



Department of Speech and Language Therapy

University of Limerick

Final Year Project

“Captain Grey and the Greedy Aliens”: Obtaining normative data on a new test for the production of verbs at sentence level in children

Submitted in partial fulfilment of the regulations for the reward of the Masters Degree in Speech and Language Therapy, University of Limerick

Proposed Journal of Submission:

International Journal of Language and Communication Disorders

Submitted by: Ellen Bridgeman (0737348)

Supervisor: Carol-Anne Murphy

Word Count: 8238

May 2014

Acknowledgements

I would firstly like to convey my gratitude to the participants and their parents and also all the schools involved for taking part in this study, without them this would not have been possible.

I would like to thank Carol-Anne Murphy for all of her support and guidance during the completion of this assignment. Her endless patience and willingness and availability to assist in answering even the smallest of queries were very much appreciated.

I wish to acknowledge the support of my family and family over both this project and my entire course of study and thank you all for keeping me sane even at the most stressful of times.

Contents

Abstract	1
1.0 Introduction.....	2
1.1 Typically Developing Children	2
1.1.1 Typically developing verb knowledge	2
1.1.2 Typically developing verb argument structure	3
1.1.3 Typically Developing Sentence Complexity.....	4
1.2 Children with Specific Language Impairment	4
1.2.1 Verb Knowledge and CwSLI.....	4
1.2.2 Verb argument structure and CwSLI	5
1.2.3 Sentence Complexity in Specific Language Impairment	5
1.3 Elicitation Techniques	6
1.4 Current Assessments.....	6
1.5 Current Study	7
2.0 Method.....	8
2.1 Ethics	8
2.2 Participants.....	8
2.3 The assessors.....	9
2.4 Design and Materials.....	9
2.5 Instrument Development.....	10
2.5.1 Captain Grey and the Greedy Aliens	10
2.5.2 Test of Verb Knowledge	10
2.6 Procedure	10
2.6.1 Test of Verb Knowledge	10
2.6.2 Captain Grey and the Greedy Aliens	11
2.6.3 Verb Error Analysis	12
2.7 Inter-rater Reliability	12
3.0 Results	14
3.1 Verb Argument Structure Production Accuracy.....	14
3.1.1 Verb argument structure accuracy scores according to Age Bands	15
3.1.2 Verb Knowledge as a predictor of Verb Argument Structure accuracy	16
3.2 Mean Sentence Complexity	17
3.2.1 MSC according to Age Bands.....	18
3.2.2 Verb Knowledge as a predictor of Mean Sentence Complexity.....	19

3.3 Standardisation of the Test Data	20
3.4 Error Analysis.....	21
3.5 Inter-rater reliability.....	21
4.0 Discussion.....	22
4.1 Verbs and Verb Argument Structure.....	22
4.2 Mean Sentence Complexity	23
4.3 Verb Error Analysis	24
4.4 Verb Knowledge as a predictor of Verb argument Production accuracy and Sentence Complexity.....	25
4.5 Test Related Behaviours.....	26
4.6 Considerations.....	26
4.7 Limitations	27
4.8 Recommendations	27
4.9 Clinical Implications and conclusions	27
Reference List.....	29
Appendix 1	
Appendix 2	
Appendix 3	

Abstract

Background: Children with specific language impairment (CwSLI) exhibit significant difficulties in language ability affecting areas such verb knowledge and verb production, producing fewer verbs and verb argument structures than typically developing (TD) children. As verbs are critical to sentence production it is also not surprising that they also produce fewer complex sentences. Although these areas of impairment are widely acknowledged there is a dearth of clinically available instruments that comprehensively assess verbs production and complex sentences in this population.

Aims: To gather normative data from typically developing children aged 4;06-12;00 on their sentence complexity and production of verbs and their related argument structures using the newly developed story retell tool 'Captain Grey and the Greedy Aliens'; to investigate the sensitivity of the assessment across the age ranges; to investigate the inter-rater reliability and clinical utility of the assessment; and finally to investigate the relationship between verb knowledge and verb argument structure accuracy and sentence complexity.

Method & Procedure: Normative data was gathered on 91 typically developing children aged 4;06-12;00 using the story retell assessment; Captain Grey and the Greedy Aliens. Participant's verb knowledge was also assessed via the Test of Verb Knowledge to allow for investigation of the relationship between verb knowledge and verb argument structure accuracy and sentence complexity.

Outcome & Results: The story retell assessment was sensitive to developmental progression across age groups for mean accuracy of verb argument structures production and mean sentence complexity and was found to be highly reliable for inter-rater reliability. Positive relationships were found between verb knowledge and verb argument structure accuracy and sentence complexity.

Conclusions & Implications: The story retell tool was shown to be effective in eliciting and analysing verb and verb argument structure production and sentence complexity. The data collected in this study can be used to provide normative data on this assessment however further sampling with larger sample sizes is warranted. Captain Grey and the Greedy Aliens may be a clinically useful tool.

1.0 Introduction

Children with specific language impairment (CwSLI) exhibit significant difficulties in language ability affecting areas such as; syntax, morphology, semantics and pragmatic skills which may persist into adolescence (Schwartz 2009; IASLT 2007; Leonard 2000). One area of particular difficulty is sentence production; CwSLI produce sentences that are frequently shorter and less complex than their typically developing (TD) age matched peers (Thordardottir and Weismer 2002; King and Fletcher 1993). There is increasing evidence that verbs are central in language development and sentence production, as verbs dictate sentence structure through obligatory arguments and assignment of thematic roles (Conti-Ramsden and Jones 1997; Tomasello 1992). Verb driven models of sentence production such as Garrett's model (Chiat *et al* 1997) recognise that both syntactic and semantic weaknesses may result in sentence production errors. Errors may arise at the semantic level due to inaccurate verb semantics or at the positional level due to inaccurate selection of syntactic forms. Research conducted on this area to date has resulted in mixed findings due to variety of assessment methods used (Owen Van Horne and Lin 2011; Thordardottir and Weismer 2002,).

1.1 Typically Developing Children

In order to identify areas of language difficulty for CwSLI it is important to be aware of typical development. This section will outline; the acquisition of verb knowledge, verb production and argument structures and the development of complex sentences of typically developing (TD) children.

1.1.1 Typically developing verb knowledge

The acquisition of verb semantics is more challenging than the acquisition of nouns as evident in early language productions which contain more nouns than verbs (Goldfield 2000; Fenson *et al* 1994). Verbs are more complex than nouns as they refer to actions and events which only last for a short space of time limiting the time a child has to associate the spoken word with the observable action (Rowland 2014). As children developed greater syntactic knowledge they are able to increase their verb knowledge through the use of syntactic bootstrapping. This allows children to use word order, function words and inflections to increase their semantic verb knowledge (Rowland 2014; McGregor 2009).

While some verbs have observable events or actions others do not which can impact on the saliency of their meaning and make acquiring them more difficult (Nils Johnston and

Wellman 1980). Examples of such verbs are cognitive state verbs which are more difficult to acquire as they are semantically abstract in nature and are complex at syntactic level as they often co-occur with complement clauses (Owen van Horne and Lin 2011; Nixon 2005; Diessel and Tomasello 2001). Although they are reportedly more difficult to acquire according to Nixon (2005) (n=40, 4;0-4;2) they appear at a relatively young age; all participants in the study produced at least one CSV in 15 minute language samples of child-caregiver interactions.

1.1.2 Typically developing verb argument structure

Verb argument structure has shown to be an area of continued development for TD children, with full acquisition of structures for complex verbs continuing into adolescence (Thordardottir and Weismer 2002; Naigles *et al* 1990).

Overgeneralizations of verb argument structures are common in TD children's productions. According to Bowerman (1974) children first overgeneralise the causative alternation (when change of state verbs can be used both intransitively and transitively) to fixed intransitive verbs at approximately 2;04 and continue to do so at 11 years of age. Overgeneralisations decrease with age as shown by Brooks *et al* (1999) in their study of 3 groups of TD children; Group 1 (3;01-3;07), Group 2 (4;06-5;07) and Group 3 (7:11-9;00). Similar findings were reported by Ambridge *et al* (2008) (n=93) in participants rating of grammatical acceptability of overgeneralised sentences using the transitive causative of real and novel verbs. Overgeneralizations are associated with frequency with low frequency verbs being overgeneralized more often (Ambridge *et al* 2008; Brooks *et al* 1999).

In order for children to accurately produce verb argument structures they must be able to accurately map the verb semantic properties to their syntactic structures (McKoon and McFarland 2000); inability can result in argument structure errors. Change-of-state verbs are typically achieved later than other verbs due to mapping errors (Pinker 2013; Gentner 1978). When change-of-state verbs included both motion and state change children are more likely to interpret the meaning as the motion change as it is more salient (Gentner 1978).

