRDLS production[a]	NRDLS comprehension[b]	Targeted BEST content[c]	Targeted BEST morphology[c]	Generalised BEST content[d]	Generalised BEST morphology[d]
Intercept	74.5***	77.4***	-12.5***	-19.4***	-3.0	-14.0***
	[71.6, 77.3]	[73.0, 81.8]	[-18.9, -6.2]	[-28.6, -10.3]	[-10.8, 4.9]	[-23.5, -4.6]
Outcome	0.5	4.9	0.5	12.3**	1.4	12.8**
	[-4.3, 5.2]	[-2.5, 12.3]	[-8.8, 9.8]	[1.5, 23.1]	[-5.7, 8.6]	[2.1, 23.5]
Follow-up	3.8	8.6**	8.8*	19.0***	-0.1	16.0**
	[-2.4, 9.9]	[0.9, 16.4]	[-0.7, 18.4]	[6.9, 31.2]	[-9.0, 8.7]	[3.7, 28.4]
BEST without sign at Baseline	-1.7	-2.9	-13.1	-4.0	-17.3	-11.8
	[-11.7, 8.3]	[-13.3, 7.5]	[-34.8, 8.7]	[-35.9, 27.9]	[-40.9, 6.3]	[-44.1, 20.5]
BEST with sign at Baseline	-0.5	-1.5	-9.0	6.6	-10.6	-2.8
	[-10.7, 9.7]	[-11.9, 9.0]	[-31.1, 13.2]	[-25.7, 38.9]	[-34.4, 13.2]	[-35.3, 29.8]
Age (centred)	-0.2*	0.1	0.8**	0.3	0.8*	0.4
	[-0.5, 0.0]	[-0.2, 0.5]	[0.2, 1.4]	[-0.9, 1.4]	[0.0, 1.5]	[-0.7, 1.5]
Dosage	0.2	0.2	0.3	-0.2	0.6	0.5
	[-0.6, 0.9]	[-0.5, 0.9]	[-1.2, 1.8]	[-2.5, 2.1]	[-1.0, 2.2]	[-1.8, 2.7]
IDACI (centred)	-5.9	-19.3***	-34.0***	-51.3***	-22.2*	-48.8***
	[-14.4, 2.5]	[-29.9, -8.7]	[-52.5, -15.6]	[-79.6, -22.9]	[-41.9, -2.6]	[-76.5, -21.2]
Expressive Profile at Baseline	0.5	10.4***	1.7	-10.2*	-1.3	-10.3*

	[-2.5, 3.5]	[7.85, 13.0]	[-4.8, 8.1]	[-20.4, -0.1]	[-8.6, 5.8]	[-20.6, -0.1]
Receptive Profile at Baseline	21.8***	-1.1	0.5	9.9	-0.8	5.5
	[17.1, 26.5]	[-5.4, 3.2]	[-9.7, 10.7]	[-5.9, 25.8]	[-12.0, 10.5]	[-10.3, 21.4]
BEST without sign at Outcome	**5.3**	**2.7**	**34.7*** **	**22.4*** **	**6.1**	**8.5**
	[-1.1, 11.8]	**[-7.4, 12.9]**	**[22.1, 47.3]**	**[7.7, 37.1]**	**[-3.7, 15.9]**	**[-6.1, 23.1]**
BEST without sign at Follow-up	**2.1**	**-1.8**	**23.3*** **	**18.9** **	**17.5*** **	**15.9***
	[-6.3, 10.6]	**[-12.4, 8.9]**	**[10.2, 36.4]**	**[2.0, 35.7]**	**[5.3, 29.8]**	**[-0.9, 32.8]**
BEST with sign at Outcome	**10.1*** **	**6.7**	**32.1*** **	**23.8*** **	**7.6**	**11.2**
	[3.7, 16.6]	**[-3.6, 16.9]**	**[19.4, 44.8]**	**[9.0, 38.6]**	**[-2.0, 17.2]**	**[-3.2, 25.6]**
BEST with sign at Follow-up	**9.4** **	**4.1**	**22.0*** **	**16.6***	**15.5** **	**13.4**
	[0.9, 17.9]	**[-6.7, 14.9]**	**[8.6, 35.3]**	**[-0.7, 33.8]**	**[2.9, 28.1]**	**[-3.8, 30.6]**
Observations	143	143	137	137	137	137
Log Likelihood	-472.0	-480.1	-537.2	-581.6	-535.7	-585.7
Akaike Inf. Crit.	1,016.1	1,032.2	1,146.5	1,235.2	1,143.5	1,243.4
Bayesian Inf. Crit.	1,122.7	1,138.8	1,251.6	1,340.3	1,248.6	1,348.5

*Note: *p<0.1; **p<0.05; ***p<0.001. Emboldened text displays results for the parameters of interest (interaction terms for intervention arm and timepoint)*

[a] *New Reynell Developmental Language Scales* (Edwards et al., 2011) *Production Subscale;* [b] *New Reynell Developmental Language Scales* (Edwards et al., 2011) *Comprehension Subscale;* [c] *Targeted BEST Assessment* (McKean et al., 2013)*;* [d] *Generalised BEST Assessment (Designed by the Author)*