1.1.3 Typically Developing Sentence Complexity

Complex sentences are sentences which contain more than one verb phrase and are produced through the use of embedding and co-ordination (Paul and Norbury 2012; Bowerman 1979). According to Miller (1981) the emergence of complex sentences occurs as follows; children begin to use simple conjoining between 3;00-3;06 years, embedding with/without simple conjoining in the one sentence between 3;06-4;00, relative clauses appear between 5-6 years and complex conjoining for example 'because, so, when' appear around this time also. The emergence of complex sentences allows children to convey more complex information and increases their generative language ability (Fletcher 2009; Bowerman 1979).

1.2 Children with Specific Language Impairment

The following will examine the current knowledge available on language difficulties of CwSLI in relation to; verb knowledge, verb argument structure and sentence complexity.

1.2.1 Verb Knowledge and CwSLI

CwSLI have delayed receptive vocabulary (Leonard *et al* 1999; Thal *et al* 1999) which according to Windfuhr *et al* (2002) includes delayed acquisition of verb knowledge. Windfuhr *et al* (2002) (n=56) compared novel verb learning of CwSLI (4;04-5;10) to TD children (2;04-3;07) matched by the number of regular verbs in their lexicon. The participants were exposed to four novel verbs during naturalistic play sessions over a two week period (4 sessions, 40 exposures to each verb). CwSLI learned the novel verbs at a slower rate, requiring twice as many exposures than younger TD children before using them spontaneously. Riches *et al* (2005) and Rice *et al* (1994) both found that CwSLI had a faster rate of forgetting newly acquired novel verbs than TD language-matched children. Faster forgetting rates may be due to CwSLI having poor working memory impairing their ability to retain verb knowledge in their short term memory while developing the meaning and successfully mapping it into their long term memory (McGregor 2009).

Due to the apparent lexical verb weaknesses of CwSLI the semantics of more complex verbs such as cognitive state verbs (CSV) may pose a challenge (King & Fletcher 1993). However studies have shown that when matched on language no differences have been found between TD and CwSLI (Oven van Horne and Lin 2011; Johnston *et al* 2001). CwSLI do however continue to lag behind their age-matched peers according to Johnston *et al*

(2001) who examined CSV productions of CwSLI over a 4 year longitudinal study (aged 4;4 at time 1, n=26) and a cross sectional study (4-6 years).

1.2.2 Verb argument structure and CwSLI

Verb knowledge is a prerequisite to verb production therefore it is unsurprising given CwSLI's delayed verb knowledge that they present with verb argument structure difficulties (Goldfield 2000; King and Fletcher 1993). Murphy (2013) found that the incidence of verb argument structure errors and omissions was higher among a group of CwSLI (n=20, age range 6;05-10;01) than in TD children (n=92, age range 3;02-10;10) on sentence generation data. Ebbels 2005 (cited in Ebbels 2007) found obligatory argument omissions continue to be an area of difficulty for CwSLI into adolescence (11;00-14;11) with CwSLI omitting more arguments than age-matched and vocabulary-matched controls for ditransitive verbs. Ditransitive verbs require three verb arguments increase processing demands which may account for argument omissions due to CwSLI having poor working memory. Thordardottir and Weismer (2002) (n=11) found that CwSLI (5;05-9;08) used fewer verb argument structures and verb argument types than age-matched controls, similar findings were also reported in a study by Conti-Ramsden and Jones (1997).

As the accurate production of verb argument structure is dependent on the accurate mapping of verb semantic properties to syntactic structures (McKoon and McFarland 2000) lexical verb semantic deficits in CwSLI may result in a greater number of errors for complex verb arguments. Ebbels et al (2012), n=70, examined the ability of CwSLI (11;03-14;10) and TD children matched on; age, vocabulary and grammar to use change-of-state and change-of-location verbs which undergo the locative alteration. CwSLI did not differ from any other group on their production of change-of-location verbs however they performed significantly worse for change-of-state verbs than all controls. Similar to TD children CwSLI make more verb semantic errors if motion changes are more salient than state change (Ebbels *et al* 2012; Shulman and Gudeman 2007; Van der Lely 1994).

1.2.3 Sentence Complexity in Specific Language Impairment

The production of complex sentences is an area of difficulty for CwSLI which has been found to persist into adolescence (Levy and Friedmann 2009; Marinelle 2004; Scheule and Nicholls 2000). A study by Marinelle (2004) (n=30) found that CwSLI (mean age 10;08) produced fewer complex sentences, fewer combined complex sentences and less clauses per utterance than their TD age-matched peers in a 100 utterance conversational

language sample. A single subject longitudinal study of CwSLI from 3;03-7;10 conducted by Schuele and Dykes (2005) found that evidence of complex sentence productions did not emerge until 4;08 which is delayed according to Miller's (1981) stages of development. Although complex language forms were delayed they emerged according to the typical stages of emergence outlined by Miller (1981).

1.3 Elicitation Techniques

Method of elicitation can affect language output. While spontaneous speech samples may give a more realistic view of participant's everyday language use it may not provide a true representation of the participant's most complex language ability (Owen Van Horne and Lin 2011; Johnston *et al* 2001). According to Owen van Horne and Lin (2011) interview style conversations, story retells and personal narratives may provide longer utterances and greater complexity which in turn may elicit a greater number of errors. Story retell may be more effective than other methods such as spontaneous speech and story generation tasks for analysing verb and sentence production as stimulus input can be controlled to target specific verbs (Hesketh 2004; Gazella and Stockman 2003). However this elicitation method may be difficult for younger TD children as they have been found to produce shorter narratives, due to levels of attention, memory and retrieval (Gazella and Stockman 2003).

1.4 Current Assessments

Many currently used standardised language assessments include some verbs in their subtests, for example the Clinical Evaluation of Language Fundamentals 4th edition (CELF 4) (Semel *et al* 2006) and The New Reynell Developmental Language Scales 4th edition (NRDLS) (Edwards *et al* 2011). However none provide a comprehensive assessment of verbs and verb argument structures. The NRDLS, which includes an assessment of production of simple verb argument structures is standardised only up to children aged 7;06 years. Tests for verb argument production developed by Ebbels *et al* (2007) and Loeb *et al* (1996) assess a wide variety of verbs but are not widely available for clinical use.

'Captain Grey and the Greedy Aliens' (Murphy 2013) may fill this current gap in sentence and verb production assessments for children. Murphy (2013) previously trialled the assessment on a small number of TD children and CwSLI as a baseline for an intervention study. Since then additional scenes have been added to target cognitive state verbs. Loeb

et al (1996) have cited the need to include these verbs in assessment due to their importance in children repertoires.

1.5 Current Study

It is apparent that there is a need for clinicians to have a comprehensive, clinically useful assessment of verbs and their argument structures for CwSLI. The aim of the current study is to gather normative data for 'Captain Grey and the Greedy Aliens' from TD children aged 4;06-12;00 years on their mean length of sentence, production of verbs and their associated structures. Other objectives are to assess the sensitivity of the assessment tool across the age range, to consider clinical utility and to assess the inter-rater reliability of the assessment tool.

In light of the observed relationship between verb semantic knowledge in the literature, we also aim to investigate the relationship between stored verb knowledge, as determined using a verb comprehension assessment, and sentence production length and accuracy as determined from the Story Retell tool.

2.0 Methodology

2.1 Ethics

Ethical approval for this study was received from the Faculty of Education and Health Sciences Research Ethics Committee of the University of Limerick. All participants and their parents/guardians were provided with information regarding the study prior to it being carried out. Written parental consent was obtained for those who expressed interest in the study, written child consent was also obtained where possible, and only those who provided written parental consent were included in the study.

2.2 Participants

The proposed sample for this study was as follows; 85 typically developing children across the following age bands; 4;06-5;06 (n=20), 5;07-6;06 (n=15), 6;07-7;06 (n=15), 7;07-8;06 (n=15) and 8;07- 12;00 (n=20). These age bands were chosen as they represented the age groups of children across primary school years.

Participants were recruited from 4 primary schools in the South of Ireland from counties from both suburban and rural areas. The schools were selected for convenience with the principals of each school being known to either the primary researcher or student researchers. School principals were sent explanatory letters, with information sheets and consent forms.

Inclusion criteria for participants was as follows; English as their first language, no diagnosis of an intellectual disability, no previous or current speech and language therapy involvement or suspected speech and language difficulty, no hearing impairment.

The obtained sample from this recruitment process comprised of 93 children aged between 58 and 144 months. Two participants were omitted from the study, one due to suspected language difficulties and the other due to lack of responses to prompting. This resulted in a final sample of 91 children. The sample consisted of 38 males and 52 females. The participants were grouped according to the age bands (see Table 1). The age bands were revised and altered from the original proposed age bands as fewer participants were gathered for the (4;06-5;05) group and a larger than expected number of the (8;07-12;00) age group were recruited. The lower numbers in the (4;06-5;05) group arose due to the

time of data collection which began in the second school semester (i.e. Jan-May) meaning many children in Junior Infants had moved up this age band.

Age Bands	N	Mean Age (months) (SD)	Range (months)	Percentage of the sample
4;06-5;05	7	62.1 (2.4)	58-65	7.69%
5;06-6;05	17	71.6 (3.4)	65-77	18.68%
6;06-7;05	17	82.6 (4.2)	78-89	18.68%
7;06-8;05	9	94.6 (4.3)	90-101	9.89%
8;06-9;05	9	109.0 (4.2)	102-113	9.89%
9;06-10;05	12	119.3 (3.7)	114-125	13.18%
10;06-12;00	20	133.3 (6.0)	127-144	21.97%
Combined sample	91	98.7 (25.4)	58-144	100%

Table 1 Mean age and age range for participants according to age groups

2.3 The assessors

Five post-graduate speech and language therapy students carried out the assessments. Testers were given instruction on how to administer the assessments used in data collection including details on elicitation techniques, prompting, recording of responses and scoring of the assessments. Testers followed specific guidelines to ensure consistency and fidelity of administration.

2.4 Design and Materials

The study was conducted using a cross sectional study design. This design was deemed appropriate for this study as the aims of the study require data to be collected from typically developing children across different age bands from one point in time and it has been used in other similar studies of normative data collection (Conti-Ramsden *et al* 2011; Dodd *et al* 2003). The participants were assessed using two assessments; a verb comprehension assessment, the Test of Verb Knowledge and a story retell tool designed to elicit targeted verbs and their argument structures as well as sentence complexity; Captain Grey and the Greedy Aliens.

2.5 Instrument Development

2.5.1 Captain Grey and the Greedy Aliens

The story retell tool 'Captain Grey and the Greedy Aliens' was developed as a method of assessing verb argument structure production and sentence complexity of CwSLI. The assessment was developed as a story retell tool, a constrained task which allows for control over the verbs elicited. A story retell task also provides a more representative sample of what the child is likely to produce in a spontaneous speech sample than would be provided by a de-contextualised structured task such as the Clinical Evaluation of Language Fundamentals (CELF-4) (Epstein and Phillips 2009; Gazella & Stockman 2003).

The verbs included in the story retell tool were chosen to capture a variety of verb types; transitive, optionally transitive, intransitive and di-transitive verbs. Different semantic-syntactic classes of verbs were also covered, such as change of state, change of location, manner of motion, direct motion, verbs of communication and cognitive state verbs. Verbs were both high and low frequency based on the BNC spoken corpus (Leech *et al* 2001) (See Appendix 1 for a full list of target verbs).

The story used in the assessment was written to have a simple transparent plot which was accompanied by a wordless picture book with colourful drawings depicting different scenes of the story.

2.5.2 Test of Verb Knowledge

The Test of Verb Knowledge is a verb comprehension assessment encompassing 36 test verbs and three trial items. The assessment is a picture pointing task; the child must select the correct verb from a choice of four pictures depicting different actions with two related and one unrelated distractors.

2.6 Procedure

Each child was assessed individually. They were seen in their school, in a quiet area outside of the classroom. Two testers were present for each assessment; one administered the assessments while the other recorded responses.

2.6.1 Test of Verb Knowledge

The Test of Verb Knowledge (TOVK) was administered first as it required no verbal response and therefore placed no pressure on the child to verbalise in the initial part of

the session allowing them to relax and become familiar with the testers. Each participant's responses were calculated for the total number of items answered correctly.

2.6.2 Captain Grey and the Greedy Aliens

Captain Grey and the Greedy Aliens, was administered second, administration was as follows;

1. The tester told the child the story using the script provided to ensure consistency and fidelity across administrators while the child looked at the wordless picture book depicting the story.
2. The child then retold the story with the use of the picture book, one tester recorded the responses. Responses were also audio-recorded.
3. Any target verb not spontaneously produced was prompted by the tester as follows:
 - a. Directing the child's attention to the relevant scene and asking the child the following: "Can you tell me about this picture?"
 - b. If this was unsuccessful the child would be given the target verb and asked: "Can you make a sentence about this picture with the word ' ____ '?"

Each assessment was audio recorded using a Dictaphone. As soon as possible after each assessment the audio recordings were uploaded to a password protected laptop and deleted from the Dictaphone. Audio recordings were kept on the password protected laptop for the duration of the study. Once responses were transcribed they were scored for the following; verb and verb argument structure accuracy and mean sentence complexity.

2.6.2.1 Verb Argument Structure Accuracy

The verb argument structure production of the target verbs was scored either correct or incorrect, scoring 1 or 0 respectively. An incorrect score was given if the participant omitted an obligatory argument, produced a mapping error or an argument ordering error. An exception to the omission of an obligatory argument was on some occasions when the subject argument was omitted, where the subject was judged to be implied its omission was not counted as an error (Thordardottir and Weismer 2002). If the participant produced the verb as a noun e.g. 'email' or misunderstood the meaning of the

verb or a verb was inadvertently not elicited due to tester omission it was omitted from the scoring. The total number of verbs produced with the correct argument structure by each participant was calculated as a percentage of the total verbs elicited.

Examples of verb argument structures errors:	
Omission of obligatory argument	'The alien cleared the food', '
Mapping error	'Captain Grey filled the petrol into the spaceship
Examples of omissions from scoring:	
Verb produced as a noun (target: email)	'He sent an email to the soldiers'
Misinterpretation (target: bet)	'Captain Grey beat the aliens'

Table 2 Examples of errors and omissions from the scoring of verb argument structure accuracy

2.6.2.2 Mean Sentence Complexity

To calculate Mean Sentence Complexity each sentence was given a score from 1-6 for complexity (see Table 2.); the total score for all sentences was calculated and divided by the number of sentences produced. The rubric for mean sentence complexity was developed by Murphy (2013) based on a previous study by Webster *et al* (2007) and the literature on developmental progression in sentence complexity in children.

Score	Example
1-3	One to three arguments/adjuncts with verb
4	2 arguments + 2 adjuncts; 3 arguments +adjunct; simple infinitive sentences and simple conjoining
5	Subordination/embedding, unmarked infinitives , relative clauses
6	Conjoining and embedding occurring in the same utterance or multiple embedding

Table 3 Mean Sentence Complexity scoring rubric

2.6.3 Verb Error Analysis

The number of times each target verb was produced in error across the entire data set of children, was recorded to allow for the analysis of the difficulty of different verbs and highlight those that were less and more problematic.

2.7 Inter-rater Reliability

To examine the reliability of scoring agreement for Captain Grey and the Greedy Aliens, 20% of the transcripts from younger participants and 10% of the transcripts from older

participants were independently analysed by two scorers. Inter-rater reliability was calculated using the equation $(X/Y) \times 100$ where X and Y were the scores obtained by the two scorers, with X being the smaller value.

3.0 Results

The following variables were analysed using IBM SPSS Statistics 20; the percentage of verb argument structure accuracy (VAS) and mean sentence complexity (MSC) as gathered from Captain Grey and the Greedy Aliens and the Test of Verb Knowledge scores (TOVK).

Variable	N	Mean	Standard Deviation	5% Trim. Mean	Range	Skewness	Test of Normality
VAS	91	92.22	10.67	93.73	27-100	-3.321	.000
MSC	91	2.83	.425	2.85	1.33-3.61	-.705	.012
TOVK	91	29.35	3.40	29.50	21-36	-.608	.001

Table 4 Analysis of the scores for Verb argument structure accuracy, Mean Sentence Complexity and the Test of Verb Knowledge for the entire data set

3.1 Verb Argument Structure Production Accuracy

Verb argument structure accuracy (Md= 95.83, IQR: 88.8, 97.92) scores did not follow normal distribution and violated the Shapiro-Wilks Test of normality ($p = .000$), strong negative skew is apparent from inspection of the Histogram (Figure 1a). A number of outliers were present in the data with three outliers more than three box lengths from the edge of the box (Figure 1b), as the Mean and 5% Trimmed Mean are similar the cases were retained.

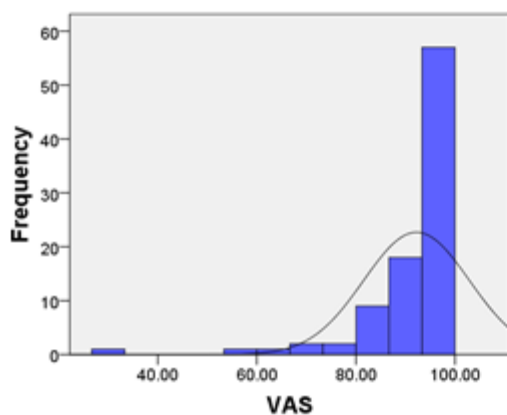


Figure 1a Distribution of raw scores for Verb argument

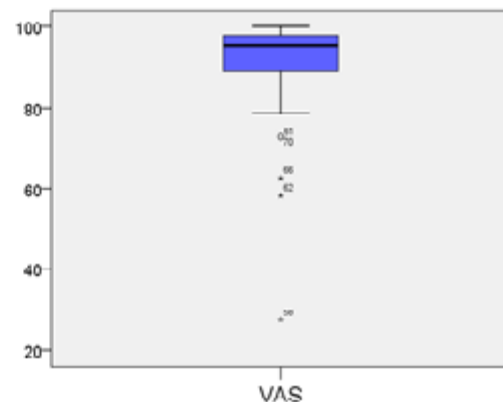


Figure 2b Boxplot of raw scores for Verb argument

3.1.1 Verb argument structure accuracy scores according to Age Bands

The participants were divided into age bands and argument structure accuracy (VAS) analysed for each Age group (see table 5). Mean scores increased for each age range from; (4;06-5;05) to (9;06-10;05) with the oldest age range (10;06-12;00) showing a decrease. The range of scores decreased as ages increased from the ages (4;06-5;05) to (7;06-8;05), the later having the lowest range of scores across all ages.

Age Range	N	Mean	SD	Range	CI for Mean	Median	% at Ceiling
4;06-5;05	7	74.15	24.68	27.1-95.74	51.32-96.96	80.43	None
5;06-6;05	17	87.65	9.25	62.5-100	82.89-92.40	89.58	5.9
6;06-7;05	17	91.20	7.53	78.72-100	88.32-96.06	95.83	17.6
7;06-8;05	9	95.51	3.61	91.11-100	92.74-98.28	95.83	22.2
8;06-9;05	9	95.59	5.15	83.33-100	91.63-99.55	97.91	11.1
9;06-10;05	12	97.65	3.24	88.63-100	95.58-99.71	97.91	41.66
10;06-12;00	20	96.19	4.41	86.95-100	94.12-98.25	97.91	35.0

Table 5 Mean scores for Verb argument structure according to age groups.

Some early developmental progression of VAS scores with age can be seen in the scatter plot (Figure 2). Due to non-normal distributions, the relationship between Verb argument structure accuracy and age was investigated using Spearman's rho; there was a strong positive correlation ($r=.53$, $p<.001$, 2 tailed) between the two variables suggesting that the test reflects developmental progression

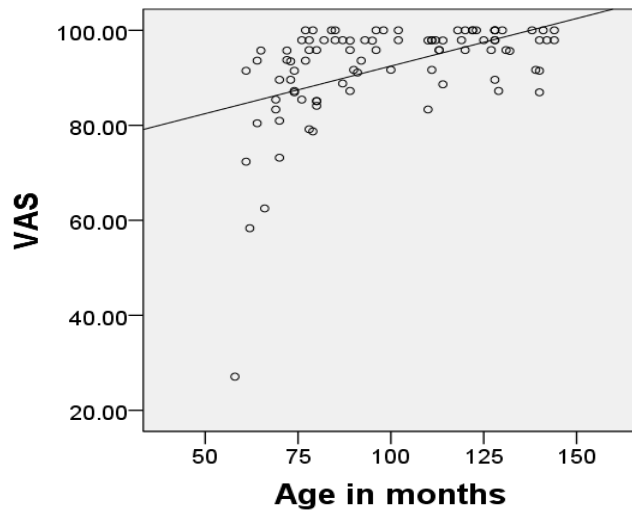


Figure 2 Scatter plot of raw verb argument structure scores

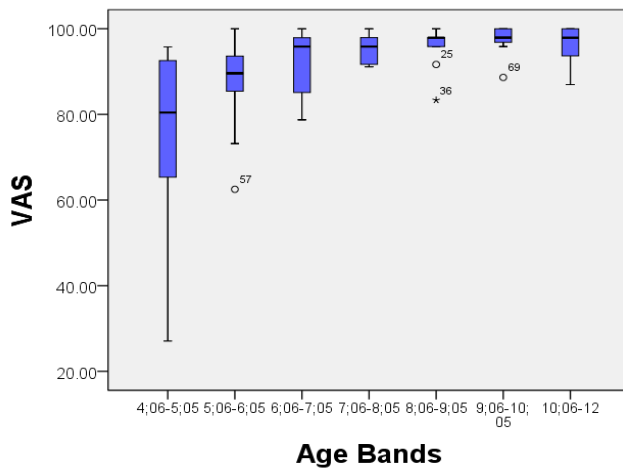


Figure 3 Boxplot of raw Verb argument structure scores according to age bands

Ceiling scores of 100% accuracy was achieved by 20.9% of participants with children in the age range (10;06-12;00) accounting for the largest of this percentage (7.7%). No floor effects were observed as all participants engaged in the task and no child obtained a score of 0. There was an increase in the percentage of participants at ceiling within each group with children at ceiling from the (5;06-6;05) age range onward.

3.1.2 Verb Knowledge as a predictor of Verb Argument Structure accuracy

Partial Correlation was used to explore the relationship between verb knowledge (as assessed using the Test of Verb Knowledge) and verb argument structure accuracy while controlling for the influence of age. The strength of the relationship between verb knowledge and verb argument structure decreased from a large strength positive correlation ($r=.55$, $n=91$, $p<0.001$) when age was not controlled for to a medium strength positive correlation ($r=.37$, $n=91$, $p<0.001$).

Standard Multiple Regression was used to explore how much of the variance in the dependent measure (verb argument structure accuracy) can be explained by verb knowledge (TOVK) and Age. One outlier, with a Cook's Distance value exceeding 1 was detected in this analysis resulting in this outlier being removed and the test rerun.

Using the enter method with verb knowledge and age as the predictor variables, the model which emerged was significant at the 5% level: $F(2, 87) = 20.079, p < .001$, explaining 30% of the variance (Adjusted R squared = .300). Table 6 provides information regarding the different predictor variables entered into the model. Verb Knowledge made the strongest unique contribution according to the Beta Values (TOVK=.36 and Age=.26).

Variables	VAS	Age	TOVK	Unstandardized coefficients		Standardized Coefficients
				B	Stand. Error	Beta
Age	.490		.642	.084	.037	.260**
TOVK	.525	.642		.888	.286	.359**

Table 6 Standard Multiple Regression of age and verb knowledge (TOVK) on Verb argument structure accuracy
** $P < 0.05$

3.2 Mean Sentence Complexity

Mean sentence complexity (MSC) (Md=2.83, IQR: 2.57, 3.2) scores did not follow normal distribution and violated the Shapiro-Wilks Test of normality (Sig. = .012), negative skew is apparent from inspection of the histogram (fig. 4). An outlier was present in this data set. As the Mean and 5% Trimmed Mean were similar the case was retained. Due to the skew of the data non-parametric tests were utilised.

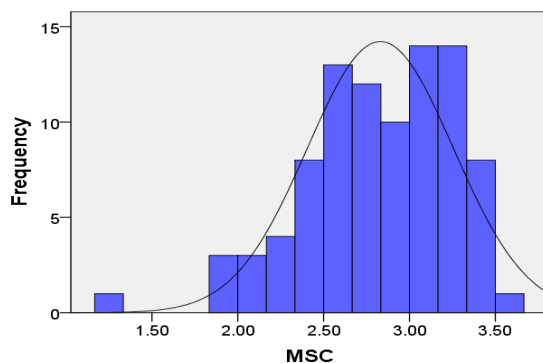


Figure 4a Distribution of mean MSC scores

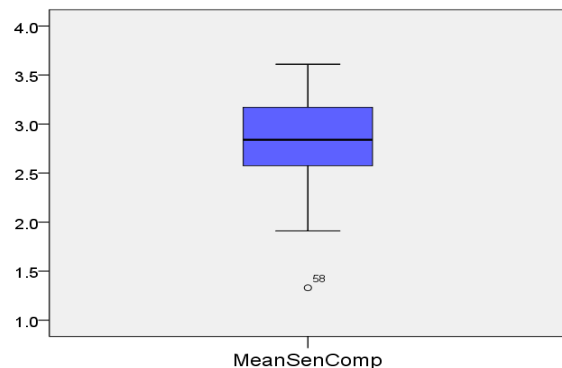


Figure 4b Boxplot of mean MSC scores

3.2.1 MSC according to Age Bands

The mean sentence complexity scores according to age ranges were analysed, mean scores increased from the age groups (4;06-5;05) to (8;06-9;05). The mean scores for age groups (9;06-10;05) and (10;06-12;00) were lower than the (8;06-9;05) age group. A number of outliers are present both in the scatter plot and in the box plot for Age Bands (Fig. 5a, Fig. 5b). The relationship between MSC and Age was investigated using Spearman's rho. There was a strong positive correlation ($r=.65$, $p<.001$, 2 tailed) between the two variables.

Age Range	N	Mean	SD	Range	CI for Mean	5%Trim. mean	Median
4;06-5;05	7	2.28	.61	1.33-3.13	1.71-2.84	2.28	2.1
5;06-6;05	17	2.50	.28	2.0-3.02	2.35-2.64	2.50	2.5
6;06-7;05	17	2.67	.31	1.91-3.29	2.51-2.82	2.68	2.69
7;06-8;05	9	2.97	.42	2.47-3.41	2.64-3.28	2.97	3.14
8;06-9;05	9	3.25	.26	2.83-3.61	3.04-3.44	3.25	3.25
9;06-10;05	12	3.04	.24	2.6-3.29	2.88-3.19	3.05	3.06
10;06-12;00	20	3.07	.24	2.59-3.45	2.97-3.17	3.08	3.09

Table 7 Mean scores for Mean Sentence Complexity according to age groups

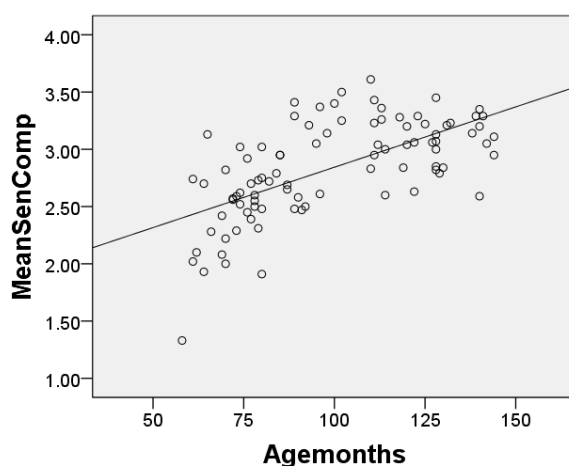


Figure 5a Scatter plot of MSC raw scores and Age

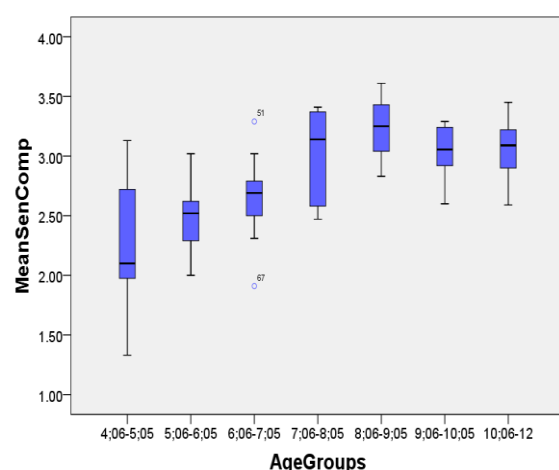


Figure 5b Boxplot of MSC according to Age

Due to the nature of Mean Sentence Complexity there are no ceiling scores to report. However 42.9% of participants scored at or above 3.0. The oldest age band (10;06-12;00) accounted for 15.4% of the participants who scored at or above 3.0, (9;06-10;05); 9.9%, (8;06-9;05); 7.7%, (7;06-8;05) 5.5%, (6;06-7;05); 2.2%, (5;06-6;05); 1.1% and (4;06-5;05); 1.1%. The higher mean sentence complexity scores achieved by the older participants is due to their greater use of complex sentences for example: “Then the soldier opened his eyes and told Captain Grey the story about how the aliens had stolen his lunch”.

3.2.2 Verb Knowledge as a predictor of Mean Sentence Complexity

Partial Correlation was used to explore the relationship between verb knowledge (as assessed using the Test of Verb Knowledge) and Mean Sentence Complexity while controlling for Age. The strength of the relationship between verb knowledge and Mean Sentence Complexity decreased from a large strength positive correlation ($r=.62$, $n=91$, $p<0.001$) when age was not controlled for to a medium strength positive correlation ($r=.35$, $n=91$, $p=.001$) when age was controlled.

Standard Multiple Regression was used to explore how much of the variance in the dependent measure (mean sentence complexity) can be explained by verb knowledge and Age.

Using the enter method with verb knowledge (TOVK) and age as the predictor variables, the model which emerged was significant at the 5% level: $F(2, 88) = 39.449$, $p<.001$, explaining 46.1% of the variance (Adjusted R squared = .461). Table 8 provides information regarding the different predictor variables entered into the model. Age made the strongest unique contribution according to the Beta Values with Verb Knowledge contributing slightly less.

Variables	MSC	Age	TOVK	Unstandardized coefficients		Standardized Coefficients
				B	Stand. Error	Beta
Age	.631		.655	.007	.002	.394**
TOVK	.620	.655		.045	.013	.362**

Table 8 Standard Multiple Regression of age and verb knowledge (TOVK) on Mean Sentence Complexity. $p<.05$

3.3 Standardisation of the Test Data

The scores for Verb Argument Structure accuracy and Mean Sentence Complexity were standardised by transforming the variables to z-scores using SPSS producing each participants (n=91) standard deviation from the mean score for each variable. From visual inspection the z-scores for Mean Sentence Complexity and Verb argument structure accuracy became more positive as age increased, indicating that both Mean Sentence Complexity and Verb Argument Structure accuracy increase with age.

The mean and standard deviation scores (SD) were also calculated for both Verb Argument Structure accuracy and Mean Sentence Complexity within each age group;

Age Bands	-3 SD	-2 SD	-1.5 SD	-1 SD	Mean	+1 SD	+1.5 SD	+2 SD
4;06-5;05	0.11	24.79	37.13	49.47	74.15	98.83	100	100
5;06-6;05	59.9	69.15	73.78	78.4	87.65	96.90	100	100
6;06-7;05	68.61	76.14	79.91	83.67	91.2	98.73	100	100
7;06-8;05	84.68	88.29	90.10	91.9	95.51	99.12	100	100
8;06-9;05	80.14	85.29	87.87	90.44	95.59	100	100	100
9;06-10;05	87.93	91.17	92.79	94.41	97.65	100	100	100
10;06-12;00	82.96	87.37	89.58	91.78	96.19	100	100	100

Table 9 VAS percentage correct raw scores by standard deviations from the mean for each age group

Age Bands	-3 SD	-2 SD	-1.5 SD	-1 SD	Mean	+1 SD	+1.5 SD	+2 SD	+3 SD
4;06-5;05	0.45	1.06	1.37	1.67	2.28	3.5	3.81	4.11	4.72
5;06-6;05	1.66	1.94	2.08	2.22	2.5	2.78	2.89	3.06	3.34
6;06-7;05	1.74	2.05	2.21	2.36	2.67	2.98	3.14	3.29	3.6
7;06-8;05	1.71	2.13	2.34	2.55	2.97	3.39	3.6	3.81	4.23
8;06-9;05	2.47	2.73	2.86	2.99	3.25	3.51	3.64	3.77	4.03
9;06-10;05	2.32	2.56	2.68	2.8	3.04	3.28	3.4	3.52	3.76
10;06-12;00	2.35	2.59	2.71	2.83	3.07	3.31	3.43	3.55	3.79

Table 10 MSC raw scores by standard deviation from the mean for each age group

These tables are the normative scores for standard deviations from the mean of both Verb argument structure accuracy and Mean sentence complexity from the results of the data collected in this study (see Appendix 2 for the equivalent in percentile ranks for each age group). These tables, (subject to any future modifications of the tool) could be used to identify where future children’s results place them in relation to the scores of children in this study for example if a child of 6;09 scored 77.5% for VAS they would be between -1.5 and -2 SD’s from the mean and between the 25th and 10th percentile.

3.4 Error Analysis

The total number of times each verb was produced in error was analysed to determine which verbs were most problematic. The five most problematic verbs were; clear (24), pour (21), fill (17), bet (17) and told (12) (See Appendix 2 for the entire list of verbs and their number of errors). “Clear”, “pour” and “fill” were all targeted as di transitive verbs requiring three arguments in the retell narrative. The verbs “pour” and “fill” are reverse role verbs “Bet” and “tell” are two cognitive state verbs requiring a sentence complement with two arguments.

Verb	Most frequent error	Example
Clear	Obligatory argument omission	the aliens cleared the food
Pour	Obligatory argument omission	Captain Grey poured water
Fill	Mapping error	Captain Grey filled the petrol into the ship
Bet	Misunderstood meaning as the verb ‘to beat’	Captain Grey beat the aliens
Told	Obligatory argument omission	The soldier told about the aliens

Table 11 Examples of Verb Errors

3.5 Inter-rater reliability

The inter-rater scoring reliability for both VAS and MSC was found to be high, with 97.34% and 95.54% agreement respectively as calculated on 20% of the transcripts from younger children (4;06-8;00) and 10% of older children (8;01-12;00). The scoring agreement for Test of Verb Knowledge was 100%. Items with scores which did not agree were discussed with the PI to agree scoring and ensure consistency across the samples.

4.0 Discussion

Normative data was gathered from TD children aged 4;06-12;00 using the story retell tool 'Captain Grey and the Greedy Aliens'. The data gathered was used to analyse children's production of the target verbs and their associated argument structures and sentence complexity to identify if the assessment was suitably sensitive to capture developmental progression. Verb comprehension was also assessed using the Test of Verb Knowledge to investigate the relationship between verb comprehension and verb argument structure accuracy and sentence complexity. The inter-rater reliability of the story retell assessment was also examined and found to be highly reliable for verb argument structure and Mean Sentence Complexity; 97.34% and 95.54% respectively.

4.1 Verbs and Verb Argument Structure

The findings of this study show that the verb argument structure accuracy increases with age. Mean scores across the age groups increased until the (9;06-10;05) with a decrease shown for the oldest age group (10;06-12;00). Statistical analysis indicated a positive correlation between Verb argument structure accuracy and age. The findings are supported by evidence that children produce less verb argument structure errors as they get older (Ambridge *et al* 2008, Pinker 2013). The set of standardised scores generated for this measure could be used as a developmental index for other children but with some caveats. The mean score for the oldest age group (10;06-12;00) being lower than the preceding group (9;06-10;05) was unexpected. Rather than lower scores representing a decline in verb argument structure accuracy for this age group the finding may reflect a sampling error. As the number in this age group was small the sample of participants may be under representative of this population as a whole with an unusually high proportion of low scores.

Verb argument structure accuracy was high across all age bands. The majority of participants scored at 80% or above for accuracy with only 7 participants scoring lower; of whom were within or below the (6;06-7;05) age group. This finding indicates that children develop a high degree of verb argument structure accuracy at a relatively young age.

Although verb accuracy is relatively high it is apparent that some verbs and verb types continue to cause difficulty into late childhood as only 35% of participants in the oldest

age group (10;06-12;00) reached ceiling. This is consistent with research by Ebbels (2012) who found that TD children (5;04-14;10) continued to make verb argument structure errors for change-of-state verbs even in the older age ranges.

Overall the small sample size may impact on the reliability of the mean scores and reduce the power of the assessment as an index of developmental progression (Paul and Norbury 2012). Larger sample sizes are required.

4.2 Mean Sentence Complexity

The complexity of sentences produced by the participants during their retelling of Captain Grey and the Greedy Aliens was assessed by calculating their Mean Sentence Complexity. As expected the mean scores for Mean Sentence complexity increased positively with the age. The Mean scores for Mean Sentence Complexity increased from the (4;06-5;05) to the (8;06-9;05) age group providing evidence of developmental progression in sentence complexity. The increase in Mean Sentence Complexity scores is expected as TD children produce a higher percentage of complex sentences per total number of sentences as they age (Paul and Norbury 2012; Bowerman 1979). The mean scores for Mean Sentence Complexity was unexpectedly lower for the two oldest age groups; (9;06-10;05) and (10;06-12;00) compared to the (8;06-9;05) age group. Similar to the findings in verb argument structure accuracy the low scores for these age groups may reflect a sampling error. The numbers in these age groups were small and therefore the sample of participants may be under representative of this population as a whole. Another possible explanation is that some participants may not be performing to the best of their ability; due to the nature of the assessment complex language is not required for the task to be completed (Hesketh 2004; Hadley 1998).

Age was found to make the largest contribution to the variation of scores for Mean Sentence Complexity when compared to verb knowledge. This provides further evidence that Mean sentence complexity is appropriately sensitive to the developmental progression of sentence complexity and supports the use of Mean sentence complexity mean scores to produce a developmental index. However these findings should be interpreted with caution due to the small sample size which may negatively impact on the reliability of the mean scores and reduce the power of the assessment as an index of developmental progression (Paul and Norbury 2012). Larger sample sizes are required.

4.3 Verb Error Analysis

As expected some verbs had higher rates for accuracy of argument structure production than others, indicating that they are mastered earlier than other more complex verbs. Only one target verb “fall” was never produced in error indicating that all children have mastered this verb. Fall is an intransitive verb of high frequency (Leech *et al* 2001) and belongs to the semantic-syntactic class of ‘manner of movement’. Other intransitive, manner of movement verbs targeted in the assessment were; fly, jump and crawl, these also had low total errors indicating that intransitive verbs of manner of motion are mastered in early childhood.

Some verbs had high error rates; clear, fill, pour, bet and tell indicating that they are more difficult to master. The verb produced in error most frequently; clear (24 errors) may be a di-transitive verb in the context due to with complex mapping between the verb semantics and sentence structure. Hence if the child produced “the aliens cleared the food”, it would be incorrect as the sentence requires the source to be specified in this instance, and on the other hand “they cleared the shelves” is correct. The majority of errors for this verb were due to obligatory argument omissions when food occupied the sentence object position and was the patient of the verb. Mapping errors or omissions were also common for the verb “fill” which is also a di-transitive verb. The majority of errors were due to argument mapping errors for example “Captain Grey filled the petrol into the spaceship”. Mapping errors for “fill” appear to be due to the overgeneralisation of the argument structure for “pour”. “Fill” and “pour” are reverse role verbs as they describe the same event but from different thematic perspectives (Marshall 2002) which can influence the occurrence of overgeneralisations. Similar to “clear”, “pour” may be a di-transitive verb in the context due to with complex mapping between the verb semantics and sentence structure. Errors were due to obligatory argument omissions for example “poured the water into bottles” which may have been due to the complexity of the verb increasing capacity demands resulting in high frequency of omission (Grela and Leonard 2000). “Bet” and “tell” are both cognitive state verbs that take sentence complements and require two argument structures. Difficulty with cognitive state verbs is in keeping with the literature which cites cognitive state verbs are more difficult to acquire (Owen van Horne and Lin 2011, Diessel and Tomasello, 2001, Nixon 2005). All five

verbs were low frequency which may have contributed to the high error rate according to the findings of Ambridge *et al* (2008) and Brooks *et al* (2008).

4.4 Verb Knowledge as a predictor of Verb argument Production accuracy and Sentence Complexity

This study also explored the relationship between stored verb knowledge and sentence complexity and verb production. Stored verb knowledge was assessed via the Test of Verb Knowledge.

Verb knowledge had a medium strength positive correlation with verb argument structure accuracy when the effects of age were controlled for and was found to be the largest contributor to variance in verb argument structure accuracy when compared to age. Indicating that stored verb knowledge can be used to predict verb argument structure accuracy. The relationship between verb knowledge and accurate production of verb argument structures is also evident in the results of Ebbels *et al's* (2012) study on verb argument structure production, CwSLI were compared to Grammar-matched (Test for Reception of Grammar, TROG) and Age-matched controls, Age-matched controls scored higher on the TROG than Grammar-matched and also achieved higher verb argument structure accuracy. In light of the indication of poor lexical verb knowledge in children with SLI (Windfuhr *et al* 2002), this aspect of their development is clinically important. Where a child does poorly in sentence production with respect to verb argument structure accuracy, further assessment of verb semantic knowledge may be warranted. Alternatively, where a child does poorly on an assessment of verb semantic knowledge, they may be considered at risk for verb argument structure production difficulties and intervention may need to be directed at this area. Intervention targeted at this area may be focused on increasing knowledge of verb semantic features as therapy targeted on this area was found in a study by Ebbels *et al* (2007) to be effective in increasing verb argument structure accuracy of CwSLI (11;00-16;01).

Verb knowledge was also found to have a medium strength positive correlation with Mean Sentence Complexity when the effects of age were controlled for, however unlike Verb Argument production although verb knowledge was a significant contributor age was the largest contributor to the variance in Mean sentence complexity. Therefore age is a better predictor of sentence complexity than stored verb knowledge. This is not

surprising given that the Mean Sentence Complexity scoring rubric awards marks for complex sentences e.g. use of embedding, subordination and conjoining. Development in the use of complex sentences would call on a wide range of abilities and not just verb knowledge.

4.5 Test Related Behaviours

Test related behaviours are import factors which may have affected the reliability of the findings and also may have implications for future development and clinical utility (Bishop 2014). Some participants required additional explanation of the task and examples of target verb use to comprehend what was expected of them when given the 2nd prompt (child is provided with the target word to put in a sentence). A number of children made a sentence unrelated to the scene for example 'The baby crawled'. One participant began all sentences using second prompts using the subject 'I' but related the sentence to the scene for example 'I captured the aliens'. Some participants presented as being shy and required extra encouragement and prompting for them to retell the story. Although some children overcame this quickly it may have affected their responses and resulted in an inaccurate representation of their true language abilities.

4.6 Considerations

In the process of administering the assessment a number of assessment procedure related findings were identified which may have future implications on clinical utility and future development.

The length of the story was found to affect younger participant's attention to the task. The length of the story resulted in some younger participants requiring greater prompting in the initial retelling of the story. Difficulty in attending to the task was anticipated for the younger children due to lower levels of attention, memory and retrieval (Gazella and Stockman 2003). Splitting the story and telling half of it at a time may be effective in reducing the demands placed on younger children however this may result in longer administration time.

The length of time required for administration varied from approximately 20-50 minutes depending on the degree of prompting required. In general greater prompting was required for younger participants which may have impacted on their results due to the

aforementioned affect the length of the assessment had on their attention to task. This may be a limitation in the clinical utility of the assessment.

4.7 Limitations

As mentioned throughout this study is limited by its small sample size. The small sample affects the ability of the mean scores for verb argument structure and mean sentence complexity to be reliably used to formulate an index of developmental progression. Further normative sampling with a larger sample size is required for the assessment to be sufficiently standardised.

Another limitation is the lack of administration of a comparable assessment to the story retell tool. Administration of a comparable assessment would have increased the validity of the story retell assessment by ensuring the assessment was accurately eliciting and assessing verb argument structure and sentence complexity of the participants (Bishop 2014). This limitation is the result of a lack of comparable appropriate assessment for this area of language for children.

4.8 Recommendations

Further normative sampling of the story retell assessment is required for to increase the power of the study results (e.g. 100 participants per age group; Paul and Norbury 2012) and for the assessment to be sufficiently standardised.

Further research may be conducted into whether the initial story retell without prompting is a sufficiently assessment of verb argument structure and mean sentence complexity. Ability to administer the initial story retell independent of the prompts would increase clinical utility allowing clinicians to administer only part of the assessment if required due to time constraints or participant participation issues.

The level of prompting required is also an area which may benefit from future research. It was qualitatively noted that younger participants required longer administration time due to prompting.

4.9 Clinical Implications and conclusions

The current research provides further evidence that verb knowledge is strongly associated with verb argument structure with greater verb knowledge resulting in fewer argument structure errors. To a lesser extent this research also reveals a positive

relationship between verb knowledge and sentence complexity. Both of these findings are clinically important for language intervention. The story retell tool 'Captain Grey and the Greedy Aliens' has shown to be an effective method of eliciting and analysing a variety of verb and verb argument structure productions as well as sentence complexity. The findings support the sensitivity of the assessment to capture the developmental progression of verb and verb argument structure production and also sentence complexity of typically developing children from the ages of 4;06 to 12;00 years. Although further normative sampling is required the findings suggest that the assessment can be a useful tool in analysing verb production and sentence complexity in children with specific language impairment.

Reference List

- Ambridge, B., Pine, J.M., Rowland, C.F., Young, C.R. (2008) 'The effects of verb semantic class and verb frequency (entrenchment) on children's and adults' graded judgements of argument-structure overgeneralization errors', *Cognition*, 106(1), 87-129.
- Bishop, D.V.M. (2014) *Uncommon Understanding: Development and disorders of language comprehension in children*, Classic ed., Sussex: Psychology Press.
- Bowerman, M. (1979) 'The acquisition of complex sentences' in Garman, M. and Fletcher, P., (Eds.), *Studies in language acquisition*, Cambridge: Cambridge University Press, 285-305.
- Brooks, P. J., Tomasello, M., Dodson, K., & Lewis, L. B. (1999) 'Young children's overgeneralizations with fixed transitivity verbs', *Child Development*, 70(6), 1325-1337.
- Chiat, S., Law, J., and Marshall, J. (1997) *Language Disorders in children and adults*, London: Whurr Publishers
- Conti-Ramsden, G., and Jones, M. (1997). 'Verb use in specific language impairment', *Journal of Speech, Language and Hearing Research*, 40(6), 1298-1313.
- Conti- Ramsden, G., Durkin, K., Simkin, Z., Lum, J.A.G., and Marchman, V. (2011) 'The PTT-20: UK normative data for 5- to 11-year olds on a 20-item past-tense task', *International Journal of Language and Communication Disorders*, 46(2), 243-248.

- Dodd, B., Holm, A., Hue, Z., and Crosbie, S. (2003) 'Phonological development: a normative study of British English-speaking children', *Clinical Linguistics and Phonetics*, 17(8), 617-643.
- Diessel, H. and Tomasello, M. (2001) 'The acquisition of finite complement clauses in English: A corpus based analysis', *Cognitive Linguistics*, 12(2), 97-141.
- Ebbels, S.H., Van der Lely, H.K., and Dockrell, J.E. (2007) 'Intervention for Verb Argument Structure in Children with persistent SLI: A randomized control trial', *Journal of Speech, Language and Hearing Research*, 50(5), 1330-1349.
- Ebbels, S.H., Dockrell, J.E., and Van der Lely, H.K. (2012) 'Production of change-of-state, change-of-location and alternating verbs: A comparison of children with specific language impairment and typically developing children', *Language and Cognitive Processes*, 27(9), 1312-1333.
- Edwards, S., Letts, C., and Sinka, I. (2011) *The New Renell Developmental Language Scales, 4th editions*, England: Nfer-Nelson.
- Epstein, S.A., and Phillips, J. (2009) 'Storytelling skills of children with specific language impairment', *Child Language and Teaching Therapy*, 25(3), 285-300.
- Fenson, L., Dale, P., Reznick, J. S., Bates, E., Thal, D. J. and Pethick, S. (1994) 'Variability in early communicative development', *Monographs of the Society for Research in Child Development*, 59(5), 1-185.
- Fletcher, P. (2009) 'Syntax in Child Language Disorders', in Schwartz, R.G., *Handbook of Child Language Disorders*, Hove: Psychology Press, 388-405.
- Gazella, J., and Stockman, I.J. (2003) 'Children's story retelling under different modality and task conditions: Implications for standardizing language sample procedures', *American Journal of Speech and Language Pathology*, 12(1), 61-72.

- Gentner, D. (1978) 'On relational meaning: The acquisition of verb meaning', *Child development*, 49(4), 988-998.
- Goldfield, B.A. (2000) 'Nouns before verbs in comprehension vs. production: the view from pragmatics', *Journal of Child Language*, 27(3), 501-520.
- Grela, B. G. (2003) 'The omission of subject arguments in children with specific language impairment', *Clinical linguistics & phonetics*, 17(2), 153-169.
- Grela, B.G., and Leonard, L. (2000) 'The influence of argument-structure complexity on the use of auxiliary verbs by children with SLI', *Journal of Speech, Language and Hearing Research*, 43(5), 1115-1125.
- Hadley, P. (1998) 'Language sampling protocols for eliciting text-level discourse', *Language, Speech, and Hearing Services in Schools*, 29(3), 132-147.
- Hesketh, A. (2004) 'Grammatical performance of children with language disorders on structure elicitation and narrative tasks', *Clinical Linguistics and Phonetics*, 18(3), 161-182.
- IASLT (2007) 'Specific Speech and Language Impairment in Children: Definition, Service Provision and Recommendations for Change, available: <http://www.iaslt.ie/docs/public/exec/IASLT%20SSLI%20Position%20Paper%20Oct%202007.pdf> [accessed 10 November 2013].
- Johnston, J.R., Miller, J., and Tallal, P. (2001) 'Use of cognitive state predicates by language-impaired children', *International Journal of Language and Communication Disorders*, 36(3), 349-370.
- King, G., and Fletcher, P. (1993) 'Grammatical problems in school-age children with Specific Language Impairment', *Clinical Linguistics and Phonetics*, 7(4), 339-352.
- Leech, G., Rayson, P. and Wilson, A. (2001) *Word Frequencies in Written and Spoken English: Based on the British National Corpus*, London: Longman.

- Leonard, L., Miller, C., and Gerber, E. (1999) 'Grammatical morphology and the lexicon in children with specific language impairment', *Journal of Speech, Language and Hearing Research*, 42 (3), 678-689.
- Loeb, D.F., Pye, C., Redmond, S., and Richardson, L.Z. (1996) 'Eliciting Verbs from Children with Specific Language Impairment', *American Journal of Speech-Language Pathology*, 5(4), 17-30.
- Leonard, L.B. (2000) *Children with Specific Language Impairment*, London: MIT Press.
- Levy, H., and Friedmann, N. (2009) 'Treatment of syntactic movement in syntactic SLI: A case study', *First Language*, 29(1), 15-50.
- Marinellie, S.A. (2004) 'Complex syntax used by school-aged children with specific language impairment (SLI) in child-adult conversation', *Journal of Communication Disorders*, 37(6), 517-533.
- Marshall, J. (2002) 'Assessment and Treatment of Sentence Processing Disorders: A Review of the Literature', in Hillis, A.E., ed., *The Handbook of Adult Language Disorders: Integrating Cognitive Neuropsychology, Neurology and Rehabilitation*, London: Psychology Press, 351-374.
- McGregor, K.K. (2009) 'Semantics in Child Language Disorders', in Schwartz, R.G. *Handbook of Child Language Disorders*, Hove: Psychology Press, 365-387.
- McKoon, G., and MacFarland, T. (2000) 'Externally and Internally cause Change of State Verbs', *Language*, 76(4), 833-858.
- Miller, J.F. (1981) *Assessing Language Production in Children*, London: Edward Arnold
- Murphy, C.A. (2013) *Profiles and Characteristics of Sentence Production difficulties in Children with Specific Language Impairment*, PhD Thesis , School of Education, Communication and Language Sciences, Newcastle University.

- Naigles, L.R. (1990) 'Children use syntax to learn verb meaning', *Journal of Child Language*, 17(2), 357-374.
- Nils Johnson, C., and Wellman, H.M. (1980) 'Children's Developing Understanding of Mental Verbs: Remember, Know and Guess', *Child Development*, 51(4), 1095-1102.
- Nixon, S. (2005) 'Mental state verb production and sentential complements in four-year-old children', *First Language*, 25(1), 19-37.
- Owen Van Horne, A.J., and Lin, S. (2011) 'Cognitive state verbs and compliment clauses in children with SLI and their typically developing peers', *Clinical Linguistics and Phonetics*, 25(10), 881-898.
- Paul, R., and Norbury, C.F. (2012) *Language Disorders from Infancy through Adolescence: Listening, Speaking, Reading, Writing, and Communicating*, 4th ed., Missouri: Elsevier Mosby.
- Pinker, S. (2013) *Learnability and Cognition: The Acquisition of Argument Structure*, 2nd ed., Cambridge: MIT Press.
- Rice, M.L., Oetting, J.B., Marquis, J., Bode, J., and Pae, S. (1994) 'Frequency Input Effects in Word Comprehension of Children with Specific Language Impairment', *Journal of Speech and Hearing Research*, 37(1), 106-122.
- Riches, N.G., Tomasello, M., and Conti-Ramsden, G. (2005) 'Verb Learning in Children with SLI: Frequency and Spacing Effects', *Journal of Speech, Language and hearing Research*, 48(6), 1397-1411.
- Rowland, C. (2014) *Understanding Child Language Acquisition*, Oxon: Routledge
- Schuele, C., and Dykes, J. (2005) 'Complex syntax acquisition: A longitudinal case study of a child with specific language impairment', *Clinical Linguistics and Phonetics*, 19(4), 295-318.

- Schuele, C., and Nicholls, L. (2000) 'Relative clause: evidence of continued linguistic vulnerability in children with specific language impairment', *Clinical Linguistics and Phonetics*, 14(8), 563- 585.
- Schwartz, R.G. (2009) *Handbook of Child Language Disorders*, Hove: Psychology
- Semel, E., Wigg, E.H., and Secord, W. (2006) *Clinical Evaluation of Language Fundamentals*, UK, 4th Ed., United Kingdom: Pearson Education , Inc.
- Shulman, C., and Guberman, A. (2007) 'Acquisition of verb meaning through syntactic cues: A comparison of children with autism, children with specific language impairment (SLI) and children with typical language development (TLD)', *Journal of child language*, 34(2), 411-423.
- Thal, D., O'Hanlon, L., Clemmons, M., and Fralin, L. (1999) 'Validity of a parent report measure of vocabulary and syntax for preschool children with language impairment', *Journal of Speech, Language and Hearing Research*, 42(2), 482-496.
- Thordardottir, E.T., Weismer, S.E. (2002) 'Verb argument structure weakness in specific language impairment in relation to age and utterance length', *Clinical Linguistics and Phonetics*, 16(4), 233-250.
- Tomasello, M. (1992) *First Verbs: A Case Study of Early Grammatical Development*, Cambridge: Cambridge University Press.
- Van der Lely, H. K. (1994) 'Canonical linking rules: Forward versus reverse linking in normally developing and specifically language-impaired children', *Cognition*, 51(1), 29-72.
- Webster, J., Franklin, S., and Howard, D. (2007) 'An analysis of thematic and phrasal structure in people with aphasia: What more can we learn from the story of Cinderella?', *Journal of Neurolinguistics*, 20(5), 363-394.

Windfuhr, K.L., Faragher, B., and Conti-Ramsden, G. (2002) 'Lexical learning skills in young children with specific language impairment (SLI)', *International Journal of Language and Communication Disorders*, 37(4), 415-432.

Appendix 1

Captain Grey and the Greedy Aliens list of target verbs:

Live	Steal
Capture	Chase
Sweeping	Destroy
Clear	Empty
Phone	Flew
Email	Lean (x2)
Push	Run
Pour	Fell
Pick	Spread
Peel	Grate
Cover (x2)	Turn into
Fill	Give
Stuff	Land
Stay	Yawn
Take	Shake
Jump	Drop
Groan	Throw
Sink	Crawl
Pretend	Think
Tell	Shout
March	Bet
Capture	Bring
Put	

Appendix 2

Raw scores for Verb argument structure accuracy and Mean Sentence Complexity by Percentiles:

Raw scores for Verb argument structure accuracy by Percentiles (Weighted Average)							
Age Bands	5	10	25	50	75	90	95
4;06-5;05	27.0800	27.0800	58.3300	80.4300	93.6200	-	-
5;06-6;05	62.5000	71.0360	84.3700	89.5800	93.6800	98.3360	-
6;06-7;05	78.7200	79.0800	85.1000	95.8300	97.9100	100.000	-
7;06-8;05	91.1100	91.1100	91.6650	95.8300	98.600	-	-
8;06-9;05	83.3300	83.3300	93.7450	97.9100	97.9200	-	-
9;06-10;05	88.6300	90.7900	96.3275	97.9150	100.0000	100.0000	-
10;06-12;00	86.9640	87.4650	92.6575	97.9100	100.0000	100.0000	-

Raw scores for Mean Sentence Complexity by Percentiles (Weighted Average)							
Age Bands	5	10	25	50	75	90	95
4;06-5;05	1.33	1.33	1.93	2.10	2.74	-	-
5;06-6;05	2.00	2.06	2.85	2.52	2.66	2.94	-
6;06-7;05	1.91	2.23	2.49	2.69	2.87	3.07	-
7;06-8;05	2.47	2.47	2.54	3.14	3.38	-	-
8;06-9;05	2.83	2.83	2.99	3.25	3.46	-	-
9;06-10;05	2.60	2.60	2.88	3.05	3.25	3.28	-
10;06-12;00	2.60	2.79	2.87	3.09	3.22	3.34	3.44

Appendix 3

Number of Verb argument structure errors produced for each target verb

Target Verb:	Number of Total Errors:
Live	6
Steal	5
Capture (1)	6
Chase	5
Sweeping	6
Destroy	3
Clear	24
Empty	4
Phone	11
Flew	2
Email	7
Lean (1)	11
Push	2
Run (1)	2
Pour	21
Fell	0
Pick	3
Spread	9
Peel	2
Grate	4
Cover (1)	8
Turn into	5
Fill	17
Give	4
Stuff	6
Land	10
Cover (2)	6
Stay	8
Yawn	4
Lean (2)	10
Take	9
Shake	6
Jump	2
Drop	1
Run (2)	1
Groan	11
Throw	3
Sink	9
Crawl	4

Pretend	2
Think	8
Tell	12
Shout	2
March	3
Bet	17
Capture	5
Bring	4
Put	1
Total Errors	311